NETWORK STRUCTURE AND OPTIMAL TECHNOLOGICAL INNOVATION

By

Tamer Khraisha

Submitted to

Central European University

Department of Network and Data Science

In partial fulfillment of the requirements for the degree of Doctor of Philosophy in Network Science

Supervisor: János Kertész

Co-supervisor: Rosario Nunzio Mantegna

Budapest, Hungary

2019

Abstract

This thesis aims at investigating the role of inter-firm network structure in the development of technological innovations. Technological innovation is a complex economic problem, in that it admits a multidimensional landscape of potential solutions that are non-trivial to explore. To model this complexity, the approach of fitness landscapes offers a flexible and detailed framework. Furthermore, the process of technological innovation is increasingly stylized as a collective problem of interacting economic agents, and to model these interactions, network science offers useful tools and models. The analysis conducted in this thesis has two goals: first, to understand the role of network average path length, degree-heterogeneity, and edge directionality in the process of technological innovation, and second to attempt to offer empirical case study where a technological innovation was collectively developed. An evolutionary agent-based model and an applied case study from financial markets are presented and discussed. In both cases, a group of firms (in the case study, banks) collectively search a complex (rugged) technological landscape and observe each other's solutions through different relationship networks. Concerning the agent-based model, two families of networks are used in the analysis; the first family includes undirected networks which vary in terms of average path length. The second family includes directed networks which vary in terms of degree heterogeneity. Results for the agent-based model show that average path length and degree heterogeneity are important factors influencing the average performance of the system. As for the case study, I used a projected financial network that is derived from the syndicated lending database DealScan, provided by Thomson and Reuters, which I employ to produce a heterogeneous and directed network. The results obtained for the case study show an interesting trade-off between market efficiency and system stability.

Acknowledgments

My doctoral studies started under the supervision of Professor Rosario Mantegna, who was at that time affiliated with CEU. Since he left CEU, the formal supervision was taken by Professor Janos Kertesz, but I have continued collaborating with Professor Mantegna, who got this way, my co-supervisor; Therefore, I am very thankful for all the support that professor Rosario provided during my studies and in the final thesis preparation. Rosario has been very kind and helpful to provide me with extensive advice, conference and travel recommendations, in addition to hosting me for two times at his Observatory of Complex Systems at the University of Palermo in Italy where he and his colleagues showed me unparalleled hospitality.

I consider the support and guidance by professor Janos Kertesz to be equally essential in helping me complete my thesis. Despite the extraordinary amount of work Janos has to deal with daily, he would always find time for meeting and discussing any issue. I am delighted to have had the opportunity to discuss my research, get advice, and chat with Janos in office and over lunch.

With Balazs Vedres, I had several discussions about my research, asked him for recommendation letters, and thanks to his advise I managed to write and submit two of my manuscripts for publication. Thanks to the efforts of Balazs, the CEU PhD program in Network Science is the first in the world, and thanks to him I will be among the first generation PhD graduates in Network Science. I am thrilled to have learnt a lot of things from Balazs.

I would like to thanks professor Doyne Farmer from the University of Oxford for having given me the opportunity to spend three months at his Complexity Economics research group where I had the chance to meet great minds and learn about the world of complexity economics.

I am particularly thankful for the European Cooperation for Statistics of Network Data Science (COSTNET) for having provided me twice with funding to visit the University of Palermo.

Finally, I would like to thank all PhD students and post-docs at the Department of Network and Data Science with whom I had the pleasure to chat, discuss research ideas, and go for drinks.

Thank you all for your support.

Contents

1 Introduction

Both from the economic and social points of view, technological innovation represents an important phenomenon with significant consequences. Given its notable effects on economic growth, human progress, social development, economic stability, industrial development, and the environment, technological innovation requires continuous efforts to scientifically understand it. Understanding technological innovation, however, is non-trivial! First, one must decide on what it means to study technological innovation: does the research question concern technological innovation analyzed as a complex object (e.g. an aeroplane) or its development process? if technological innovation is analyzed as a complex object, the question which arises is how to model its components and figure out the relationships between them; If on the other hand, one wants to analyze the development process of a technological innovation, then one needs to specify which aspect or phase of the development process is the subject of the analyses: the emergence, diffusion, or evolution of a technology? Finally, a frequent subject of study in the technological innovation literature concerns the measurement and prediction of technological innovation: is technological innovation measurable? is there data that represents technological innovation? can we predict it? It is important to note that these questions are not necessarily independent and more than one question could be the subject of the same study. Crucially, each of these questions has been investigated with different depth. As far as the development process is concerned, most of the literature is focused on the problem of the diffusion of technological innovations. Diffusion of innovation is an empirically well-established area and benefits from sound mathematical models like the S-shaped and logistic curves [72; 61; 7; 51]. However, the focus on diffusion has created a gap in literature represented by the scarcity of studies that analyse the other aspects of the development process which are emergence and evolution [81; 70]. How do technologies emerge and why? Can we understand their evolution through time? Filling this gap is important because the mechanisms of emergence, diffusion, and evolution of technological innovation are linked through feedback mechanisms and complex interactions and therefore should be analyzed together. For this reason, to understand technological innovation as a process, then emergence, diffusion, and evolution need to (ideally) be combined in the same study. This this thesis is presented as an original contribution to this gap.

Often, the development process of technological innovations is modelled as a collective process of interacting agents. With this assumption (or fact) in mind, a large amount of research has been conducted to understand the process of technological innovation with network science. In this regard, the concept of network has been used to study several configurations, such as individuals or problem solvers in a research project or organisation [4], technological artefacts and designs [106; 164], or innovating firms which are linked through knowledge flow relationships. Substantial work has been done showing the role of network features on inter-firm knowledge transfer, such as network centrality and node embeddedness [90], network range, and cohesion [174]. One particular line of research on complex problem solving by collectives has concluded that the average path length of networks, defined as the speed at which networks disseminate information, can have a relevant effect on the performance of the collective[14; 151; 132]. Average path length correlates with speed of diffusion; Networks with short average path length allow for information flow throughout the network to be fast. The opposite case is with long average path length networks, which disseminate information slowly. In reference [132], the authors argued that in situations involving collective problem solving, short average path length networks performed better than long average path ones in the long term but not in the short term. The explanation for this is that short average path length networks circulate information about immediate solutions quickly and the system is likely to experience an early convergence to low-quality solutions. Long average path length networks, given their slow rate of information dissemination, allow agents to explore a wider variety of solutions and discover better ones. In another experimental study, authors of reference [151] show that short average path length networks perform better than long average path length ones. The study showed that the reason for this was that agents were able to rationally adapt their search strategies when they receive information more rapidly. Building on this, authors of reference [14] show that the performance of short and long average path length networks depends on the search strategy employed by the nodes.

As far as technological innovation is concerned, there is evidence that shorter path length correlates positively with system-level technological innovation [74]. Although these studies have spurred an interesting debate on whether network average path length can improve collective performance, little has been done to understand the role of other network features like degree-heterogeneity, edge direction, as well as different interaction rules. Do heterogeneous networks facilitate technological innovation in an innovation network? Under what conditions? Does edge direction have an effect? Does the interaction rule between agents affect the outcome of innovation? The aim of this thesis is to contribute to answering these questions by doing two things: first, develop and simulate agent-based models for investigating the process of technological innovation as a collective evolutionary process involving a technological space (also called fitness or technological landscape) and a network of innovating firms, and second, to present a historical case study on Financial Risk Management (FRM) to illustrate the role of network in real life collective innovation. The case study is constructed using stylized facts from the history of FRM. To my knowledge, this is one of few studies to analyze a real-life case study of collective innovations and is the first study to incorporate a regulator as a 'super-node' that influences the decisions of searching agents. The results from the agent-based model showed that network average path length and degree-heterogeneity are important factors affecting the average performance of the system. Another finding shows that there could be a trade-off between the frequency with which agents seek information from their neighbours and the rate with which they conduct an independent search. As for the case study, I used a projected financial network that is derived from the syndicated lending database DealScan, provided by Thomson Reuters, which I use to produce a heterogeneous and directed network. The results obtained so far for the case study of FRM show an interesting trade-off between market efficiency and system stability: in absence of regulatory intervention, the system achieves high performance but at the cost of lower diversity; in presence of the regulator, the system might achieve lower performance but keeping more diversity in the system.

The thesis is structured as follows: the second chapter investigates the problem of technological innovation: What is it, how to study it, why is it a complex problem, how to model it as a complex problem. The third chapter investigates the network and collective nature of the process of technological innovation: why is the process of technological innovation development a network and collective phenomena, how to model technological innovation as a collective problem? In the fourth chapter, I report on the results of the evolutionary agent-based model simulated with different network structures and interaction rules. Chapter 5 presents the history, model, data, and results of the case study provided as evidence for the collective and complex nature of technological innovation. Chapter 6 offers a discussion on how network studies like this thesis might be useful for policy-making. Conclusions and future extensions are offered in chapter 7.

2 Technological innovation as a complex problem

Technological innovation is a significant problem both from the economic and social point of view [56]. From an economic point of view, technological innovation can be considered as one of the main factors influencing the economic standards of living and the quality of social life in that it correlates with the wealth of a nation [196]. In the 1920s, the Russian Economist Kondratiev observed that there is a relationship between technological innovation and economic cycles of growth, recession, and depression. These cycles came to be known as Kondratiev Waves. Later, the Austrian economist Joseph Schumpeter sustained that the economy can be modelled as following business cycles whose ups and downs are caused mainly by the introduction of successful technological innovations by entrepreneurs [188]. For Schumpeter, technological innovations arrive in business cycles in the form of *clusters* of innovations. Schumpeter also argued that entrepreneurs would require financial capital for them to create technological innovation, thus emphasising the role of finance in the process of economic growth. Carlota Perez described how the episodes of technical changes which we have witnessed so far are heavily dependent on financial capital [166]; This is the reason why many critics argued that, had financial capital been invested in useful and productive technological innovations, the financial crisis of 2008 could have been avoided or results less consequential. In fact, following the crisis of 2008, the topic of technological innovation has re-emerged into the scene as a frequently proposed solution to fight economic meltdown [62]. Finally, it worth mentioning that technological innovation is important because it acts as a means by which nations and firms compete globally and allow countries to establish a comparative advantage.

From a social point of view, technological innovation is closely related to social development. Technological changes affect almost all aspects of people's life at home, work, or anywhere else. It can improve social welfare by increasing the efficiency and quality of services, free the society from the repetitive and time-consuming activities, enhance the delivery of health-care, education, transportation, and information-related services. The ways technologies are used by societies are of great importance. Technologies can be used for good as well as evil purposes. Problems like global warming, technological employment, ecological disasters, freedom restriction, and others happen mostly thanks to advances in technological innovation. Controlling technological innovation might not be feasible, however as Kevin Kelly sustained, the optimal approach to guarantee the best outcome out of technological innovation is to know how to direct technological innovation for the best uses [125]. According to reference [56], several players can influence the way technology is created and shaped: governments through their role as technology supplier (financing science), regulators through patent laws, intellectual property, and technical standards, and customers through their economic and political choices.

2.1 What is technological innovation?

Technological innovation has been the subject of inquiry for a wide range of scholars and scientists; therefore, different viewpoints have emerged to understand how to define and conceptualise technological innovation. In most definitions, the main discussion revolves around the idea that innovation encompasses a *creation process* focused to a large extent on the element of *newness*. Here, I should mention that innovation needs to be distinguished from what is called *invention*, in that invention is about the conception of a new idea, while innovation is about the actual development and implementation of an invention in life or in the production process. The term ’technological’ is often attached to the concept of innovation to indicate that innovation is developed within an existing technological paradigm that includes tools, instruments, knowledge (know-how and know-what), patterns of inquiry and behaviours used to tackle and solve specific problems in certain domains. In this thesis, I will adopt the definition according to which technological innovation is a creative activity, conducted by one or more agents, in order to develop a new solution to an existing problem within a specific domain.

Several types and classes of technological innovations exist. Among the economists who emphasised the importance of technological innovation in economic life is the Austrian economist Joseph Schumpeter. In his seminal work *Theory of Economic Development*, Schumpeter distinguishes five classes of innovations:

1. The introduction of a new good, like a product which is not yet familiar to customers, or unique quality of the same good. This type of innovations usually takes the name of product innovation.

2. The introduction of a new method of production, i.e. one that is not yet used or tried in the sector or industry concerned. In many cases this type of innovation is called process innovation.

3. The opening of a new market: a market that has not been exposed yet to the products or services of the concerned firm.

4. New source of supply of raw materials or intermediate goods.

5. A new organisation of an industry, like the creation or breaking up of a monopoly position.

The classification offered by Schumpeter presents itself as a general classification and does not take into account the nature of the innovation being introduced. To provide a more accurate and detailed distinction between the different types of technological innovations, scholars and technologists have proposed several classification schemes. For example, authors in reference [149] proposed a distinction between the following types of innovations:

- **Radical Innovations**: Innovations that leave a substantial impact or cause a significant change in the market or industry where they are introduced. In some contexts, radical innovations are also referred to as *revolutionary* or *breakthrough* innovations.

- **Incremental innovations**: Innovations that come in the form of small improvement or change to products, services or processes. Another term used to describe incremental innovations is *evolutionary* innovations.

- **System innovations**: Innovations that require a long time and several resources to realise. Examples are IT systems.

Distinguishing between a radical and non-radical innovation is a non-trivial task and it represents an important research question (see for example [50; 39; 45; 208]). Authors in reference [50] defined a radical innovation as one that is: (1) novel; (2) unique; and (3) has an impact on future technology. From this definition, it can be seen that defining a radical innovation requires a novelty criterion as well as a method to estimate the future impact of an innovation on the marketplace and society. Additionally, radical innovation is not to be confused with *disruptive innovation*, which according to the author of [40] is an innovation that takes a produce of service that is accessible to only a small category of people and make it affordable or accessible a large part of the society.

In another study, authors of [100] proposed the following categorization of innovations:

- **Radical**: These innovations establish a new dominant design with new defining concepts that link components in a new architecture. For example, the move from room air fan to central air conditioning.
- **Incremental**: Incremental innovations refine and extend an established design. Most of the improvement occurs in individual components, without altering the underlying concepts and links between them, for example, introducing an improvement to the air fan's blade design or in the power of the motor.
- **Modular**: It is an innovation that changes only the core design concept without changing the product's architecture. For example, replacing the analog with digital telephones, it turns the core design concept but not the product's architecture.
- **Architectural**: Innovations that change only the product's architecture without changing the components or core design concept. For example, the introduction of small, portable fans in addition to ceiling-mounted room fans. Here, the primary components (blade, motor, control system) are the same, but the architecture is different.

2.2 What does it mean to study technological innovation?

Studying technological innovation can refer to different things. Within the research community that investigates the phenomenon of technological innovation, four main topics have been the centre of focus: emergence, diffusion, evolution, measurement and prediction. In the next sections, I offer a brief discussion of each of these subjects of study.

2.2.1 Emergence

The first subject of study that has received attention in the technological innovation literature is the emergence of innovations. Researchers in this line of inquiry try to answer the question of how do innovations emerge and why they are initially developed. Concerning the why technological innovations emerge, the existing literature has focused on two main perspectives. In the first perspective, an innovation emerges due to an identified market/customer need or a business problem which require firms to innovate to satisfy the demand or solve that specific problem. This theory is called *market pull*. In a recent study on the Swedish firms, author of reference [200] developed an analytical framework of the historical driving forces of innovation in Sweden and found that Swedish innovations were mostly developed as a response to problems or new opportunities, which can be seen as evidence for the market pull hypothesis. The other perspective is known as *technology push* and perceives an innovation as the result of a new technology or idea that is developed independently of market demand, therefore emphasising the supply aspect of technological innovations. For example, a firm supplies a new technology to increase profits, or a new technology emerges because it has been enabled by another technology [209]. Proponents of each perspective debated for several years and a conclusion is reached according to which both factors (push and pull), are crucial in explaining the emergence of technological innovations [58; 154]. Following this conclusion, the shift has been towards understanding the mix of economic, political, institutional and technological factors that drive innovation [210].

In addition to the question of why technological innovations emerge, scholars have been interested in understanding how innovations emerge. In this regard, the most dominant view is the combinatorial mechanism which models the emergence of a new technology as a novel combination of *existing* objects or technologies [9]. The combinatorial theory can explain the emergence of a wide range of technological innovations; however, it fails at explaining the rise of breakthrough innovations (or novelty) which are not merely a combination of existing technologies. One explanation was offered by Brian Arthur who sustained that technological novelty emerges *"not just from combination of what exists already but from the constant capturing and harnessing of natural phenomena. At the very start of technological time, we directly picked up and used phenomena: the heat of fire, the sharpness of flaked obsidian, the momentum of stone in motion. All that we have achieved since comes from harnessing these and other phenomena, and combining the pieces that result"*.

Another important theory that is gaining popularity in the scientific community is the theory of *adjacent possible* introduced by Stuart Kauffman [122]. For Kauffman, the adjacent possible is a consequence of the dynamic nature of the biosphere which is continuously creating novelties in ways that typically cannot be foretold [123]. In other words, the adjacent possible consists of all possible novelties that could potentially emerge (through recombination) in the next time step, given the present state of the world. It is a subset of the whole space of possibilities which is conditional on the present [206]. Kauffman shows that like the biosphere, *"the 'econosphere' is a self-consistently co-constructing whole, persistently evolving, with small and large extinctions of old ways of making a living, and the persistent small and large avalanches of the emergence of new ways of making a living"*. Whenever a

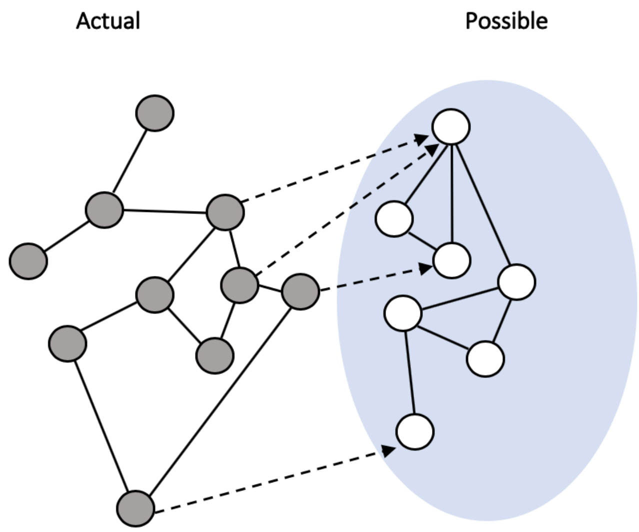


Figure 1: Graph-theoretic illustration of the adjacent possible

novelty emerges in this way, then part of what was possible becomes actual, and this actual in turn is bounded by a new adjacent possible [207]. Figure 1 presents a graph-theoretic representation of the adjacent possible as suggested in [140]. Nodes in grey are assumed to represent the actual technologies abstractly. A link between two nodes implies that these two technologies can be combined together to produce new technology. The nodes in white are the novelties that have not been discovered so far, but yet they are possible given the present state of the world. Some problems in economics can be assumed to have an adjacent possible that is dynamically changing and expanding with time. Technological innovation is perhaps the most practical example [215; 114; 123].

2.2.2 Measurement and prediction

The second line of inquiry in technological innovation concerns the problem of measurement and prediction of technological innovations.

Several conventional measures have been used to quantify technological innovation. The two most common are expenditure on research and development and patents, which can be treated as direct measures of technological innovation. Another way to measure technological innovation is to use a functional approach, where the examiner observes not the innovation itself, but its desired effect. For example, authors in reference [67] measured technological innovation in the energy sector indirectly by analysing the reduction in costs of production of different forms of energy and attributing such a decrease in costs to innovation.

Measuring technological innovation entails several challenges. As I stated in the previous section, the conventional definition of technological innovation is the Schumpeterian definition, according to which technology consists of novelty, i.e. the creation of completely new products and services. However, this definition might suffer from two shortcomings: first, important innovations should not necessarily consist in totally novel things as Schumpeter claimed; innovative activities which cause relatively small changes in product performance may also have important economic and technological consequences [178]. Innovation can also be cumulative , in that small inventions accumulate over time to produce a final innovation [191; 53; 203]. Newness may also be subjective, in that innovation is considered new if it is perceived as such by its potential users or developers. Finally, since *'new'* does not necessarily translate into *'better'*, the measurement of technological innovations might require the establishment of a criterion to classify a technological innovation as such [65].

The measurement of technological innovation can assist researchers in predicting future innovative activities. And if it was possible to predict technological innovation, then it might be possible to predict economic growth, which is among the most critical prediction tasks for policymakers. However, similar to the problem of measurement, the prediction of technological innovation can be challenging for two main reasons. First, to predict an innovation we must identify the forces driving its emergence. As discussed in the previous section on emergence, pull vs push theories had dominated the scene for a long time, where the question was whether technological innovation is the result of market demand or of independent technological trajectory which is evolving through time. In addition to whether the driving factors are push and pull, these factors can change and evolve. This, in turn, might require a dynamic theory of technological innovation. The other reason why we cannot easily predict technological innovation is the non-ergodicity of dynamics in a technological landscape. Assuming that innovations emerge by combinatorial process [9], then a successful innovation will depend on putting together existing products and services in new, unanticipated ways. Because it is impossible to define in advance all the possible combinations, the prediction of future technological innovations becomes almost impossible [118], unless a formulation is done in a statistical sense.

Crucially, although the prediction of single technological innovations may be impossible, Giovanni Dosi states that it might be possible instead to identify *technological paradigms* and *technological trajectories* as an indicator of the general innovating behaviour in the market [58]. According [60, p. 16], a technological paradigm entails *"a definition of the relevant prob-*

lems that must be tackled, the tasks to be fulfilled, a pattern of inquiry, the material technology to be used, and the types of basic artifacts to be developed and improved", while technological trajectories describe progress along one paradigm. Some studies have confirmed that the pattern of innovative behaviour in an industry can be the product of different technological paradigms [141], while other studies have shown that there is not clear evidence whether firms in a specific industry behave in the same way [133]. Some studies modelled technological trajectories as patent citation networks [213]. The question of whether technological paradigm and trajectories exist, and how to describe them remains one of the challenging tasks and open questions in evolutionary economics.

2.2.3 Diffusion

Diffusion of innovations is a research topic that seeks to explain why and how innovations are adopted by participants in a social system, the speed and mechanism behind the diffusion of certain types of technologies rather than others, and the characteristics of innovation users [177]. Innovation diffusion occupies a special place in the technological innovation literature, mainly due to the availability of data on several case studies and the use of mathematical tools like S-curves and logistic models [190]. In this research line, the hypothesis that specific technologies may win over others due to technological *lock-in*, as well as positive feedbacks, has gained great importance [33; 7]. The importance of technological lock-ins derives from two main reasons: first, a technological lock-in can mean that the system has converged to a sub-optimal solution, and second, technological lock-ins might induce a reduction of the diversity in the system which can lead to fragility. If a system of agents (e.g. a market) adopts a unique solution to a problem, then an unfavourable change in the environment would put the whole system at risk. If however agents adopt a variety of solutions within every generation, then the system as a whole is likely to survive changes in the environment.

Another well-known mechanism for innovation diffusion is the so-called *network effect* or *network externality*, i.e. the fact that the use of one product or service by one agent is influenced by the number (network) of other agents who use the same product or service [117; 63]. Last but not least, an influential theory that has received considerable attention in the diffusion of innovation literature was proposed by Mark Granovetter, who sustained that the presence of *'weak ties'* could lie behind the process of diffusion of specific innovations, especially the unconventional ones [94]. Innovators outside the conventional system may have more freedom in experimenting new ideas and methods, and should they succeed in achieving an important discovery, players from the conventional system who has weak ties to the innovators outside the system may produce better results.

2.2.4 Evolution

In most studies that use the concept of diffusion of innovations, there is an implicit assumption that the innovation being adopted has reached its final development or commercial shape. In reality, innovations can change in their composition, shape, increase in details and evolve into different forms through time as a response to feedback, enabling technologies, and other factors. In this case, researchers often use the concept of 'evolution' of innovations. When studying the evolution of technological innovations, the problem is not simply why and how certain technologies are adopted rather than others, but to understand the mechanism behind the temporal evolution of innovations as if they were like living organisms which undergo mutations and recombination based on a common descent [9]. According to Brian Arthur, understanding how technologies evolve is the essential question in technological innovation because *"without evolution - without a sense of common relatedness- technologies seem to be born independently and improve independently. Each must come from some unexplained mental process, some form of 'creativity' or 'thinking outside the box' that brings it into existence and separately develops it. With evolution (if we can find how it works), new technologies would be birthed in some precise way from previous ones, albeit with considerable mental midwifing, and develop through some understood process of adaptation. In other words, if we could understand evolution, we could understand that most mysterious of processes: innovation."* The diffusion of technological innovations can happen simultaneously with evolution, and therefore a deeper understanding of technological innovation would require studying evolution and diffusion as two mechanisms operating simultaneously [153].

2.3 Complexity and technological innovation

In this thesis, technological innovation will be treated as a complex economic problem. Before moving to a discussion of what complex problems are in economics, it is worth illustrating what economic means. In the mainstream approach, an economist would think of an economic problem or economic decision-making as constrained optimisation. At the micro level, the problem for economic agents is constrained choice: the individual customer exercises rational choice between baskets of goods; the producing firm chooses between production plans, and the investor chooses between portfolios of investments. At the macro level, the problem for society is the efficient allocation of scarce resources to the many uses. For the regulator, the problem is to guarantee fair competition and avoid costly bailouts. Typically, the mainstream approach to economic problems assumes the existence of a single optimal solution, the market equilibrium point. This followed mathematically from the shapes of the supply and demand curves, including the assumption of diminishing returns to scale for demand curves. Among the promising alternative approaches to economic problems is what I call here the *complexity approach*, which traces its origins from evolutionary theory and organizational learning literature and differs from the mainstream approach in that it emphasizes the role of combinatorial evolution, bounded rationality, local search, and trial and error in the study of economic problems [157; 145]. Economic problems addressed by the complexity approach are assumed to include the possibility of multiple optima and the premature convergence on a (sub-optimal) local peak.

To operationalize complexity in the complexity approach, it is worth clarifying what is usually meant by complexity. To this end, I would discuss two perspectives. The first perspective is based on the difference between ontological and epistemological complexity [41]. Simply, ontology is concerned with the nature of a problem; epistemology is how we know that problem [10; 23; 28; 91]. Building on this distinction, ontological complexity can be defined as the complexity of the problem being analysed such as the complexity of an organisation or a technological artefact. Epistemological complexity, in contrast, deals with the potentiality of what might be known and what may never be known about a complex problem [184]. In other words, epistemological complexity refers to *"how complex are our descriptions"* [41]. In this thesis, I assume that the epistemological complexity is a pale shadow of ontological complexity. This point can be clarified by relating to the classification of ontology and epistemology based on the assumption of an objective 'out there' existence of reality, or a subjective reality that is concluded through meditated social interpretation [23; 28]. Objective ontological complexity assumes that a phenomenon is complex regardless of human cognition; based on this assumption, we can say that objective epistemological complexity looks for the variables and their relationships, which exist 'out there', to model the objective ontological complexity of a problem. The role of the researcher here is to discover these variables and locate reality using observation, measurement, taste, and touch [113; 28]. The contrasting alternative is the subjective view (often called interpretivist or constructivist), and it argues that *"the reality that people confront is the reality they construe"* [91]. The subjective view would lead to the argument that ontological and epistemological complexity of a problem are socially constructed and dependent upon human cognition and social relationships. While objective epistemological complexity assumes the existence of objective ontological complexity, the challenge is with subjective epistemological complexity which can presuppose either an objective or a subjective ontological complexity (see [113]). The other perspective which I adapt to understand complexity is more practical and relies on the distinction between static and dynamic complexity [32], which I illustrate next in more detail.

In the static sense, a problem is said to be complex if at least two conditions are satisfied. First, the solution to the problem is made up of several components where each component can assume one of the different values or attributes [132; 192]. Mathematically this can be modelled as a set W of N components:

$$W = (\omega^1, \omega^2, ..., \omega^i, ...\omega^N)$$

Where ω^i represents the choice for component i, for i = 1,..., N. For each component, the set of possible options can be assumed to be discrete or continuous, and such choices could be qualitative or quantitative. For example, in the discrete case, we can assume that there are s choices to choose from and therefore

$$\omega^i \in \{1, ..., s\}$$

For i=1,......,N, where *s* is a positive integer. Hence the number of solutions is finite and given by

$$\#\Omega = s^N$$

The second condition for static complexity requires the existence of non-linear inter-dependencies between the components of a solution such that the total evaluation of the goodness of a solution is not merely the sum of the goodness of the individual parts. Mathematically, dependencies can be modelled using matrix or network representation. For example, considering the formula for W defined above, we can think of a graph G=(V,E) where V are nodes representing the components of a problem $(\omega^1, \omega^2, ..., \omega^i, ...\omega^N)$, and E is the set of edges (could be weighted edges) that represent the dependence between two connected components. By definition, the presence of multiple dimensions that need to be optimized, and complex inter-dependencies between them implies that there are different solutions to choose from, thus giving rise to what is called a solution space, or what I shall call in this chapter a fitness landscape (discussed in detail in the next sections).

In many cases, the static complexity of a problem might change with time, giving rise to what is called *dynamic complexity*. The dynamic view is mainly concerned with the nature and rate of change of both the components of a problem and their relationships, as well as the behavioural changes that follow from such dynamics (e.g. dynamic optimisation). In other words, dynamic complexity can be viewed as representing the changing patterns of static complexity. Facts supporting the presence of dynamic complexity can be: increase in the dimensionality of a problem, change in the pattern or nature of connectivity among the components of a problem, feedback loops that can alter the landscape of solutions, and the emergence or expansion of a landscape. Throughout the thesis, it will be shown that both the static and dynamic complexity views are important in understanding the fitness landscapes of economic problems as well as the adaptation strategies of agents.

The question now is how do these definitions of complexity (ontological, epistemological, static, and dynamic) apply to the problem of technological innovation? The next section provides an answer to this question and provides a review of the most important models of fitness landscape.

2.3.1 Technological innovation as a complex problem

Technological innovation is an outstanding example of a complex problem to which the four forms of complexity addressed in the previous section apply (ontological, epistemological, static, and dynamic). Technological innovations are complex in the ontological sense, in that they are real human-made objects or know-how that emerge and develop for economic and social reasons [9]. They are also complex in the epistemological sense, in that humans might have limited ability to understand and discover the dimensions of a technological problem and the nature of interactions between its components. To theoretically model the ontological complexity of technological innovation, the static view discussed previously can offer a practical approach. Typically, a technological innovation is made up of many components, where each component can assume one of different shapes or values, with different degrees of independence or complementarity among the components [152; 81; 73]. For example, to build a jetfighter, one may need to optimise dimensions such as speed, firing power, weight, and manoeuvrability. However, these dimensions are to some degree dependent on each other so that an improvement in one might result in reduced performance in another. To make jetfighters faster implicates that they should be lighter, yet this can translate into less firing power. The higher is the number of components and complementary relationships of a technological problem, the more complex is the solution space of that problem [200; 105; 73; 136; 120]. When complementary interactions are dense, then a change in the value of one component can have a large impact on the goodness (fitness) of a solution, which I refer to as ruggedness in the next section. Viewed from a mainstream economic perspective, before-mentioned problems are typically assumed to have single optimal solutions. For example, production functions and production frontiers are modelled as smooth and convex functions that admit unique solutions, mostly to simplify analysis [120; 178]. However, problems like technological innovation require optimisation which entails conflicting constraints, implicating the need to search in a complex landscape of many potential solutions. As I illustrate in the next section, complexity theorists often use the approach of fitness landscape to model the solution space of an economic problem.

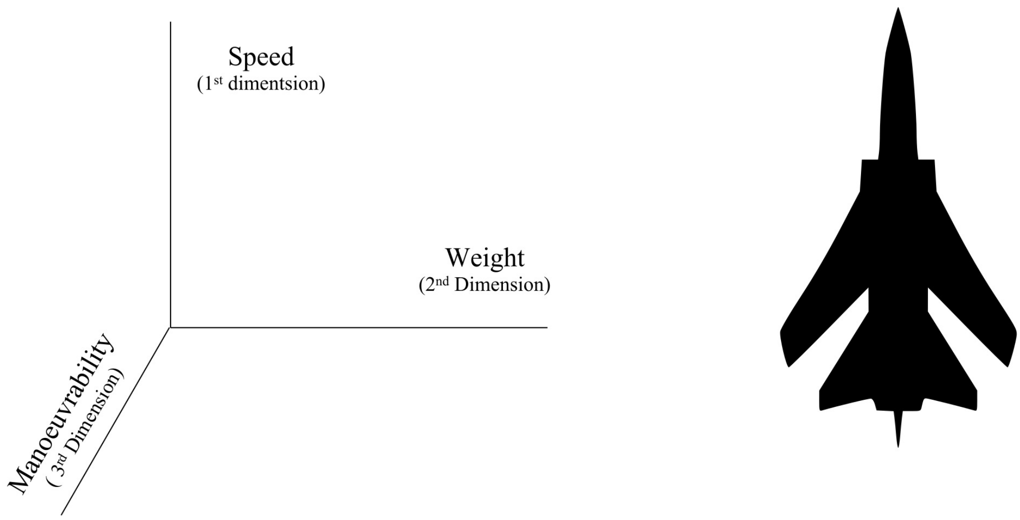
Moving from the static view of complexity, it is reasonable to assume that the complexity of technological innovations increases over time [8]. For example, technological artefacts like an aeroplane become larger and more sophisticated and composed of more parts. Similarly, the complementary structure of a technological innovation (number and weight of links between components) may change as time passes, for example connecting more parts of an aeroplane.

As I mentioned earlier, the most commonly adopted approach to the modelling of technological innovation is the fitness landscape approach. To give an idea about this approach, the next section illustrates its main features.

2.3.2 The variety of fitness landscape models

Theoretically, to model a fitness landscape, three main ingredients are necessary. First, given a space **S** of possible (or available) solutions to a problem, an encoding function enc: **S** $\Rightarrow F$ that encodes elements from the solution space to some representation space so that an algorithm can process them. Examples of representations spaces are the real vector $(F = \mathbb{R}^n)$, the bit-vector (F = C(n)), a permutation $(F = P_n)$, or more complicated configurations. Figure 2 illustrates a simple example of a fitness landscape.

The second ingredient in a fitness landscape model is a fitness function $f : F \Rightarrow \mathbb{R}$ that assigns fitness value f(x) to an encoded solution x. The fitness function is not enough to define a fitness landscape. What is needed is a notion of neighborhood or connectedness between the



Speed: high or low ——— Encoding ——→ High: 1
Weight: high or low ———→
Manoeuvrability: high or low ——— ——→ Low: 0

One solution: (high speed, low weight, high maneuverability) ——→ (1,0,1)

All possible solutions: (1,1,1), (1,1,0), (1,0,1), (0,1,1), (0,0,1), (0,1,0), (0,0,1), (0,0,0)

Fitness values:
(1,1,1): **1**
(1,1,0): **2**
(1,0,1): **3**
(1,0,0): **4**
(0,1,1): **5**
(0,0,1): **6**
(0,1,0): **7**
(0,0,0): **8**

Neighborhood (Hamming distance):
(1,1,1): **{(0,1,1), (1,1,0), (1,0,1)}**
(1,1,0): **{(1,1,1), (1,0,0), (0,1,0)}**
(1,0,1): **{(1,1,1), (1,0,0), (0,0,1)}**
(0,1,1): **{(1,1,1), (0,0,1), (0,1,0)}**
(0,0,1): **{(0,0,0), (0,1,1), (1,0,1)}**
(0,1,0): **{(1,1,0), (0,0,0), (0,1,1)}**
(0,0,1): **{(0,0,0), (0,1,1), (1,0,1)}**
(0,0,0): **{(1,0,0), (0,1,1), (0,1,0)}**

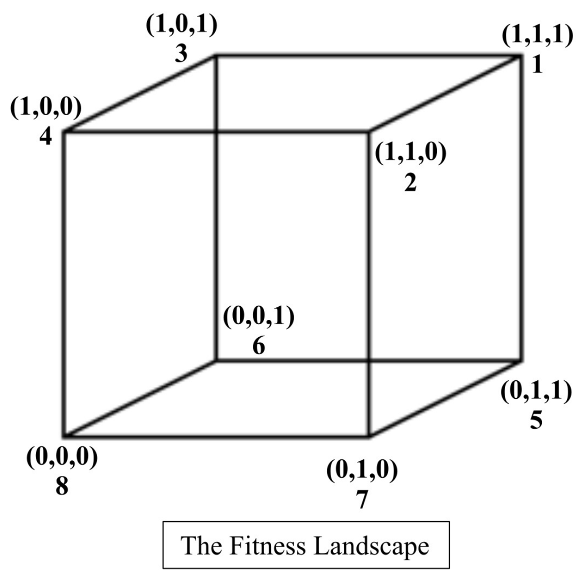


Figure 2: Simple illustration of how fitness landscapes are constructed. This hypothetical example takes the jetfighter as a complex technological problem characterised by three dimensions: speed, weight, and manoeuvrability. Each one of these dimensions can assume one of two values: high or low. High is encoded as 1, and low is encoded as 0. The total number of possible solutions is 8, where each of the three positions in the solution represents one dimension of the problem. Fitness values are assigned randomly by taking the numbers from 1 - 8. Finally, each solution is connected with those solutions which are at a Hamming distance of 1.

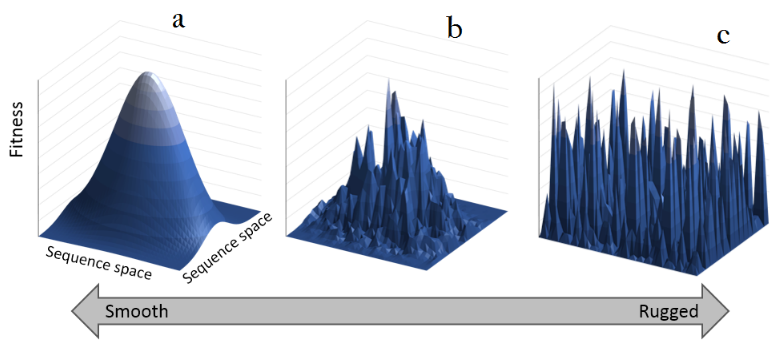


Figure 3: Different typologies of fitness landscapes. A smooth fitness landscape (left) has a small number of local optima. Smooth landscapes are easy to explore since all that is required to reach the optimal point is hill-climbing heuristic. By increasing the number of local optima (figures on the left and right), search becomes more complicated, and a simple hill-climbing heuristic would eventually end up in a local optimum that is sub-optimal compared to the other local optima. Figure source: [189]

different solutions in S. Here, a neighborhood function $\mathcal{N} : S \Rightarrow 2^S$ (2^S is the power set) is often used to associate a set of neighbor solutions to a candidate solution x. An essential part of a neighborhood function is the choice of the operator of the algorithm, for example:

$$\mathcal{N}(x) = \{y \in S | Pr(y = op(x) > 0\}$$

$$or \quad \mathcal{N}(x) = \{y \in S | Pr(y = op(x) > \alpha\}$$

Where op(x) is often replaced with a distance measure. To illustrate the concept of neighbourhood, suppose we choose a distance measure like the Hamming distance which measures the number of positions at which the two strings or combinations differ:

$$\mathcal{N}(x) = \{y \in S | d_{Hamming}(x, y) = 1\}$$

And suppose we want to obtain the neighbors of candidate (0,0,0) in the fitness landscape in figure 2, then

$$\mathcal{N}(0, 0, 0) = \{(0, 0, 1), (0, 1, 0), (1, 0, 0)\}$$

In summary, a fitness landscape $\mathcal{F}$ is defined as a set of encoded search space F, a fitness function f(x) and neighborhood function $\mathcal{N}(x)$ that assign fitness value and distances between solutions in F .

$$\mathcal{F} = (F, f(x), \mathcal{N}(x))$$

A variety of fitness landscape models have been used to model complex economic problems, and in particular technological innovation. The first model of fitness landscapes is the rank model (often called the House of Cards model) which was proposed by Kauffman and Levin [119]. In the rank model, a problem is represented as a string of length N, where each element in the string can assume a value of 0 or 1, producing a total number of combinations equal to 2^N . Integers from the interval [1, N] are assumed to represent fitness values which can be assigned either randomly or with a correlation constraint. Example of correlation constraint could be that the difference in fitness between neighbouring combinations is $\leq \alpha$. One example of a study that used similar fitness landscapes is reference [59], in which the authors examined the optimal organisational power structure by modelling the space of different power structures as a fitness landscape where fitness values of the various structures were ranked from best to worst.

The second and most popular model of fitness landscapes is the NK model proposed by Stuart Kauffman to model the effect of 'epistasis' on the structure of fitness landscapes [119]. In biology, epistasis refers to the fact that genes affect other genes, which can be translated into economic settings as institutional complementarity or technical complementarity. In the NK model, N represents the dimension of the problem, and K controls the level of interdependence among the components. If k=0, then the interactions between components is absent and, therefore, each component contributes independently to the fitness of a given solution. On the other extreme, if k=N, then every component interacts with all the other components and the fitness contribution of one component influences the fitness contribution of all the other components. By varying K, one can control for the ruggedness (complexity) of the fitness landscape (see Figure 3 for illustration of the concept of ruggedness). The NK model has been extensively used to model technological and organisational landscapes [200; 73; 135; 216]. In assigning fitness values to the different combinations, the NK model relies on the idea of averaging fitness contributions. The fitness of a given combination V is the sum of contributions from each locus V_i in the combination:

$$F(V) = \sum_i f(V_i)$$

The contribution from each locus in V depends generally on the value of k other loci:

$$f(V_i) = f(V_i, V_1^i, ..., V_k^i)$$

For instance, assume you want to model an NK model with binary strings, with N=4 and K=1. An example of a string is [1,0,1,1]. The components of a string are often called locus (plural loci). In this case, the fitness of one string is the sum of the fitness contributions of all 4 loci (because the length of the string is 4). Given that K is 1, then the fitness contribution of one locus will depend on the value of the locus and one other loci in the string. For the sake of simplicity, I shall use the convention $f(S_i) = f(S_i, S_{i+1})$, which means that each locus i is affected by its immediate neighbor i+1, and given the possibility of cyclicity $f(S_4) = f(S_4, S_1)$. If we take the fitness function to be as follows: f(0, 0) = 1; f(0, 1) = 0; f(1, 0) = 0; f(1, 1) = 2, then the fitness value of the string [1,0,1,1] is

$$F(1011) = f(1,0) + f(0,1) + f(1,1) + f(1,1) = 0 + 0 + 2 + 2 = 4$$

Hence, the fitness function $f(V_i, V_1^i, ..., V_k^i)$ can be considered as a mapping between combinations of length k+1 and scalars, which are usually called "fitness contributions". In obtaining these fitness contributions, the original NK model relies mostly on random draws from a uniform distribution between 0 and 1 [120]. In the original NK model, the fitness value can be written as the average of the combination' fitness contributions:

$$z = \frac{1}{N} \sum f(V_i)$$

For example, assuming N=3 and k=2, then a fitness function can be obtained as the average of three numbers drawn randomly from the interval [0, 1]. Figure 4 shows an example of an NK constructed following this method.

	f_1	f_2	f_3	f
000:	0.5	0.3	0.5	0.4
001:	0.6	0.4	0.9	0.6
010:	0.1	0.7	0.3	0.3
011:	0.2	0.4	0.6	0.4
100:	0.9	0.9	0.9	0.9
101:	0.5	0.7	0.1	0.5
110:	0.6	0.2	0.1	0.3
111:	0.9	0.4	0.8	0.7

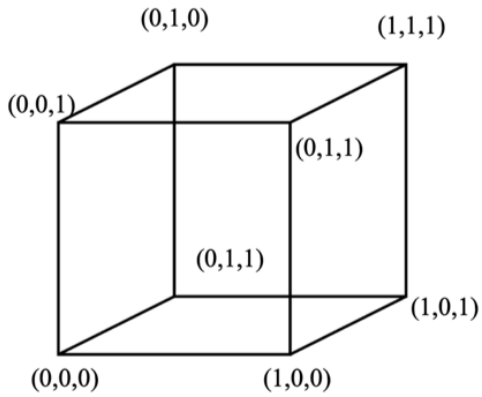


Figure 4: NK landscape with N=3 and K=2

Two main variants of the NK have been proposed. The first extension is known as the generalised NK model [5]. Generalised NK landscapes are constructed following the same logic as in the original NK model. Besides, the basic properties of the NK landscapes, which relate the interdependence parameter k to both the number and the fitness of local optima, remains the same. There are however two main differences; First, whereas in the standard NK model the k parameter is uniform and symmetric across all elements of a combination, in the generalized NK the k parameter can be heterogeneous and asymmetric, meaning that K is not the same for all components of a solution and influence of one element on another can go one way only. In this way, some components may have many connections (high K) with other components, in which case they are called *core* components, while others have few connections (low k) and are called *periphery* components. Second, the number of elements in a combination does not necessarily need to be equal to the number of fitness components F. In the original NK model, it is assumed that each element in a combination is identified with one fitness component. To illustrate more the idea of the generalized NK model, I follow the reference [5], which used matrix representation to explain it.

Figure 5 shows three examples. Panel (a) presents a standard NK model with N=4 elements (denoted with E) and k=1. There is a total of four fitness components (denoted with C). An X in the matrix indicates the element influences fitness contribution. Panel (b) instead presents a generalised model with N=4 and non-uniform k. In this generalised model, we still have a number of fitness components equal to N. In panel (c), the number of elements N is equal to

	E_1	E_2	E_3	E_4
C_1	**X**			
C_2		**X**		
C_3			**X**	
C_4				**X**

(a)

	E_1	E_2	E_3	E_4
C_1	**X**			**X**
C_2	**X**	**X**		**X**
C_3			**X**	**X**
C_4		**X**	**X**	**X**

(b)

	E_1	E_2	E_3	E_4	E_5	E_6	E_7
C_1	**X**				**X**	**X**	
C_2		**X**			**X**		**X**
C_3			**X**			**X**	
C_4				**X**	**X**	**X**	

(c)

Figure 5: Generalized NK models. Panel(a) is a standard NK model with N=4 elements(denoted with E) and fixed K (K=1). Panel(b) is a generalized NK model with N=4 and variable K. Panel(c) is a generalized NK model with N=7 and variable K

7, while the number of fitness components is 4. Panel (c) shows that elements E5 and E6 are strongly interdependent, while the other elements are loosely connected. The interdependent elements can be thought of as the core elements of the system, while the loosely connected elements as the periphery. It could be noticed also that in panel(b), the system presents no modular structure, while in panel (c) we have five almost modular elements $(E_1, E_2, E_3, E_4, E_7)$ and two strongly interdependent elements (E_5, E_6).

Versions of the generalised NK models have been used to model complex economic problems like the financial system [105] and modular product architectures [80]. The second variant of the NK model is the N(K+C) which was proposed to account for co-evolution [124]. In the N (K+C) model, the fitness contribution of each element of a group depends both on K other intra-group features as well as on C features for other groups. Given a solution with N elements, the fitness of such solution may be either independent of the attributes of other organisations (i.e., C=0), or it may depend on a variety of characteristics of other agents (i.e., C >0). The rationale behind N(K+C) is that different agents or groups can have different fitness landscapes and the actions of one agent in her landscape can influence the landscape of another agent. Authors often use the term 'coupled' or 'linked' fitness landscapes to model co-evolution.

Other fitness landscape models used in the literature include random field models, which are random functions of multidimensional parameters that can be simulated using different correlation lengths and probability distributions [121]. Figure 6 shows the plots of two random fields generated on a 2-dimensional lattice using multivariate standard normal distribution with an exponential covariance function:

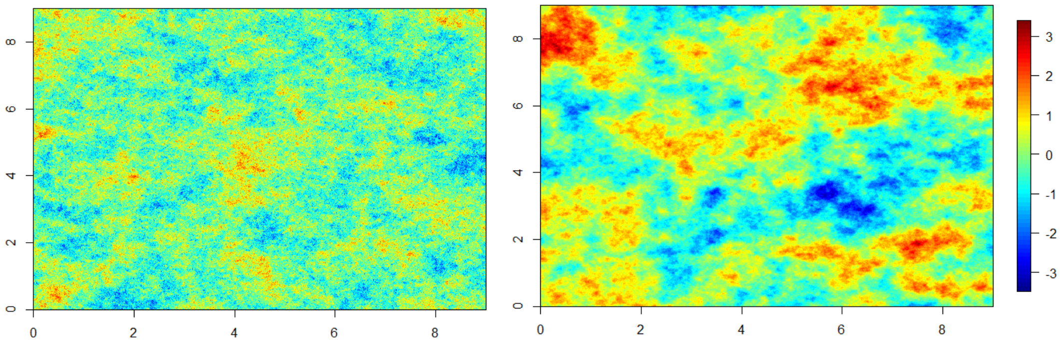


Figure 6: Realizations of a mean zero Gaussian random field model with correlation length 0.1 (left) and 0.5 (right)

$$cov(\mathbf{z}(x), \mathbf{z}(y)) = exp(-\frac{||x - y||}{l})$$

where $\mathbf{z}()$ is a zero mean multivariate Gaussian function, and y and x are two dimensional parameters expressing a pair of tuples of the form (x_i, x_j) and (y_i, y_j). Each item in the tuple assume values in the range 0-1 separated by an interval of 0.001. This implies that the final lattice (figure 6) have a 1000 x 1000 dimension, and a total of one million tuples. The plots were generated using correlation length 0.1 (left plot) and 0.5 (right plot). The correlation structure of a fitness landscape is an important concept since it conveys information about the extent to which nearby locations on the fitness landscape have similar fitness values. For large values of correlation length, the landscape is typically smooth and has fewer peaks, while for smaller values of correlation length, the resulting landscape is rugged with multiple peaks and valleys. A necessary consequence of highly correlated landscapes is that it will require more steps to escape a correlated area [119].

Another promising model for fitness landscapes is called *Design Structure Matrix*. Design Structure Matrices (DSM) is a network modelling approach used in engineering management and other fields and can offer a workable solution [195; 64; 152]. DSM are equivalent to an n x n adjacency matrix = A_{ij} , with $A_{ij} \neq 0$ if component i impacts component j (and $A_{ij} = 0$ otherwise). The advantage of DSM is the high modelling flexibility achieved thanks to network science tools (see figure 7).

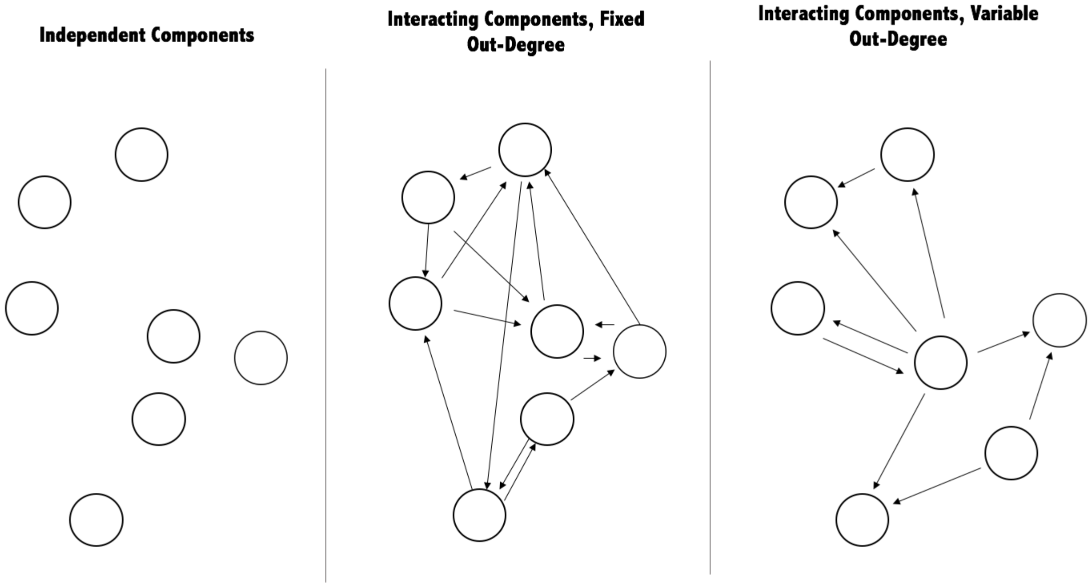


Figure 7: Interaction topologies in DMS. Assume a problem made of seven components (nodes). A DSM can model a variety of scenarios such as the simple but unrealistic scenario of completely independent components (left), fixed out-degree, where each component affects exactly 2 components (middle), and most realistically the case with variable out-degree (right). Such logic was used in [152]. Weighted interactions have also been considered [36]

Finally, a recent model has been proposed with the name 'local optima network', which is a network that as nodes has the local optima of the fitness landscape (the concept of local optima is explained in the next section) and links are the transition probabilities between them [160]. Local optima networks have not been used to model economic problems, but I believe they have some potential. In one scenario, the researcher might be interested in abstracting away the existence of search trajectories and focus on local optima and multiple equilibrium, in such case a local optima network can be a good model. Using a network of local optima one can also save on the computational complexity required to build and represent large multidimensional spaces like in the case of NK and random fields. In another scenario, the searcher or problem solver might know already the local optima of a problem, in such case, there would be no need to draw the entire landscape and rely only on a network representation of the phenomena. To give an illustration of how a local optima network can be modeled, figure 8 presents a directed network where each link points from one local optimum to another with an attached probability of moving between nodes. This network representation can allow for using tools from network science such as random walks on networks, PageRank, and other

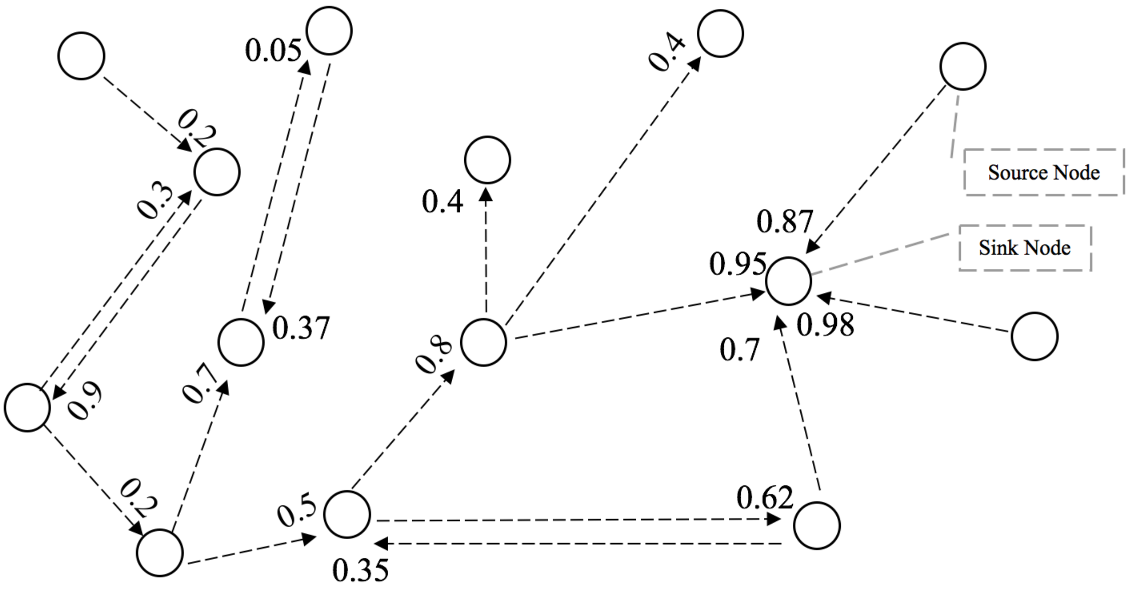


Figure 8: Local Optima Network

tools to study the dynamics of the system. In network science, there are concepts like sink and source node where the first refers to nodes that have only incoming links, while the later has out-coming links. A lock-in situation in a local optima network can be thought of as a sink node that once reached the system is trapped there (see figure 8).

As explained later in section 4.1, in this thesis, I construct and use an artificial fitness landscapes having as a main feature presence of multiple local peaks and one global optimum.

2.3.3 Solving the technological innovation problem: evolutionary dynamics on fitness landscapes

A crucial assumption in a fitness landscape model of an economic problem is the rule used to optimise or adapt in such landscapes. In describing these rules, researchers often use concepts like 'problem solving' [148], 'optimal search' [121], 'adaptation' [135], or 'exploration and exploitation' [144]. Behavioural rules can be classified into two categories. The first category includes *algorithmic* rules, which entail a search for an optimal solution that avoids sub-optimal lock-ins. Algorithmic search can be treated as a complex (detailed) version of the problem of optimization in mainstream economics. Ideally, the algorithmic search would work if solutions were enumerable, explorable and comparable, at least as an assumption. Second are *adaptive* rules, which are used to model and analyse the trajectory of evolution over time. Adaptive behaviour is commonly associated with the process of evolution. Evolution on fitness landscapes can be treated as a formalization of Herbert Simon's concept of bounded rationality, which is based on costly decision making and limited cognitive capabilities [136; 145]. The adaptive perspective differs from the algorithmic one in that adaptation is not about finding an exact and optimal solution, rather it involves a change in an attribute of a problem to increase fitness. Here, the adaptive agent is not assumed to be able to enumerate and explore the whole landscape, which otherwise would make adaptation an empty concept. Researchers often adopt the adaptive perspective if the intention is to show how agents actually deal with complex problems in the real world. For example, adaptive search might resemble the decision of a firm to improve the quality of its product based on changing market preferences. An example of algorithmic search might be the search for a financial investment strategy that guarantees the highest return for a given level of risk. Crucially, whether the intention is to search for optimal solutions or to adapt, a common assumption is that an agent or a population of connected agents perform exploratory moves in their fitness landscape. Next, I offer a discussion of how to model these moves.

In order to model search or adaptive behaviour, the simplest scenario would be that of a single agent or a population of independent agents located somewhere on the fitness landscape. Assuming that we are dealing with economic agents, a first question might arise as to why economic agents would move in their fitness landscape? agents can be interested in solving economic problems for different reasons; a manager might be interested in designing the best organizational structure in order to improve organizational performance, regulators may search for an optimal financial architecture to avoid failures and guarantee efficiency, and firms might be interested in developing technological innovations in order to increase profits.

As far as optimal search is the behavioural rule being modeled, one of the first fundamental procedures concerns the representation of the landscape itself. If we assume for the moment that a fitness landscape is exogenously given, then a useful analytical step consists in measuring the statistical features of the landscape [169]. Perhaps the most important and basic feature is that of local optima. Given a landscape S, A point $x \in S$ is a local optimum if, given the neighbourhood $\mathcal{N}(x)$, of x and a fitness function f:

$$\forall n \in \mathcal{N}(x), f(x) - f(n) \geq 0$$

The number or frequency of local optima in a fitness landscape is often referred to as the *modality* of a fitness landscape. A related and important quantity is that of landscape *ruggedness*. A fitness landscape with many local peaks characterised by high fitness surrounded by deep valleys with low fitness is called rugged, while landscapes with relatively small fitness differences between near solutions are called smooth. Two common measures of ruggedness are often used: *autocorrelation* and c*orrelation length*. According to reference [169], autocorrelation measures the correlation of neighbouring fitness values along a random walk (defined below) in the landscape. It can thus be calculated as the average correlation of fitness values at distance τ:

$$\rho(\tau) = \frac{E(f_t.f_{t+\tau}) - E(f_t)E(f_{t+\tau})}{Var(f_t)}$$

Where E is an expectation operator. Other measures of ruggedness include the correlation length, defined as the average distance between fitness points until they become 'uncorrelated' (see [169]), and information theory methods which relate ruggedness to the amount of information needed to describe a walk on a fitness landscape [211]. Authors in reference [169] offer a review of some of the autocorrelation measures used in literature. Another related concept to local optima is that of *basins of attractions*. A basin of attraction is an area around a local optimum that always leads to that local optimum when going uphill. Similarly, a derived concept from local optima and basins of attraction is that of fitness barriers. A fitness barrier refers to the maximum of minimum fitness values required to get from one optimum to another through an arbitrary path, given all possible pathes. Finally, the concept of fitness neutrality is used to refer to those connected areas of the fitness landscape which have equal fitness.

Moving to search rules, the most common tool in fitness landscape analysis is the notion of *walk* [119; 169]. The basic type of walks is the simple random walk. Mathematically, given a solution space S, a random walk is a sequence of solutions $(s_1, s_2, ., s_t,)$ where each s_{t+1} is chosen to be a random neighbor of s_t, $(s_t, s_{t+1}) \in S$, with probability $P_{s_t \Rightarrow s_{t+1}}$ drawn from some probability distribution. Several variations of the random walk has been proposed: A *random walk* with jumps does not require that s_t and s_{t+1} be neighbors; an *adaptive walk* assumes that a move from s_t to a neighboring s_{t+1} is accepted if s_{t+1} has higher fitness than s_t or all neighbors of s_t, an *uphill-downhill* performs an adaptive walk until it cannot attain higher fitness followed by a reverse adaptive walk until the walker cannot attain lower fitness, and in a *simulated annealing* walk the move from s_t to s_{t+1} is accepted with probability $p = \frac{1}{1+e^{-\Delta F/T}}$ where ΔF is the difference between the fitness of s_t and the fitness of s_{t+1} and T is a temperature parameter used to control for randomness in accepting a new solution. These examples illustrate the need for neighborhood function $\mathcal{N}(x)$ to get the neighbor solutions to a candidate solution x and the fitness function f(x) that returns the fitness of a solution x.

An implicit assumption in most of the walks mentioned above is the combinatorial logic, which refers to the fact that known solutions in the fitness landscape are combinations of existing components or subsystems [6; 80]. The combinatorial assumption is often assumed as a stylized fact in economic problems like technological innovation [191]. Another frequent assumption in these walks, especially when applied to an NK landscape, is that a single move consists of a one-bit mutation. The one-bit mutation rule may be adequate if we assume an NK landscape with low K, in such case the fitness landscape is smooth, and the searcher can find the local peak by performing 'hill-climbing'. When interdependencies increase (higher K), problems might arise as the system is very likely to exhibit path dependence and sub-optimal lock-in [105; 135]. The issue of path dependence is caused by the fact that the local optimum that is reached is dependent on the initial position such that a walk that starts from that initial position will allow the agent to climb up to the nearest local optimum. This means that the global optimum will be reached only when the initial position is within its basin of attraction. Once a local optimum is reached, this may act to lock an agent into that solution and prevent them from exploring other points in the landscape. The phenomenon of lock-in is perhaps the most crucial concept in landscape analysis both from the modelling and policy points of view. Search on fitness landscapes is mostly about how to achieve the highest fitness and avoid getting locked in a local optimum. For policymakers, lock-in situations reduce diversity in the system which can lead to fragility.

Whenever complementarity between the components of a problem is high, lock-ins could be seen as the result of a long evolutionary process where each evolutionary unit is optimally adjusted to the other complementary units. In this case, the system exhibits rigidity, and small changes or mutations to the existing solution can never have an effect on fitness and are likely to be eliminated by selection forces [165]. The nature of lock-ins might also depend on problem complexity (dimensionality). Author of reference [165] shows that there is a trade off between complexity and evolutionary rigidity. Simple organisms are more flexible, in that they can easily mutate; however simplicity can reduce the space of possibilities. On the other hand, complex organisms have more evolutionary possibilities, but at the cost of more rigidity. For policy making, this type of situations might justify a top-down control of the system to resolve economic problems. Whereas single searchers go by one-bit mutations, a central controller can perform multiple mutations simultaneously [105]. Given its relevance to economic problems, the theory of lock-in has received considerable attention, particularly concerning the problem of diversity [33; 105; 7].

To overcome or avoid situations of lock-in, researchers proposed a variety of possible solutions. Author of reference [165] proposed two mechanisms for protecting diversity: protectionism and subsidies. The idea of protectionism consists of protecting a new species from being eliminated by the dominating species. For example, the introduction of a new technological innovation that is not in line with the dominant design may put it under the risk of extension because markets are very unlikely to adopt it. Protecting the new technology from the pressure to adapt to the dominant design can contribute to diversity and increase the potential for better innovation. This policy will depend on several factors such as competition, network externalities, sizes of populations, or interaction with old species. Some studies use the term 'technological niches' as a key instrument to escape from a dominant design paradigm [33]. The second proposed solution by [165] to lock-ins is subsidies. This strategy involves the adoption of a different type of selection forces than existing ones. For example, an innovation may survive competition not for its high profitability when compared to others, but for its environment-friendly benefits. Another potential solution to technological lock-ins relies on the idea of flexible design [2]. This solution consists in maximising the number of paths towards the different existing local optima and increasing the robustness of the system to changing evidence, changing fitness values, and preferences. In this way, the system can be reversed even if initial steps have been taken toward a specific local optimum. Furthermore, critical factors like regulation, technological breakthroughs, firm-level strategic initiatives, and the emergence of new leading users can also contribute to the diffusion and adoption of new and diverse innovations, thus reducing the probability of lock-ins.

Although the statistical analysis of fitness landscapes can offer significant modelling power, agents in economic settings are likely to be more intelligent than following hill-climbing or simple random walk. In real life situations, agents might often use heuristics and cognitive shortcuts which they construct and dynamically change in order to adapt to their environment [138; 88; 79; 148]. For example, if we take an NK landscape with high K, then the number of local optima is high and economic agents might explore the landscape by means of heuristics which decompose the landscape into searchable sub-areas which can be searched independently or quasi-independently [158; 73]. This decision to decompose a fitness landscape can be the result of bounded rationality and limited resources. It should be noted that the decomposition of a complex problem is not a trivial task since agents with bounded rationality might not be able to find the correct or optimal level of decomposition of a problem, especially when problem complementarity is high [86]. This, in turn, might introduce a trade-off between the level of decomposition (modularization) of a problem and the quality of the solutions [146; 27]. Decomposed areas of a landscape could be related to what Dosi called a technological paradigm [27], which embodies a selection of the set of technological problems to pursue and the tools to solve them. Different technological paradigms might encourage different search heuristics. Additionally, different sub-areas might offer different return opportunities; Authors of reference [147] proposed a model with consumers endowed with heterogeneous preferences and for this reason a technological fitness landscape may become divided into different areas where each area is populated by different types of consumers, thus providing different profit opportunities.

Another line of research classifies landscape heuristics based on two dimensions as proposed in [139; 85]. The first dimension concerns *where* to search, in such case a distinction could be made between local vs distant search heuristics [115; 134]. The other dimension is *how* to search, where heuristics are classified as experiential or cognitive. As far as the 'where' dimension is concerned, an optimal scenario would involve a balance of both local and distant search. In the literature, this balance is often called exploration-exploitation and agents are assumed to pursue such balance [144]. Achieving this balance however is challenging, because too much exploitation can result in sub-optimal performance in the long run, while too much exploration is costly and eventually forgoes the benefits of exploitation in the short run. Research has been done to explore how agents achieve this balance [17; 98; 173], however, it is still an open question whether and how agents achieve such balance, especially in dynamic environments. An interesting finding is offered in reference [22], in which the authors conducted an experimental study of adaptive behaviour on rugged landscapes and found that success in search narrows down search activity to the local neighbourhood of the status quo (exploitation), whereas failure incentivizes more exploratory search (exploration). In a more general sense, factors such as - search costs, growth stage, institutional context, organisational inertia, bounded rationality, absorptive capacity, and dynamic capabilities - are among the most important determinants of the exploration-exploitation strategy. Moving to the second dimension, i.e. how to search, its first form is experiential search which can be thought of as learning by doing process [168]. This means that solutions are tried and evaluated 'on-line' and feedback from trials is used to decide on the subsequent steps. The other form is cognitive search, which involves mind representation of a problem and its solutions which are tested using 'off-line' evaluation and learning before doing.

Finally, it is worth mentioning that the process of search on fitness landscapes might require another trade-off between a centralised and decentralized search. The important question here relates to whether the search should be decentralized (self-organized), where agents are free to navigate the fitness landscape, or to some extent centralized, where a central agent can influence the direction of the search of the other agents? In presence of complex landscapes (rugged fitness landscape), the trade-off between authority and power on the one hand and the adaptive behaviour of agents on the other side becomes more relevant. In a simple computational model, authors of reference [59] show that organisational performance is enhanced when there is a balance between decentralised local coordination and authority. In their model, the authors adopted a fitness landscape model where fitness values were ranked from best to worst, and agents could perform more than single-bit mutation search and *"mutate up to all the policies under their control"*. In another paper, authors of [105] sustained that when individual agents search by performing only small incremental steps in the fitness landscape, there might be a need for a central agent to introduce more radical changes, which goes in line with what I discussed in section 4.1 about escaping lock-in situations.

Crucially, search for technological innovation in a fitness landscape is likely to be performed by a multitude of connected agents rather than a single one. In this regard, technological innovation is often perceived as a collective phenomenon. In the next section, I discuss the idea of collective technological innovation and illustrate the role of networks in such a process.

3 The process of technological innovation as a collective and network problem

Besides being a primary and outstanding example for a complex economic problem, technological innovation is increasingly stylized as a 'collective' phenomenon of interactions between a multitude of innovating agents [111; 96]. As the author of reference [172] sustained "*the idea of a lonely innovator has been definitely abandoned and replaced by a view which suggests that technological change takes place in the form of both technological and non-technological interplay between different actors, such as individuals, firms, organizations and institutions*". When technological innovation is conducted collectively rather than singly, the exploration-exploitation trade-off may be pursued by balancing individual and collective search [14; 15]. Having said that, one might ask why is it that agents decide to solve a problem collectively rather than singly? next I offer a few explanations.

3.1 The rationale behind the collective nature of technological innovation

Why would an economic agent like a firm search the landscape of a problem together with other firms rather than doing it alone and seize all the benefits? In the literature, there have been several interpretations of why agents in specific business contexts may benefit from the sharing of information/knowledge between each other. Several reasons and theories are proposed to explain the collective aspect of technological innovation and to cover all of them is beyond the scope of this thesis. However, in the next few subsections, I provide a discussion of the most important and diffused of these theories.

3.1.1 High costs of technological innovation

Among the most important reasons for technological collaboration are the high costs involved in the development of specific innovations. For example, developing a new drug or a military weapon can cost hundreds of millions of dollars, and a single firm may not be willing to incur the entire cost. For this reason, firms might form networks to share the costs of development. This, however, comes at the expense of having to share the return, therefore introducing a trade-off. This trade-off can be very important for the performance and competitiveness of a firm and therefore cannot be neglected. Several empirical studies have shown that in many situations there are advantages for market participants to share information, and those who refuse to share any knowledge might end up being excluded from the settings in which these exchanges of knowledge take place [214; 187; 101; 179; 180].

3.1.2 Low protection for innovations

Since technological innovation is a costly activity, innovating firms would decide to innovate if they know that they can seize all the benefits that derive from their investments. The literature on this topic has shown that innovating firms may not be always able to appropriate all the benefits arising from their innovations, thus firms have less incentive to carry the scientific research needed to innovate [156; 214; 202]. This situation happens if, for example, the innovations in one sector are not subject to patent protection. One possible solution to such a problem is cooperative R&D where a multitude of firms shares knowledge concerning the development of a specific technological innovation [112].

3.1.3 Technological complexity and novelty

Another reason behind the decision of market agents to engage in knowledge sharing is when the knowledge base of an industry is sophisticated so that it makes sense to collaborate by sharing the costs and benefits of complex R&D projects [171]. The development of a certain technology might require a wide range of technical skills and knowledge depth which a single firm is unlikely to have, and so collaboration offers a means to gain access to these skills. In his seminal paper, the author of [203] claims that alliances can resolve problems of complex coordination between firms and therefore results in better innovations and higher welfare. By establishing different collaborations, firms can achieve *technological diversification*, which is a term used to describe the expansion of a firm's competencies into a broader range of technological areas [31]. This, however, comes at the cost of sharing the returns. In reference [203], the author states that technological collaboration could represent a challenge for the antitrust authorities since alliances can resemble cartels, and it might be difficult to distinguish between cartels and alliances.

Another factor that contributes to the emergence of collaboration is the search for novelty. Assuming that the space of technological innovations is always expanding, then novel technologies and ideas would be continuously emerging. Established firms might not have an incentive to look for novelty if they are following a well established technological paradigm, but this is not the case with players like start-ups are known for their agility and unique innovation patterns. Therefore, these startup firms might be ideal partner for established firms to collaborate with to explore novel technologies. A well-known example is the FinTech industry which is composed of start-ups searching for novel technologies to improve financial services [194]. To be embedded in the financial system, FinTech start-ups usually establish collaboration agreements with big banks who in turn benefit from gaining knowledge about alternative innovative technologies [38].

3.1.4 Technological risks

Given the high costs and uncertainty involved in the development of new technology, firms may decide to collaborate to share the risk of innovation [76]. This has been the case with biotechnology and information technology [56].

3.1.5 Overcoming local search

Some research (e.g. [181]) argues that due to organisational and relational constraints, firms might find themselves in a limited context - both technologically and geographically- in their search for new knowledge. Distant knowledge might offer good ideas that would allow a firm to innovate, however, it might be challenging for some firms to reach beyond their existing context. In this case, an excellent solution might consist in establishing cooperation or alliance relationships with other firms which can serve as bridges to distant knowledge, allowing the a firm to overcome its contextual constraints.

3.1.6 The nature of technological knowledge

Most technological knowledge is tacit, in that it is hard to codify into explicit blueprints, and firm-specific, meaning that it is not necessarily adequate to implement immediately in another firm [176]. This makes technological knowledge hard to transfer to other firms. This difficulty can be overcome by collaboration since close linkages allow for the emergence of a common understanding which allows effective transfer of technological knowledge. The tacit nature of technological knowledge makes it hard to price, and with collaboration firms avoid the problem of pricing.

3.1.7 System-wide problems and technical standards

Among the most critical factors that can justify information sharing is the situation in which a set of roughly equivalent market agents are all trying to solve the same problem. This situation can be encountered in real life as when doctors and professors are working on the same problem or state government formulating their policies [132]. In some industries, firms might be interested in promoting technical standards throughout the system in order to facilitate functioning and communication in the system. The same holds for the transfer of best practices [199].

3.1.8 Globalization and competitive strategies

From a corporate perspective, technological collaboration could be viewed as a competitive advantage where firms try to establish the right partnerships and alliances to navigate the competitive landscape. In his book about globalisation, author of [161] argued that globalisation "*mandates alliances, makes them absolutely essential to strategy*" and finding the right international partner becomes of great importance.

3.2 Theoretical approaches to collective technological innovation

In this section, I discuss some of the theoretical approaches proposed and used in literature to model technological collaboration and knowledge transfer that involve collective and network aspects.

3.2.1 Information cascades, herding and the problem of aggregation

Seen from one angle, economic problems like technological innovation are knowledge-driven. In many cases, knowledge about a certain problem happens to be dispersed among individual agents, therefore obtaining a solution to such a problem can go beyond the inventive capacities of single individuals. For this reason, some economic theorists like Friedrich Hayek claimed that the main problem that economic agents face concerns the aggregation of the dispersed pieces of knowledge [99]. In many cases, the problem of aggregation may fail and result in market behaviours that are sub-optimal. Among the variety of models proposed to explain aggregation failures is *information cascades theory* [12; 20]. Information cascade theory illustrates how non-welfare improving behaviour (fads) can diffuse among agents when agents observe the choices of others but not the quality or success of such choices. The cascade happens when a signal travels through the system (network) which causes agents to base their adoption decision on the choice of other agents, which leads to a self-reinforcing process, reflected in phenomena like bank runs, fire selling and so on. In reference [198], the information cascade model was extended to situations in which agents can observe the performance of other agents but not the reason for that success. The market for technological innovation might be exposed to the problem of information cascades if participants decide to adopt certain technology without examining its potential success. Similarly, information cascades can lead as well to the problem of technological lock-in.

3.2.2 Imitation regime

When the market landscape is populated by many firms who are behaving as a community of simultaneously searching agents, then the practice of observing and imitating others may be an effective way of saving time and resources. Economic models of competition often assume that when an agent comes up with a profitable innovation or product, then the rest of the market will soon imitate and copy such innovation and this will bring down profits to zero.

Economic models tend to assume that all agents in the economy perfectly share knowledge, therefore, neglecting the fact that imitation has an aspect of knowledge diffusion [155]. The hypothesis of perfect knowledge distribution comes weaker if we consider that learning a new technology is instead a costly activity [130]. For this reason, it would be more realistic to assume that there are networks of observation and imitation between firms which I call here an *imitation regime*.

Imitation behaviour in economics has been a puzzling phenomenon for policymakers. On the one hand, research has shown that when innovating firms do not succeed in seizing profits from their innovations, for example, because imitators may have better complementary assets, markets don't function well [202]. On the other hand, imitators might be able to exploit other's innovations better and develop them further, resulting in a higher economic performance overall. The later situation might result beneficial if innovation is both *complementary*, meaning that each potential innovator in the market follows a different research line and *'sequential'*, meaning that each innovation is built on the previous innovations [19]. In the same paper, author of [19] makes the argument that in the software industry, which was dominated by imitation in its initial stage, benefited substantially from imitation. Another viewpoint about imitation, usually called *Negative Incentive Problem*, states that competition driven by imitation tends to harm competition and growth [155]. So far there is still no clear evidence whether imitation-driven competition enhances innovation or not.

Imitation has often been treated as a learning activity, where imitating firms start by copying others and successively developing new technologies based on the knowledge which they acquired. For example, Toyota, who entered the car industry in the 1930s, started by copying the Ford production system, and then modifying and improving it to come up with a more efficient production system which they called the *lean production system* (LPS). Interestingly, some of the features of the LPS were later copied by American firms (see chapter 6 of ref. [75]). It is often argued that many of the modern Japanese innovations were mainly based on imitating foreign technology [75]. Another notable example of imitation is IBM. A company called UNIVAC was the first business computer built by John Mauchly and John Eckert. In its early years, IBM is known for having followed a strategy based on rapid imitation by acquiring and learning the most essential technological knowledge from competitors and universities (see chapter 7 of ref. [75]). Even though the first data-processing computer introduced by IBM (the 702 model) was unreliable, successive versions (especially the 705 model) were launched in 1954, and by 1955, IBM was selling more computers than UNIVAC [30]. Similarly, first to introduce a graphical operating system was Apple in 1984. In 1985, Microsoft entered the graphical operating system market by introducing Windows 1.0. The first versions of Windows (1.0 and 2.0) were not successful, however, with the introduction of Windows 3.0 in late 1990 Microsoft realised significant success [137]

Several could be the motivations for observing the behaviours and innovations of others. In a formal model introduced in reference [24], the authors modeled the probability of an agent seeking information from another agent as a function of (1) guessing what the other agent knows; (2) estimating what the other agent knows; (3) being able to access agent's knowledge; and (4) perceiving a reasonable cost of seeking information from that agent. In organizational and corporate settings, similar requirements have been studied. For example, the concept of absorptive capacity, which refers to a firm's capacity to collect, assimilate and understand knowledge from other firms or the surrounding environment, has received considerable attention [44; 43]. Tacit knowledge and know-how could be important factors influencing the ability of a firm to absorb knowledge from another firm. Authors of reference [197] showed that networks play an essential role in knowledge transfer when knowledge is moderately complex.

3.2.3 The collective invention regime

In 1983, Robert Allen coined the term Collective Invention [3] to explain the development of blast furnace design between the iron making companies in Cleveland district in Britain. The collective invention regime described by Allen has two main features: first, firms willingly release to the public information about their innovations in order for others to acquire and learn from such innovations; Second, firms which receive the information build on such knowledge to develop new innovations further and release them in their turn to the public. In this way, technological innovations are developed by repeated interactions and feedback mechanisms. Typically, a collective invention regime is assumed to end when a dominant design emerges [162]. The collective invention regime lost its importance in the early and mid-twentieth century by the establishment of internal R&D labs in firms, but recently there has been a re-emergence of new forms of collective regimes that brought into the scene some of the aspects of collective invention [170]. The collective invention model has been later extended to incorporate networks, which was first done in reference [46]. In their paper, authors of [46] show that small world regions (in the Watts-Strogatz sense) achieve higher innovation level. Another variation of the collective invention model is called the *private-collective invention regime*, of which open source development is the most frequent example [102; 162]. The main difference between collective invention regimes and the private-collective regime is that the former terminates with the emergence of a dominant design, while the latter survives it design, and thus shows continuity in the innovation process which goes beyond traditional forms of markets and formal hierarchies and alliances [162].

3.2.4 Strategic link formation

An important branch of network science takes the name of *Strategic Link Formation* (SLF), and it deals mainly with how and why networks form and take a particular form [92]. This line of research has received particular attention in the field of economic networks [92; 108; 109] given that link formation entails an evaluation of the economic consequences of establishing new links. The main idea behind strategic link formation is that in many network settings, the decision of nodes to create connections with other nodes is determined by the choice of participating nodes, and not randomly. The main assumption in strategic link formation is that the node's decision to form a link is based on cost-benefit analysis and individual preferences that together determine the outcome. Nodes will maintain useful links and drop those which are costly. Given the strategic aspect of SLF, researchers in this field usually make use of game theory to model network formation. Concepts like efficiency and utility are used to describe which networks are optimal from an economic point of view.

Within the SLF literature, the *connections model* is considered the basic model to explain the strategic formation of links [110]. The idea of the connections model is the following: given a network, a node has direct links to other nodes to whom he/she is connected, and these links have a cost and benefit to the node. Additionally, a node can derive some benefit from friends of friends, friends of friends of friends and so on but with a decaying benefit as the path length increases. The basic connections model assumes that a node pays a cost for direct links only. Given a network G, the net gain (or utility) that a node realises is the difference between the sum of benefits that the node gets from her direct and indirect links and the cost of maintaining those links. In mathematical terms, it is:

$$u_i(G) = \sum_{j \neq i; \exists p_{ij}} \delta^{l_{ij}(G)} - d_i(G)c$$

where p_{ij} indicates a path between i and j, l_{ij} is the number of links in the shortest path linking i to j, $d_i(G)$ is the degree of i (number of links i has), c > 0 is the cost for a node of maintaining a link, and $0 < \delta < 1$ is a parameter that accounts for the fact that the utility (value) that i obtains from having a connection to j is proportional to the closeness of j to i.

Several variations of the strategic link formation model have been proposed to explain different economic and social phenomena. One variant of the SLF investigates the formation of *undirected* networks where a link between two nodes is beneficial for both parties. This is usually referred to as *two-way flow networks* [83]. Since in reality many networks are directed, i.e. the benefit flows only towards the investor of the link, another class of connection model networks was developed to model this situation and is usually called is *one-way flow networks* [83].

Another parameter in the connections models proposed in the literature is *homogeneity* with respect to values and costs. Players can either have similar or different costs of establishing new links with others and based on the level of heterogeneity , different network structures might emerge [83]. In reference [11], the authors studied a non-cooperative one-way network formation model with homogeneous players. Their study showed that if a player has increasing payoffs in the number of other players accessed and decreasing payoffs in the number of edges formed, then the strict Nash equilibrium network is either an empty network with no links or a wheel network where each player creates and receives one link. In another study, authors of [83] analysed a one-way network (directed) formation model with heterogeneous players. Players are heterogeneous in terms of the costs of establishing a link and the value gained from accessing other players. In many real-life situations, this heterogeneity is likely to arise naturally. For example, in the context of information or knowledge networks, some individuals are more interested (or invest more) in certain topics and issues and therefore have more information, which automatically makes them a valuable contact. Similarly, different individuals are likely to have different social and communication skills, and consequently, it might be natural that creating links is cheaper for some individuals as compared to others (firms who don't invest in a specific technology are likely to seek knowledge from firms who invest in it).

Strategic link formation can have important applications in the field of technological innovation. This can be mainly justified by the fact that technological innovation is mostly a collaborative phenomenon and thus the choice of the firm or innovator to collaborate with others may entail cost-benefit analysis. According to reference [90], the position and embeddedness of a firm in technological innovation alliances have a positive effect on the potential for exploration and patenting. In reference [167], the authors argued that firms collaborate with other firms to become part of a knowledge network, gain experience in their industry, and collectively use their knowledge to serve their customers in a competitive environment effectively. In a computational model, authors of reference [66] examined the role of *endogenous* network formation on search involving complex fitness landscapes. The paper showed that agents achieved better results (in terms of searching for an optimal solution) when they were allowed to adjust their network compared to the case when they are connected via a static network. In all these cases, creating the right links requires strategic thinking in order to take advantage of the knowledge network fully. Another situation where strategic links formation plays a role is in the world of startups. In many cases, startups decide to collaborate with big firms for the startup to gain access to a wide and commercial platform and for the big firm to learn from the innovations of startups. An example of this behaviour is the FinTech industry, which consists of startups who are working on improving back-end and customer-face financial services. For many FinTech firms, the way to commercialise and offer their products in the financial system is via a collaboration agreement with an established financial institution (mostly banks). Choosing the right bank to collaborate with is a strategic decision which entails costs and benefits and therefore can be treated as a strategic link formation problem.

3.2.5 Infomediators, connectors and the law of the few

In many social and economic settings, a particular form of communication may emerge where the majority of agents acquire their information from a small subset of the agents. Some refer to this group of agents as *influencers*, *connectors* or *infomediators*. Work on this phenomena started with references [116; 131], where the authors studied the effect of personal contacts and media on consumer and voting decisions concerning products, movies and fashion. Their main finding was that personal contacts play a crucial role in disseminating information which in turn influence individuals decisions and choices. Reference [131], which relied on a sample of 4000 individuals, revealed that 20 per cent of the whole sample was the source of information for the rest. In a similar study, authors of reference [68] identified 20 per cent of their sample of 1400 individuals as the main source of information about food items, household products, drugs. Similarly, research conducted on online social communities revealed similar patterns of communication. In their paper, authors of [221] investigated the Java Forum which is an online group of users who ask and answer each other's questions about the programming language Java. The study found that 55 per cent of users only asks questions, 12 per cent ask and answer questions and 13 per cent only answers questions.

In an interesting study, authors of reference [84] called this phenomenon *the law of the few*: that the majority of agents acquire most of their information from a small subset of other

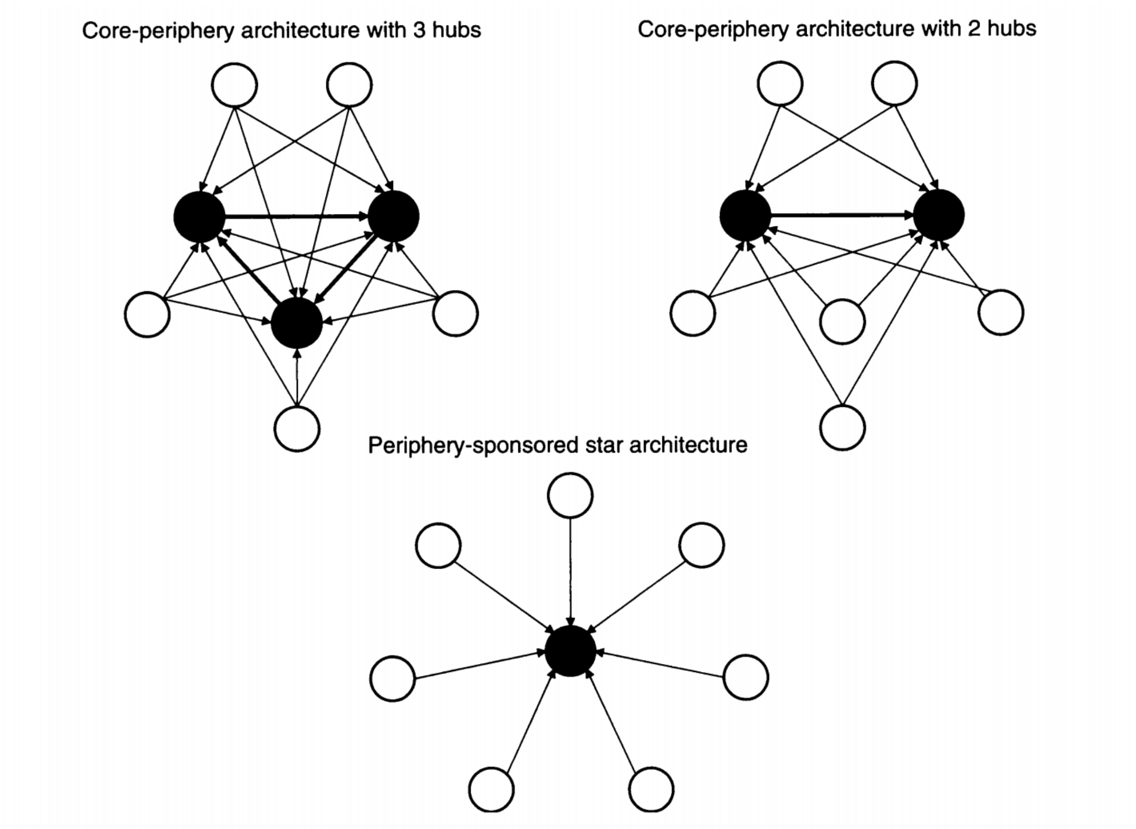


Figure 9: Examples of core-periphery networks. Core nodes are in black while periphery nodes in white. Figure from [84]

agents. In their paper, the authors asked the question of whether the law of the few can emerge in settings where there are no differences between individuals. To answer this question, they proposed a game where individuals can either acquire information personally or establish links with other individuals to access their information. The main finding of the paper was that every equilibrium of the game exhibited the law of the few. The resulting network has a *Core-Periphery Structure*; core agents acquire information personally, while periphery agents acquire information by using their links with the core. Figure 9 illustrates some configurations of core-periphery networks. In networks science, a famous mechanism that explains the emergence of networks with a few hubs is the so called *preferential attachment* [13]. In few words, a preferential attachment mechanism is one where a node in the network receives new links based on how many link it has already. In this case, nodes which have already a large number of links have higher probability of receiving new links that nodes with fewer links. From an economic perspective, there could be several factors that might favour a law of the few between firms. Perhaps the most crucial reason is resource limitations, where firms with more resources are in a position to acquire more knowledge. Another factor might be the experience. For example, authors of reference [26] sustained that the knowledge base of firms is path-dependent; therefore older firms can be in a better position that new incumbent ones to acquire knowledge.

3.2.6 Open innovation

In 2006, Henry Chesbrough, a professor at UC Berkeley's Haas Business School coined the term *open innovation* in his book 'Open innovation: The new imperative for creating and profiting from technology'. Before becoming a professor, Chesbrough worked as a manager in the computer disk drive industry in Silicon Valley where he noticed that there was no flow of ideas and advice from academia coming to the industry. For this reason, Chesbrough decided to move to the academic world to work on this gap between industry and academia. To realise his ambition, Chesbrough theorised the idea of open innovation which he defines as:

> *"open innovation is a paradigm that assumes that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as firms look to advance their technology"*
>
> - [37]

Conceptually, the open innovation regime is a distributed, participatory, decentralised approach to innovation, based on the fact that valuable knowledge is disseminated, and no company, no matter how old or how big, could innovate effectively on its own. Open innovation has two aspects. The most recognised of these aspects is the 'outside in' aspect, where external ideas and innovations are brought into the innovation process of a firm. The other facet is the 'inside out' part, where un- and under-utilised knowledge and ideas are brought from the firm to the outside world where they can be incorporated into other firms' innovation process.

A key part of the open innovation regime is the business model. To know what to look for outside and what to transfer to the outside will depend on a firm's business model. The ideas to bring from outside should be those which fit the business model, while the internal ideas and knowledge that don't fit the business model are likely to go to the outside.

Open innovation is not like a collective invention or private-collective invention regimes in that the business model for innovation in an open innovation regime is a crucial part. Thus open innovation is not to be confused with innovations like open source. On the other hand, open innovation is not to be confused with supply chain management, since the actors involved in an open innovation regime fall outside supply chains (such as universities and individuals), and these actors can be influenced but not managed or directed. Finally, open innovation is to be distinguished from user innovation. Although the user is an essential actor in the open innovation regime, so are universities, venture capitalists, startups, and corporate R&D.

The question of whether open innovation regimes are more useful than closed ones has received considerable attention [69; 49], and remains one of the main areas which requires further investigation.

3.2.7 The concept of innovation networks

In many research studies, the concept of network is used to refer to a group of innovating firms who are working together on developing technological innovations, in such case the term *innovation network* is often employed. Networks of innovators have been an important subject of study in the last few decades. In a special issue of the Research Policy dedicated to networks of innovators, the authors of reference [52] claim that among others, network of innovators might be *"supplier-user networks, networks of pioneers and adopters within the same industry, regional inter-industrial networks, international strategic technological alliances in new technologies, and professional inter-organizational networks that develop and promote a new technology"*. Innovation networks are increasingly becoming an integral part of economic systems. One estimate made by reference [95] shows that between 2002 and 2011, firms around the globe formed close to 42,000 alliance relationships.

Generally, innovation networks can be distinguished from other types of networks in that they consist in loose, informal, implicit, easily decomposable, and recombinable systems of relationships, although those which succeed can last for long period [52]. It should be noticed that although innovation networks can involve legal contracts and relationships, informal links can be equally important, especially because of the tacit nature of technological knowledge [174].

In many cases, an innovation network would emerge if it brings benefits to the inter-firm system that are partly external to single firms. In other words, for innovation networks to develop, they have to guarantee positive-sum game, where some players may lose some of the time, but a majority of players are winners most of the time. Alliance networks is perhaps the most suitable example in this context. Assuming that technological innovation is combinatorial (new combinations of existing elements), an innovation network might act as a mechanism through which one firm combines its resources with those of another firm in order to develop a technological innovation. In addition to that, innovation networks can act as a channel for firms to establish standards and public norms of the new market [205].

The flows of information in innovation networks usually assume the form of technological transactions. Examples of technological transactions include technological knowledge transfer, innovation ventures, technology adoption, or technological adaptation. Technological transactions can involve costs which can vary according to the nature of the transaction. To understand the nature of transaction costs involved in innovation network one can build on reference[203], in which the author illustrated that technologies share a number of common features including: uncertainty (in that it involves navigating a wide space of solutions), path dependency (in that technologies evolve along an established paradigm), cumulative development (especially when technologies evolve within a technological paradigm), inter-relatedness (especially with other technologies, complementary assets, and users), tacitness (in that it is difficult to codify and articulate), and inappropriability (in that it is often hard to establish which parts of the technological innovation can be legally protected). According to the extent to which these features are present in a technological innovation, one might apply a transaction costs approach to study the dynamics of technological transactions.

3.2.8 Networks and institutional isomorphism

So far, I have quoted only arguments stating that networks are a positive factor in the economy and the development of technological innovations. In some instances, this might not be the case. I borrow the following example from reference [175] to illustrate the point: in [175], the authors examined the role of inter-organisational networks in the diffusion of computer-aided production management (CAPM) innovation in the UK industrial sector during the 1980s. The study found that engagement in inter-organisational networks was an essential factor in the diffusion of CAPM since it allowed potential adopters to obtain information about the new technology. However, it was also found that the information which adopters obtained from networks was used to reinforce the image of technology suppliers. This implied that firms were led by the network to adopt an innovation not because it has higher performance, but due to mimetic and normative processes that influence the adoption process. According to the authors, the diffusion of CAPM can be an example of *'institutional isomorphism'*. This finding shows that networks are not always a positive factor in the process of technological innovation.

In seminal paper [55], Powell and DiMaggio coined the term 'institutional isomorphism' which refers to the process through which institutions, which face the same environmental conditions, resemble each other even if they evolve from different positions. The resulting similarity between organisations can be the result of the imitation of independent development under similar constraints. For example, a firm might decide to enter an established industry but wants to adopt a different business model; The theory of isomorphism suggests that this firm, once embedded in the existing inter-firm network, will end up locked into the system and therefore will be heavily influenced by the network context. The diffusion or evolution of technological innovation through institutional isomorphism can lead to suboptimal solutions which cannot be explained by the efficient-choice perspective that dominates the innovation literature [177].

There are three mechanisms of institutional isomorphic change. The first is a coercive isomorphism, which derives mainly from political power and *results from both formal and informal pressures exerted on organisations by other organisations upon which they are dependent and by cultural expectations in the society within which organisations function* [55, p 15]. According to Powell and DiMaggio *"organizations are increasingly homogeneous within given domains and increasingly organised around rituals of conformity to wider organizations"*. The second type of isomorphism is *mimetic isomorphism*, which results from *"standard responses to uncertainty"*. When faced with uncertainty, organisations are more likely to imitate each other. Powell and DiMaggio state that *"when organizational technologies are poorly understood, when goals are ambiguous, or when the environment creates symbolic uncertainty, organizations may model themselves on other organizations"*. Organisational models can be diffused through employee migration or by consulting firms. The third isomorphic change is normative pressure, which derives mainly from professionalisation. Powell and DiMaggio define professionalisation as *"the collective struggle of members of an occupation to define the conditions and methods of their work,..., and to establish a cognitive base and legitimation for their occupational autonomy"*. There are two crucial aspects of professionalisation that can lead to isomorphism. The first is the result of formal education and legitimisation in a *"cognitive base produced by university specialists"*. People from the same educational background will very likely approach a problem in much the same way. The second is the increasing role of professional and inter-organisational networks that span organisations and through which new ideas and models diffuse rapidly. The similarities caused by these professionalisation mechanisms allow firms to interact with each other more efficiently and to strengthen legitimacy among organisations. The more firms are similar, the more what they do will look legitimate, even if it is be sub-optimal.

Until now, the thesis has focused on reviewing the theoretical aspects of the problem of technological innovation and the collective nature of its development. In the next chapter, I will report on the result of an evolutionary and to some extent idealised agent-based model to illustrate the role of network average path length, edge direction, and degree-heterogeneity on the performance of a system of innovating firms. I note that the results presented in chapter 4 constitute the content of join paper with Rosario Mantegna which is currently under revision by the Journal of Complex Networks. In chapter 5, I present the results which I have obtained so far from a historical agent-based model that depicts the development of Financial Risk Management (FRM) as a collective phenomenon. I shall bring to the attention of the reader that given the current stage of my research, these models are purely explanatory models rather than predictive.

4 Path length, Degree-Heterogeneity, and Edge Directionality: An Agent-Based Model

Analysing the role of networks in the process of technological innovations can be approached from several angles. The first angle concerns the network metrics that the researcher wants to examine. Examples are degree centrality, betweenness centrality, network embeddedness, weak or strong ties, and so on. The second angle concerns the decision of whether the researcher is interested in the performance at the node-level or network level. It is possible to study performance from a micro perspective, where the researcher might be interested in understanding the role of network position on the performance of the the single node. On the other hand, it is possible to study the effect of macro-level network features like average degree or average betweenness on the overall performance of the network. The model and results presented in this chapter and chapter 5 are based on the macro-perspective, where the main metrices concern degree, betweenness, and closeness centralities, clustering, and finally network constraint. These metrics are discussed in the next sections.

One particular line of research on complex problem solving by collectives has concluded that the average path length of networks can have a relevant effect on the performance of the collective [14; 151; 132]. From a topological point of view, average path length correlates with the speed of diffusion; Networks with short average path length allow for information to flow fast. The opposite case is with long average path length networks which disseminate information slowly. In reference [132], the authors argued that in situations involving collective problem solving, long average path length networks performed better than short average path length ones in the long term but not in the short term. The explanation for this is that short average path length networks circulate information about immediate solutions quickly and the system is likely to experience an early convergence to low-quality solutions. Long average path length networks, given their slow rate of information dissemination, allow agents to explore a wider variety of solutions and discover better ones. In another experimental study, authors of reference [151] show that short average path length networks perform better than long average path length ones. The study showed that the reason for this was that agents were able to rationally adapt their search strategies when they receive information more rapidly. Building on this, authors in reference [14] show that the performance of short and long average path length networks depends on the search strategy employed by the nodes. As far as technological innovation is concerned, there is evidence that shorter path length correlates positively with system-level technological innovation [74].

Although these studies have spurred an interesting debate on whether network average path length can improve collective performance, little has been done regarding the role of other network features like degree-heterogeneity, edge direction, as well as different interaction rules. Do heterogeneous networks facilitate technological innovation in a collective innovation context? Under what conditions? Does edge direction affect performance? Does the interaction rule between agents affect the outcome of innovation? To answer these questions I constructed and simulated an agent-based model to understand the role of average path length, degree heterogeneity, and edge directionality on the system-wide average performance. To do this, I simulated a network of searching firms on a fitness landscape using networks which varied in terms of average path length, degree heterogeneity and edge directionality.

The chapter proceeds as follows: First I give a schematic description of my model, network structure, details of the search dynamics and observation rules of agents. Agents in my model are firms and all firms are assumed to face the same initial problem: finding the optimal technological innovation with the highest fitness in a given technological space (or fitness landscape) without knowing the structure of this technological space. The model considers only a generic innovation without making it a case study. Solutions are called technological innovations (or simply innovations), and the fitness landscape represents the space of all possible technological innovations. In the model setting, the fitness landscape has one global peak, and the optimal technological innovation would be to find the solution with the highest fitness possible. I assume that there is a network of relationships between firms such that neighbouring firms (two linked nodes) can observe the innovations according to their position in the network. A node observing another node means that the first is interested in understanding and eventually imitating the innovation of the other node. Specifically, for undirected networks, the observation is bi-directional whereas for the directed networks the observation will be possible according to the direction of the arc (or arcs) connecting the two nodes (uni-directional if the link points solely from A to B, and bi-directional if A points to B and B points to A). I assume that a directed link from firm A to firm B means that firm A is interested in observing and eventually imitating the innovation that firm B has found, but firm B is not interested in observing firm A. The fitness of technological innovation is assumed to be related to the profitability of an innovation such that the higher the fitness, the higher the expected profits.

4.1 The Fitness landscape

As I illustrated in section 2.3.2, a variety of fitness landscape models are used to model complex problems. Although these models have been widely used in the description of complex systems, I have decided to follow similar steps along the line suggested by [151]. In reference [151], the authors constructed a fitness landscape having in mind not the goal of replicating standard fitness landscapes like the NK or random field models; instead, they aimed at capturing more qualitative features of real-world problems. Following this logic, the feature which I want to capture is the presence of multiple local peaks and one global optimum. I note that the focus is not on the fact that there is a unique global optimum, but instead I aim to create a context where there is one solution that is significantly better than every other solution to render optimal search a challenging task that requires extended exploration. To this end, the landscape is constructed in such a way to have many peaks with values between 100 and 5000 and a single peak with fitness value of 10000 which corresponds to the global optimum. In this way, the resulting landscape, which has a global peak, is non-trivial to navigate since it requires a lot of effort to find the optimal (global) innovation. I note that firms are assumed to lack information about the structure of the landscape and the existence of a global optimum. For the sake of simplicity, I assume that the fitness of a technological innovation is related to how much profits is guarantees in return, i.e. the higher is the fitness the more profitable is the innovation.

The space of technological innovations (fitness landscape) is modelled as a 2-dimensional regular grid of size 1000 x 1000. The assignment of fitness values to the grid points is done in four stages:

1- The grid is divided into 100 x 100 smaller and equally sized blocks, resulting in a total of 100 blocks of size 100x100.

2- Evenly spaced values within the interval [100,5000] are generated with spacing between values of 50. The result is a list of 99 numbers, to which I add the value 10000 that will be used as a global optimum, i.e. [100,150,200,250,...5000,10000].

3 -To each of the blocks obtained in step 1, I assign randomly and without replacement a number from the generated values in step 2 to the point at the centre of each block.

4 - I assign values to all other coordinates inside each block by following an iterative process where nodes of distance 1 from the central point of a block are assigned fitness values equal to a fraction of the fitness assigned to the center (see step 3) , nodes at distance 2 are assigned a smaller fraction of the fitness of the central node and so on , resulting in a fitness landscapes which has many peaks and valleys and high variance in the fitness values of local peaks.

A schematic illustration of the fitness landscape is shown in figure 10. I point out that the fitness landscape used in this thesis is fixed, i.e. does not expand, and static, i.e. the fitness values of innovations do not change over time. This will be considered as limitations of the study and added as possible extensions to the model.

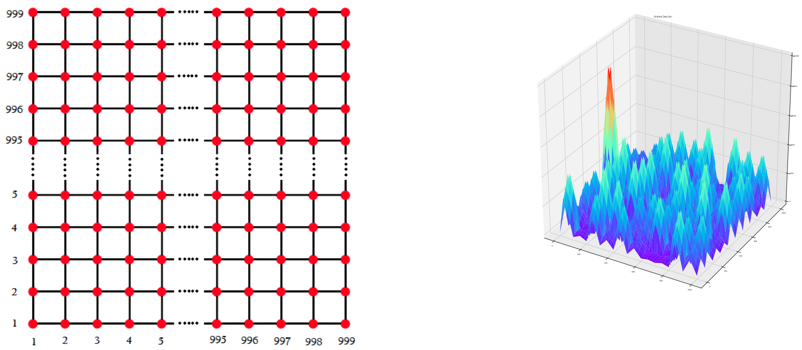


Figure 10: The Fitness landscape. The fitness landscape is generated by dividing a 1000 x 1000 grid (left panel) into 100 blocks of size 100 x 100 and assigning them different fitness values. An example of resulting landscape is shown in the right panel of the figure.

4.2 Network configuration

The key question of the model in this section is the impact of average path length, degree heterogeneity edge direction, and observation probability on the average performance of the innovation system. For this reason, firms are located on a network which defines the interaction structure. I consider only networks that have a single strongly-connected component for directed networks (i.e. there is at least one path connecting each pair of nodes in the network), and weakly-connected component for directed networks (i.e. there is at least one path that goes either from node A to node B or from node B to node A). In the simulations presented, I examine two basic types of network: a family of undirected networks which vary in terms of average path length, and a family of directed networks which vary in terms of heterogeneity of degree. In the undirected networks, observation can be done in two ways, meaning that node A can observe neighbouring node B and vice versa. In the directed networks observation is one way, meaning that if A has a directed link towards B and B doesn't have a directed link to A then only A can observe B but not vice-versa. [83].

In the investigation of the role of network average path length, I consider eight networks that are analogue to those proposed in [151]. Specifically, I generate eight networks having 64 nodes and a fixed degree of 3 for each node, but with different structural properties which I illustrate next:

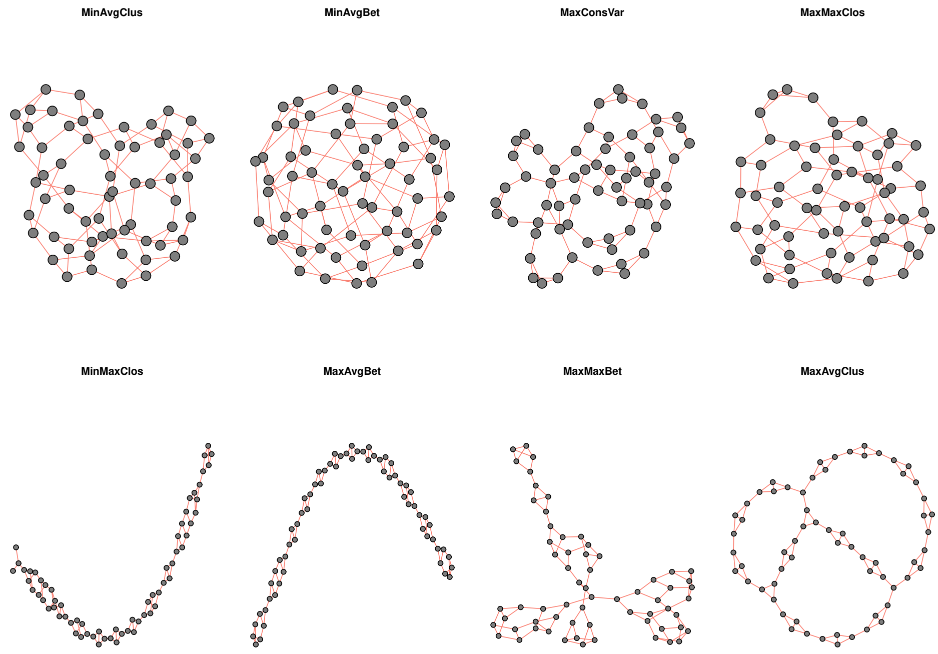


Figure 11: The networks used in simulations investigating the role of average path length. The panels of the upper row show networks which have high average path. From left to right they are characterised by minimised average clustering, minimised average betweenness, maximised variance in constraint and maximised maximum closeness. The four networks in the panels of the lower row have low average path length in terms of path length. From left to right they are characterised by minimised maximum closeness, maximised average betweenness, maximised maximum betweenness and maximised average clustering. Details of the generation of the networks will follow later.

- **Betweenness Centrality** Betweenness centrality [77] can be perceived as a measure of the amount of information that flows between nodes through aother node. Formally, let α_{ij} be the number of shortest paths from i to j and let $\alpha_{ij}(k)$ be the number of shortest paths from i to j that go through k, then the betweenness centrality of k is

$$C_{betweenness}(k) = \sum_{i \neq k \neq j \in V} \frac{\alpha_{ij}(k)}{\alpha_{ij}}$$

- **Closeness** Closeness centrality [78] measures how reachable a node is from the other nodes in a network. More formally, it can be calculated as the reciprocal of the average length of the shortest path between a node and the rest of the network

$$C_{closeness}(k) = \frac{n-1}{\sum_{v \in V} d(k, v)}$$

where n is the number of nodes in the network, and d(k,v) is the length of the shortest path from k to v.

- **Clustering Coefficient** Clustering coefficient is a measure of how connected are the neighbors of a node [183]. In formal language

$$cc(k) = \frac{|\{(i,j)|i \in \Gamma(k)\, and\, j \in \Gamma(k)\}|}{\binom{deg(k)}{2}}$$

where $\Gamma(k)$ is the set of nodes that k is connected to and deg(k) is the degree of node k.

- **Network constraint** network constraint, developed in reference [29], capture the extent to which a node acts as a bridge between different groups in the network. Burt defined network constraint of node k as follows:

$$nc(k) = \frac{1}{deg(k)^2} \sum_{v\in\Gamma(k)} \left(p_{kv} + \sum_{w\in\Gamma(k), w\neq v} p_{kw}p_{wv} \right)^2$$

where p_{kv} represents the fraction of attention (often measured as link weight) that node k gives to node v. The sum $\sum_{w\in\Gamma(k),w\neq v} p_{kw}p_{wv}$ denotes the total fraction of (indirected) attention that k gives to v going through an intermediary w. If the sum of direct and indirect attention that k gives to v is high, then k is said to waste effort giving attention to v. Smaller value of network constraint is considered to be better because it allows for more freedom in exploiting ones links. When network constraint is high, then the neighbors of a node are densely connected between each other, therefore limiting the opportunities of a node to access unique information that others don't have access to [29]. It should be noted that the equation above is valid for directed and weighted network. In this thesis I used the version for undirected and unweighted network which is as follows:

$$nc(k) = \frac{1}{deg(k)^2} \sum_{v\in\Gamma(k)} \left(1 + \sum_{w\in\Gamma(k), w\neq v} p_{wv} \right)^2$$

In both cases, the measure results minimised when none of k's neighbours is connected with other neighbours.

It has been shown that the above network average attributes influence the average path length of the network, which is used as an indicator of average path length [151; 132]. Therefore, by adjusting these network attributes, I am able to obtain networks of the same size, average degree, and the number of edges but with different levels of average path length.

In constructing the networks, I use the following algorithm. Firstly I create random networks with 64 nodes all of them with fixed degree 3, therefore with a number of edges of 96. Secondly, for a sufficiently large number of iterations, I choose randomly two edges per iteration and perform a degree preserving double edge swap. A double edge swap deletes two randomly chosen edges x-y and z-w and creates the new edges x-z and y-w. With this rewiring approach besides keeping the node degree fixed, I ensure that the graph remains connected (or weakly connected for the directed networks) by accepting rewiring that preserves the network connectedness. During my search for a network characterised by the maximum value of a given network indicator only rewiring procedures that increased or decreased one of the properties of the four network metrics mentioned above are accepted. The properties examined are the average value, the variance, the maximum value and the minimum value. For example, the network with "Maximized Maximum Closeness" is the one where the most central node (in closeness terms) is as close to other nodes as possible. So to obtain the network with maximized maximum closeness, I start with a random network with fixed degree of 3. Subsequently, for a total of one million steps, I obtain the value of closeness centrality for the node with the highest closeness centrality, perform a double-edge swap, and check the closeness value of the node with the highest closeness value after the swap. If the maximum closeness value increases and the network remains connected, I accept the swap, otherwise I undo the swap. With this procedure, I control the average path length of the network, obtaining eight networks analogue to those investigated in [151]. Figure 11 shows the undirected networks used in the simulations of our agent-based model. The parameters of these networks are summarised in Table 1.

To examine the role of degree heterogeneity and one-way (directed) observation, I also generate four networks with 64 nodes, and a total number of directed edges equals to 96. The

	AvgBet	MinBet	MaxBet	AvgClos	MaxClos	MinClos	AvgClus	VarConst
MinAvgBet	0.04	0.04	0.05	0.26	0.27	0.26	0.00	0.00
MinAvgClus	0.05	0.02	0.10	0.23	0.25	0.20	0.00	0.00
MaxMaxClos	0.05	0.01	0.08	0.24	0.27	0.19	0.03	0.00
MaxConsVar	0.06	0.00	0.17	0.20	0.25	0.16	0.22	0.02
MaxAvgClus	0.12	0.03	0.34	0.12	0.15	0.10	0.47	0.01
MaxMaxBet	0.11	0.00	0.67	0.13	0.19	0.09	0.12	0.01
MinMaxClos	0.22	0.00	0.51	0.07	0.09	0.05	0.19	0.02
MaxAvgBet	0.24	0.00	0.51	0.06	0.08	0.04	0.41	0.01

Table 1: Structural properties of communication networks used in the thesis

procedure used in the generation of these networks is the following: first I fix the number of hubs (call it H) I want to have in the network, and I connect them by directed links with probability 0.5. Secondly, with probability 0.95, I assign a directed link from any of leaf nodes to a randomly chosen hub, otherwise, with probability 0.05 a link is assigned to a randomly selected node of the periphery of the network. During the generation of the network, I make sure that the network is weakly connected by accepting link assignments that preserve connectivity, and I impose a condition where nodes with degree k can have an outgoing link only towards nodes with a degree higher than or equal k. For every value of H, I generate many realizations of these networks. The final choice is based on the value of the degree variance. As shown in figure 12, the four selected networks present different levels of degree homogeneity. It could be noticed that by increasing degree heterogeneity, the resulting networks resemble networks with a core-periphery structure where the core is densely connected and its nodes have a high degree while the periphery of the is made of sparsely connected nodes which are mostly linked to the core [48]. Figure 12 presents the four directed networks used in the paper.

4.3 Research design

Each simulation is structured as follows: I assign random initial solutions to the 64 firms in the fitness landscape. This guarantees that the system has a high level of initial heterogeneity in the adopted solutions. Solutions are labelled in the form of 2-tuples (A,B), where A and B can assume all integer values between 1-1000, generating a two dimensional grid of 1000 x 1000. For example, one solution could be (10,20), another (100,500), and so on. I then systematically vary the observation probability p, for which I consider the following values linearly spaced

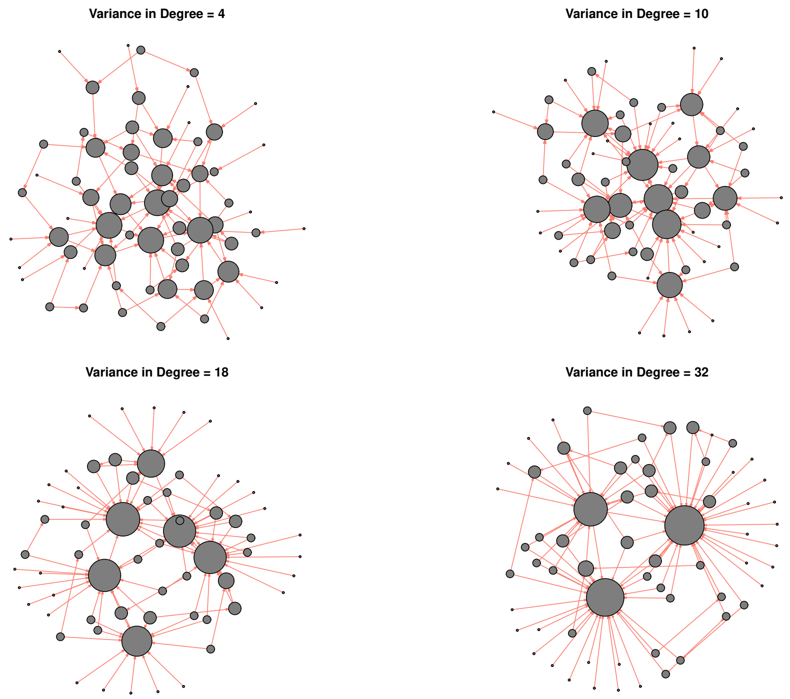


Figure 12: The networks used in the simulations of firms located in directed and heterogeneous networks. By increasing the variance of degree the network assumes a structure progressively closer to a core-periphery network.

[0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]. At each round of the simulation, I assumed that firms choose randomly between two strategies: (a) with probability p (observation probability) the firm observes the innovation of a randomly chosen neighbouring firm, and if the fitness value of the innovation of the observed firm is higher, the observing firm will move on the position of the fitness landscape of the neighbouring firm otherwise it will stay in its position [1]; (b) With probability $1-p$ (probability of autonomous search), the firm conducts independent search by visiting a random neighboring innovation on the fitness landscape. The independent search by a firm will be treated as a local exploitation process. For example, if the actual position in the fitness landscape of a firm is (50,50), then the possible neighboring innovations are (51,50),(49,50),(50,49),(50,51) [2]. If the fitness value of the newly visited innovation (F_{New}) is greater than the fitness of the previous innovation F_{Old}, then the firm will move on the fitness landscape to the new innovation with probability compared to parameter $Z = 1 - \exp^{\frac{-\Delta E}{T}}$,

[1] I note that the model implicitly assumes that knowledge about the fitness of observed neighbors has moderate level of complexity such that it can flow easily from one firm to another [197]

[2] In technological innovation studies, the behaviour of local search for new innovations, i.e., trying combinations of existing innovations which are to somehow similar to what an agent already use, is treated as a stylized fact. see [191] for a discussion of stylized facts about technological innovation

CEU eTD Collection

where ΔE is simply F_{New} - F_{Old} and T is a noise parameter that controls for randomness in accepting the new innovations. The parameter Z can assume a max of 1 and any value below (including negative values). This way, when a firm chooses a new position on the landscape, it evaluates Z and it is then compared to a random probability obtained from the interval [0,1]. If Z is higher than the random probability, then the new solution is accepted. If Z is lower or Z is <0 then the new solution is not accepted. The higher is the value of T the less noise there will be in that a larger number of visited solutions will be accepted[3]. I report the results obtained for $T = 300$, but I also run simulations with $T = 100$ and $T = 10$. I note that when $p = 0$, then firms are searching independently and don't observe each other, which is the equivalent of having no network in which observation can take place. On the other hand, when $p = 1$, then only observation takes place and search on the fitness landscape happens by observing neighboring firms and moving to their location if they have fitter solutions. Observing the solution of a neighboring firm will be treated as a process of exploration which can be distant from the current location of the firm. Each simulation is run by performing 1000 rounds with asynchronous update (meaning that all nodes must finish their moves at eath rouund before starting the next round). Results are reported by averaging the averages across 200 simulations. For each simulation, I use a different set of initial conditions of firms (positions in the fitness landscape) and these 200 sets of initial conditions are the same for all networks.

The assumption of observing neighbouring firms is justified by the fact that all firms face the same initial problem which requires them to search for an innovation (a solution) in the fitness landscape. In settings where agents can observe one another's choices, it would be a rational decision to learn from one another [82; 21; 212]. It is worth noting that in this model I am not concerned with the costs of observation or absorptive capacity of firms (see [46]), instead I focus on the role of firm network structure on performance. The performance will be measured as the average fitness of the innovations of all firms at one timestep and by the number of firms who find the global peak. Next, I turn to illustrate the model in more detail.

[3] I note that the idea of the parameter $1 - \exp^{\frac{-\Delta E}{T}}$ is inspired by the simulated annealing optimization method [128]. I note however that in simulated annealing the parameter T, which refers to Temperature, assumes the opposite role in that the higher it is the more noise there is.

4.4 Statistical quantities

For the purpose of this model, I am mainly interested in in two statistical quantities: average performance, and the average number of times the global optima was found. Denote with V(G) the set of all nodes in a network G, with $(p^1_{i,x}, p^2_{i,x})$ the position of a node i at round x, and R the number of realizations performed, then the system level average performance for network G at round x is calculated as:

$$Avg(G,x) = \frac{\sum_1^R \sum_{\forall i \in V(G)} fit(p^1_{i,x}, p^2_{i,x})}{R * |V(G)|}$$

Where fit() returns the fitness value assigned to one solution as already explained in section 4.1. The second statistical quantity I am interested in is the number of times the firms found the global optima at any given round (I shall called it Glob(G,x)):

$$Glob(G,x) = |\{i | i \in G \, and \, fit(p^1_{i,x}, p^2_{i,x}) = 10000\}|$$

4.5 Results

4.5.1 short average path length networks perform better than long average path length networks

As a first result, the simulations show that short average path length networks achieve higher average fitness than long average path length ones in the short-medium term, but long average path length networks do marginally outperform in the long run for all observation probabilities p greater than 0. Figure 13 plots the average performance over 200 simulations of the short and long average path length networks over time for one value of the observation probability ($p = 0.3$). Figure 13 shows a clear pattern. When information circulates faster, firms can get information about better local optima faster and at an earlier stage than when they are located in long average path length networks.

By rapidly circulating information about the best innovations, short average path length

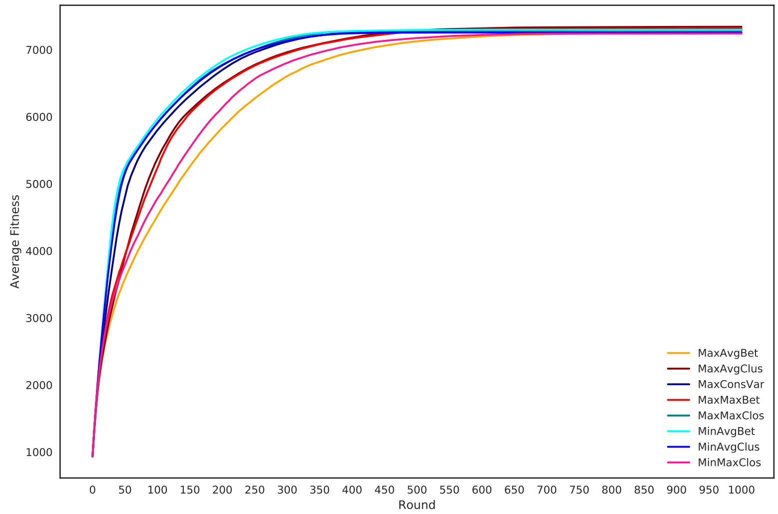


Figure 13: Average fitness for all eight undirected observation networks. The observation probability is $p = 0.3$. Short average path length networks are blue lines and long average path length ones are red lines. The value of the noise parameter T is 300

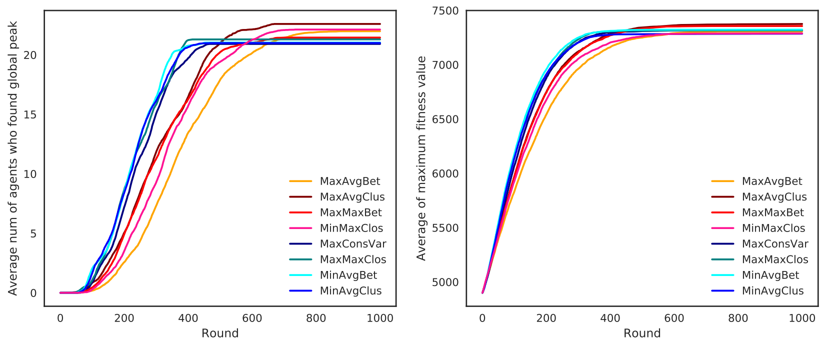


Figure 14: Simulations of firms' search located in all eight undirected observation networks. The observation probability is $p = 0.3$. Short average path length networks have colours in the blue region of the colour spectrum whereas long average path length ones have colours in the red region of the colour spectrum. Left panel: average number of agents finding the global peak. Right panel: average value of the maximum fitness value discovered by firms. The value of the noise parameter T is 300

networks allow firms to position themselves close to high fitness solutions quicker than in the case of long average path length networks. However, given that the landscape which I use requires an extended search to find the global peak, short average path length networks are more likely to experience a convergence to sub-optimal solutions. On the other hand, long average path length networks do not circulate information quickly given the higher average path length and therefore they have more time to conduct an extended search which would increase the chance of finding the global peak more often but at a later stage. Figure 14 shows this behaviour clearly. In the left panel, I report the average over 200 simulations of the number of agents who are located at the local peak at each round, and on the right panel, I show the average of the maximum fitness value attained by a firm (the best performer) at each round. From figure 14, it could be seen that short average path length networks allow firms to find the global peak more often in the short-medium term, but in the long run, more agents on average find the global peak in long average path length networks. I note that after checking the fitness values attained by firms in the last round (round 1000), it was found that approximately half the time (100 simulations) most firms were at the global peak, while in the other half of simulations no firm could find the global optimum. This can be attributed to the fact that some initial conditions are more favourable for finding the global peak. Similar behaviour was observed for the maximum fitness per round as shown in the right panel of figure 14.

Finally, results show that the performance at the system level improves on average when we I decrease the value of the parameter T. For T=100 and T=10, all networks achieve better average fitness and the global peak is found more often. This can be interpreted by considering that lower values of T allow autonomous searchers (which happens with probability 1-p) to accept new solutions more often than in the presence of higher values of T. When T is high then the average performance is mainly driven by firms observing each other and therefore the role of the network structure is more important. On the other hand, when T is low, then the autonomous search contributes more to the average performance and the role of the network becomes less pronounced. One would expect that in real life firms have limited resources and therefore it is unlikely that a firm invests heavily in both autonomous search as well as observing other innovating firms. To achieve a balance between autonomous vs. observational search, firms may be constrained by several factors like (i) existence of technological paradigms [58] which compels firms to check the prevailing innovative trend, (ii) absorptive capacity which limits firm's understanding of other firms' solutions [220], and (iii) dynamic capabilities [204] which limit firm's ability to adapt to new solutions.

4.5.2 Collective search performs better than independent search

Regardless of network average path length, my results demonstrate that search with observation probability greater than zero produces higher average fitness in the medium and long run than autonomous search. Figure 15 plots the average performance of searches performed by firms when they are located in a network characterized by a maximized value of the maximum closeness for several different values of the observation probability. What figure 15 shows is that for moderately high values of the observation probability (for example, 0.5,0.6, and 0.7), the system achieves higher performance in the short and medium run. On the contrary, for lower values (0.1,0.2, and 0.3) the system achieves higher performance in the long run. When observation probability is moderately high, then firms will be copying each others' solutions more frequently, resulting in a quick improvement in the short and medium run, but increasing the likelihood that the system will be unable to find valuable solutions, which typically requires more exploration. This is the reason why simulations with lower observation probabilities do perform poorly at the early stage, but later they achieve higher performance which is due to the fact that they are given more time to explore the space of solutions. These simulations tell us that in order to achieve optimal performance at the system level, there is a need to guarantee a balance between how often firms observe each other and how often they search in isolation.

4.5.3 Degree heterogeneity has a negative effect on average performance

Simulations for directed networks of firms with different degree show that the increase of degree heterogeneity harms the average performance of the system. As I mentioned in section 4.2, degree heterogeneity is measured by the variance in the degree of nodes, where higher variance indicates more heterogeneous networks. Figure 16 illustrates the behaviour of the system for four directed networks (simulations are performed for observation probability $p = 0.3$). As expected, the network with degree variance 4 has by far outperformed all other networks with variance 10, 18 and 32 respectively. This behaviour shows that by introducing a heterogeneous and one-way observation structure, we reduce the sources of information available to the imitating firms, thus resulting in worse performance. The inferior performance can be attributed to the fact that the location of hubs in the technological landscape constrains

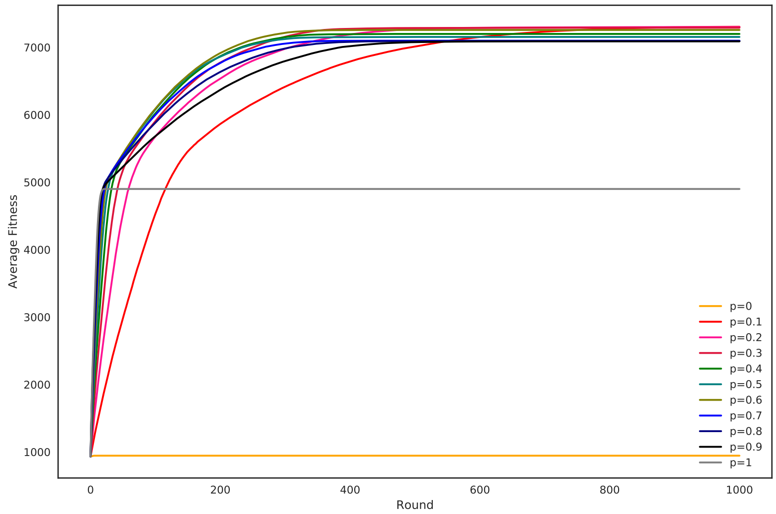


Figure 15: Average fitness for a search performed by firms located in an undirected network with a maximized value of the maximum closeness. Different lines refer to different values of the observation probability p. These values are ranging from 0 to 1 in steps of 0.1. The value of the noise parameter T is 300

their search for technological improvements and as a consequence limits the improvements of the imitating firms. If the initial position of hubs is poor (in fitness terms), then heterogeneous structures of firms with one-way observation are very likely to drive the system to perform poorly. Concerning the finding of the global peak, results show that, contrary to the family of varying average path length networks, the family of directed and heterogeneous networks are not able to find the global optimum in all cases and for all observation probabilities with $T = 300$. This, however, is not the case for lower values of the parameter T. The results show that by setting the value of T to $T = 100$, some firms in all four networks are able to find the global optimum (in average 1 or 2 times). By setting $T = 10$, the system performs better and the global optimum is found more frequently by firms located in all four networks.

It could be noticed in figure 16 that the average performance of the system is much lower for all four networks than it was with the degree regular and undirected networks examined in the previous section. To check the role of directed versus bidirectional observation, I repeated the same simulations on the four networks in figure 12 by ignoring edge direction and keeping the same network structure. Results for these simulations are shown in figure 17. Figure 17

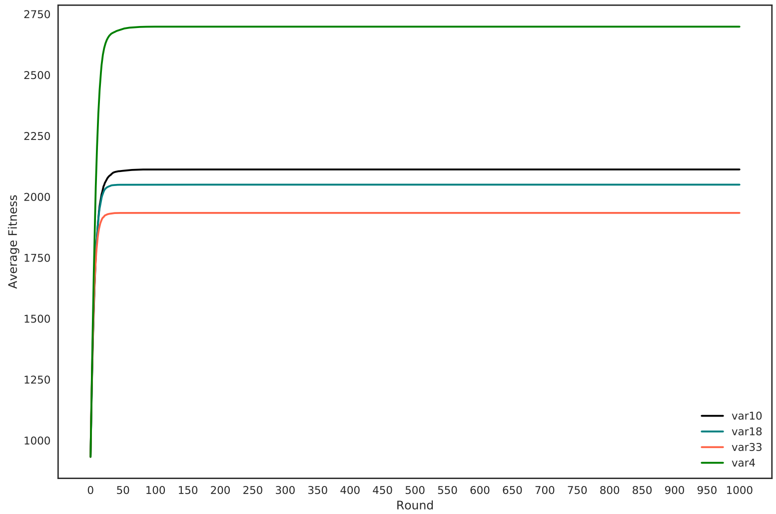


Figure 16: Average fitness for a search performed by firms located in the four networks with heterogeneity in the degree discussed in the text. Edges of the network are directed, and observations of the firms are directional. Each network is labelled by the value of the variance of the degree. Observation probability is $p = 0.3$

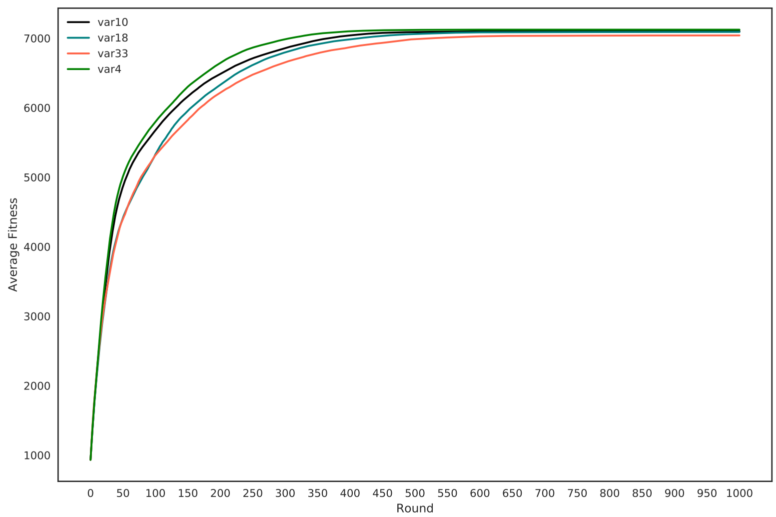


Figure 17: Average fitness for a search performed by firms located in the four networks with heterogeneity in the degree discussed in the text but ignoring edge direction. Edges of the network are undirected, and observations of the firms are bidirectional. Each network is labelled by the value of the variance of the degree. Observation probability is $p = 0.3$

shows that the average performance for all networks is higher than the case with directional observation (figure 16). When the structure of observation between firms is directional, then it is more likely that the system collectively misses improvement opportunities because of the lack of mutual observation. This raises an important question of why certain structures like directed and heterogeneous networks might form despite their lower benefit for innovation? In the real world, factors pertaining to social and economic pressures can lead to the formation of heterogeneous networks [93]. As I illustrated in section 3.2.5, this phenomenon has been called 'The Law of The Few', which is a situation where the majority of agents acquire most of their information from a small subset of other agents. One economic reason why the law of the few can emerge in firm networks relates to the role of experience [26]. The knowledge base of firms is path-dependent; therefore older firms can be in a better position that new incumbent ones. The universality of such claim, however, cannot be confirmed, and because of the importance of such theory, an interesting debate in the technological innovation literature spurred around the idea of small vs big innovators. This debate is usually referred to as the Schumpeter/Arrow debate [89]. For Joseph Schumpeter, concentrated markets with big companies have more resources to invest in R&D and therefore are in better position to innovate. On the other hand, Arrow argued that big companies may have incentives not to innovate and therefore competition should be the main driver of innovation. In his classical paper, [159] suggests that small firms may be more innovative.

Although in my thesis I am assuming that firms have the same search capability, my results highlight that in a market characterised by heterogeneous observation structure that favours the core over the periphery, the system-level performance is not optimal and more likely to experience sub-optimal lock-in.

Concerning the effect of observation probability, networks with degree heterogeneity show a less clear distinction between which observation probabilities are optimal for the system performance. The Results shows however that collective search, characterised by values of the observation probability greater than zero $(p > 0)$, always performs better than independent search ($p = 0$). However, the difference in performance between different values of the observation probability is less pronounced than the case with the undirected and regular networks examined in the previous section.

So far, I have dealt with an idealised model where all components of the model are artificially constructed. The lack of empirical case studies is still a significant gap in the literature of collective search on fitness landscapes. In the next chapter, I present a contribution to the empirical direction by building a historical agent-based model of the joint invention of Financial Risk Management constructed using empirical data.

5 Computational case study: modern market risk measurement as a collective innovation

In this section, I report on the results of a historical agent-based model constructed to provide an empirical application of the model proposed in chapter 4. According to reference [217], a historical agent-based model *"constrains parameters, interactions, and decision rules in the model in line with the specific, empirically-observable history of a particular industry. It can be interpreted as a calibration exercise with respect to unique historical traces"*. The case study presented in this chapter concerns the development of the financial risk measurement solution whose history of development followed a collective and interactive process. The model presented explains the industry-side dynamics and not the whole story. I also note that the results presented here are not the final version of the model which is still a work in progress as of the date of this thesis.

5.1 Market risk management: a short history

The object of the case study is the financial risk management (FRM) solution to measure market risk in financial institutions. Market risk is defined as *the risk of losses in on- and off-balance sheet positions arising from movements in market prices.* [47]. Market risk can result from different sources like interest rates, currency exchange rate, and equity risk factors.

The history of FRM is a history of interaction between theoretical findings, market players, and the regulatory authorities. On the theoretical side, most sources accredit the beginnings of FRM to Louis Bachelier, who was the first to use stochastic processes in the study of stock option pricing [143]. The work of Bachelier remained unknown to scholars until the period after WWII when Leonard Jimmie Savage translated the thesis of Bachelier and brought his work to the attention of Paul Samuelson. Paul Samuelson's work on random walks (and subsequently Geometric Brownian Motion) became one of the main assumptions in the market risk paradigm. In parallel with the random walk theory, the works of Markowitz on modelling risk using mean-variance analysis, Eugene Fama on efficient financial markets, William Sharpe on quantifying the worth of an asset, and Black, Scholes, and Merton on the value of risk were essential in shaping the trajectory of FRM [143]. In this thesis, I claim that financial theory is not sufficient to explain the history of FRM, whose development involves the adaptation and application of financial theory into commercial use. This is in line with the distinction between invention and innovation illustrated in section 2.1.

On the industry side, the problem of how to quantify market risk has always been driven by the fact that market risk is intrinsic to the business model of financial institutions. Additionally, financial institutions, and in particular, commercial banks, are (at least in theory) supposed to perform two primary functions: risk management and resource allocation.

The evolution of FRM went through several stages. In the years following WWII, the financial sector did not have a well-established standard for measuring financial risk. Each bank relied on their ad-hoc heuristics. One convention was to treat the maximum loss incurred after the post-world war history as a proxy for market risk [57]. However beginning from the 1970s, a wave of convergence took place among financial institutions that resulted in what I call FRM. The first stage in the convergence to FRM was repricing gap analysis (RGA), which simulated the effect of interest rate changes on interest income. RGA is easy to use and communicate, but it has limitations if managers are interested in a more precise measure of risk.

The use of duration analysis characterised the second stage. Duration analysis measures the sensitivity of a security's price to changes in interest rates and has the advantage of taking into account the effect of market rate changes to the market value of securities. Another advantage of duration analysis is that it converts a financial instrument into zero coupon equivalents, therefore creating a common unit of measurement for market risk. Duration analysis has the disadvantage of not taking into account changes in the yield curve environment, and it does not integrate correlation risk. Given these issues, a third stage followed where measures like Value at Risk (VAR) and scenario analysis were the dominant measures. The origins of Value at Risk go back to Bankers Trust who began implementing the first VAR measures to calculate what is called Risk-Adjusted Return on Capital (RAROC) [57]. The emergence of VAR models was encouraged by the shift to standard-deviation models which was also pioneered by Bankers Trust in by late 1970s. VAR models were later popularised during the 1980s and 1990s by the research team at JP Morgan [103], led by Till Guildimann. Given its rapid diffusion and simplicity of implementation, regulators adopted and encouraged the use of VAR-like and scenario analysis tools.

Finally, on the regulatory side, the primary driver of regulatory intervention in the development of FRM was the need to guarantee that financial institutions have enough capital as a buffer against losses incurred due to market risk. To determine how much capital banks should hold, regulators were also interested in a solution for market risk measurement. With a standard risk measure (or practice) in place, regulators can simplify their job and compare the risk profiles of different banks. There were several episodes of regulatory interventions in the financial market that were mostly driven by the goal of driving the market to converge to a uniform solution to the problem of market risk. For a detailed discussion of the most critical regulatory interventions in the history of FRM see [103].

5.2 Stylized facts

When constructing a historical-agent based model of an innovation or industry, the main ingredient is the set of stylized facts that characterise the history of such innovation [142].

Stylized facts are used to model the parameters and components of the agent-based model in order to produce patterns that resemble the actual history of the problem at hand. In this section, I provide five of the stylized facts about the development of FRM.

1- The extensive interaction between the public and private sector in finding the right solution ([18; 71]. As the chairman of the Federal Reserve Ben Bernanke sustained:

> *The evolution of risk management as a discipline has thus been driven by market forces on the one hand and developments in banking supervision on the other, each side operating with the other in complementary and mutually reinforcing ways. Banks and other market participants have made many of the key innovations in risk measurement and risk management, but supervisors have often helped to adapt and disseminate best practices to a broader array of financial institutions. And at times, supervisors have taken the lead, for example, by identifying emerging issues through examinations and comparisons of peer institutions or by establishing guidelines that codify evolving practices.*

2- The dominance of a few institutions in the innovation process. These institutions were mostly the largest ones who were actively investing in finding a solution to the problem of measurement and management of market risk. This stylized fact can be treated as an example of institutional isomorphism where powerful agents create an environment that encourages others to belong to it in order to conform. Evidence for institutional isomorphism can be supported by the group of thirty meeting in 1991. The group of thirty is an international body of leading financiers and academics which aims to deepen understanding of economic and financial issues and to examine consequences of decisions made in the public and private sectors related to these issues. In 1993, the G30 published the famous report in which they made it explicit that market risk measure like Value-at-risk and stress scenario are to be treated as best practice and encouraged the system to adopt them. See Appendix A for a list of banks who were part of the 1993 G30 project.

3- The dimension of the FRM problem as seen from an industrial perspective. As with any complex economic problem, finding the best risk measure entails a constrained optimization along multiple dimensions. The most critical dimensions of the problem of FRM were to find a measure that is comprehensive (i.e. deals with all or nearly all aspects of financial risk), and simple (i.e. easy to implement). Simplicity and breadth, however, resulted in being qualities where an increase in one would results in decrease in the other. As Kenneth Garbade, who worked in Bankers Trust Cross Markets Research Group, noted:

> *In view of the importance of risk assessment and capital adequacy to regulatory agencies and market participants, it is not surprising that many analysts have tried to devise procedures for computing risk and/or capital adequacy which are (a) comprehensive and (b) simple to implement. Without exception, however, those who make the effort quickly discover that the twin goals of breadth and simplicity are seemingly impossible to attain simultaneously. As a result, risk and capital adequacy formulas are either complex or of limited applicability, and are sometimes both.*

4- development of FRM lead to a situation of lock-in. By the beginning of the 1990s, the financial sector managed to agree on a final solution for market risk that involves VAR and stress scenario, and this solution was quickly generalised throughout the system. Two main factors contributed to the lock-in situation of FRM. First, in the 1990s, complimentary services were developed around FRM such as software implementation and university courses. Second, regulators adopted similar measures as their solution for evaluating capital requirements [97; 150; 71]. Adopting a uniform approach to risk measurement throughout the financial system might give rise to a form of systemic risk driven by model homogeneity. The source of the systemic risk lies the fact that a unique approach might generate a system-wide asymmetric risk profile that exposes the entire system to the same risk factors. As Bernanke sustained *a single firm may have an acceptable exposure to a particular type of risk that would be unacceptable if replicated across many firms*. Even the theory of forecasting sustain that aggregating different forecasts tend to outperform a single forecast [16; 54]. Following the crisis of 2008, measures of market risk like Value-at-Risk have been heavily criticized for being a main factor leading to the crisis. Future development in the area of market risk management will show whether it will be possible to lock-out from the Value-at-Risk solution to market risk measurement.

5- I would argue that a stylized fact of financial markets concerns the fact that financial innovations are typically not subject to legal protection ([127]); thus financial institutions have high incentive to observe and copy the innovations of other institutions. Evidence shows that during the years leading to the development of FRM, small financial institutions actively attended conferences and workshops offered by major financial institutions on the subject of risk management [57; 97].

5.3 Data collection

The data employed in this section come from two primary sources. Data on the financial network of banks is obtained from the DealSscan database. DealScan is the most comprehensive database of the international syndicated lending market. It is published by Reuters Loan Pricing Corporation (LPC) and contains information on over $2 trillion large corporate syndicated loans originated since 1981. Academic researchers and economists leverage DealScan data for investigating several problems like bank-firm relationships analysis, lending behaviours, banking structure, information asymmetries and cross-country lending activities. DealScan provides detailed information on loan contract terms, type, amount, currency, country of origination, purpose, type, borrower, and lenders. For each participant in a loan, information is available regarding its role (lead arrangers and participant lenders), and its share in the loan. Information is available both on the deal (package) level as well as on the tranche (facility) level. A deal can be made of several tranches. The data used in this thesis had 263299 facilities and 17377 distinct participating banks. The size of the deals in the database may vary from a small amount of several thousand to as much as several billion US dollars. The number of lenders in the deals and tranches varies greatly and ranges between one to more than 190 lenders. The market covers borrowing firms from more than 180 countries. The reason DealScan is chosen is that of its comprehensive coverage that includes most of the

	Degree	Betweenness	Cloneness	Clustering	Eigenvector	Frequency
Average	131	1.038766e-04	0.391831	0.728495	0030180	84
Min	1	0.000000e+00	0.000069	0	3.198113e-19	1
Max	5021	4.626842e-02	0.593151	1	1	29881
10th Perc	4	0.000000e+00	0.329445	0.379447	1.223761e-06	1
50th Perc	37	3.408369e-07	0.397505	0.767007	2.085476e-05	4
80th Perc	168	2.161768e-05	0.440529	1	1.050072e-04	24
90th Perc	351	8.935352e-05	0.456229	1	2.008113e-04	68
95th Perc	563	2.662830e-04	0.468303	1	2.926613e-04	172
99th Perc	1362	1.758312e-03	0.498557	1	5.938556e-04	1161

Table 2: Summary Statistics of the Lender-Lender projected graph

international banks (more than 180 countries). Additionally, since the development of FRM took place mostly in the United States, DealScan has particularly good coverage for the US loan markets as sustained in [35]. Finally, the closing of a syndicated loan requires interaction between banks' directors and professionals that implies, at least partial, access and sharing of information concerning the process of risk management followed by each bank. The network will be used as a proxy for an observation network where a link from one bank to another means that the first bank observes the risk management solution of another bank. The fact that the syndicated loan market is highly exposed to information asymmetries problems implies that banks who participate with others in a facility do usually acquire information about each other and this will be used as an argument in favor of the use of DealScan as a proxy for observation network.

To get the proxy network, I first extract a bipartite graph where on one side there are lenders and on the other side borrowers. Most of this research will focus on lender-lender relationships, and for this reason, the lender-side projected network will be the primary data input. In this network, a node represents a bank, and a link exists between two banks if they participated together in at least loan. Weights are excluded for the sake of simplicity. Figure 18 shows both the original Bipartite network and the lender-side projected one. Figure 19 shows the log-log plot of the degree of the projected graph. Applying the methodology proposed in [42], the hypothesis of power-law distribution was confirmed for the degree distribution. Additional summary statistics are shown in table 2.

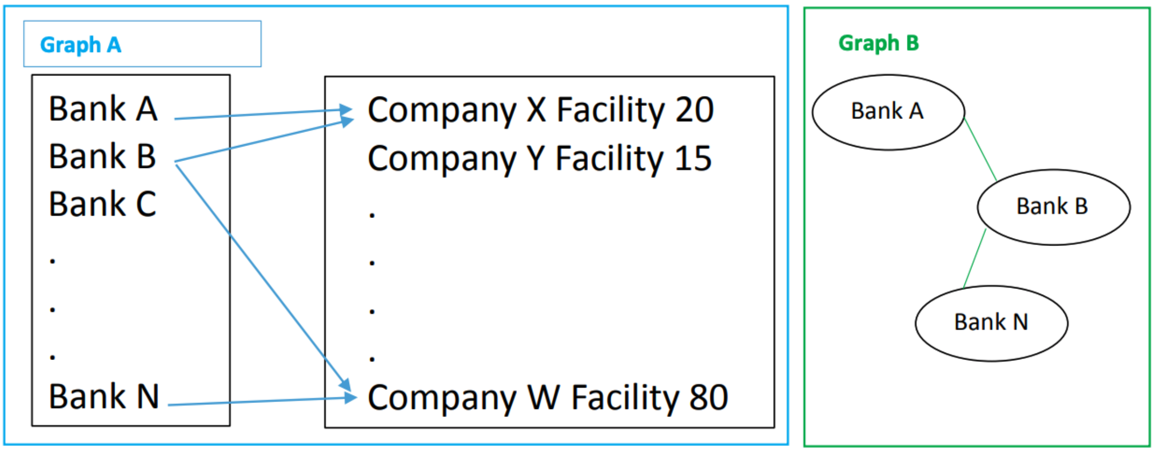


Figure 18: The first graph to the left (Graph A)is the bipartite graph where on one side we have lenders, and on the other side, we have borrowers. For example, it is shown that Bank A and Bank B participated together to give a loan to Company X through Facility N.20. The graph to the Left (Graph B) is the Lender Side projected Network. For example, if on Graph A both Bank A and Bank B participated together in the same facility (Facility 20) they are connected.

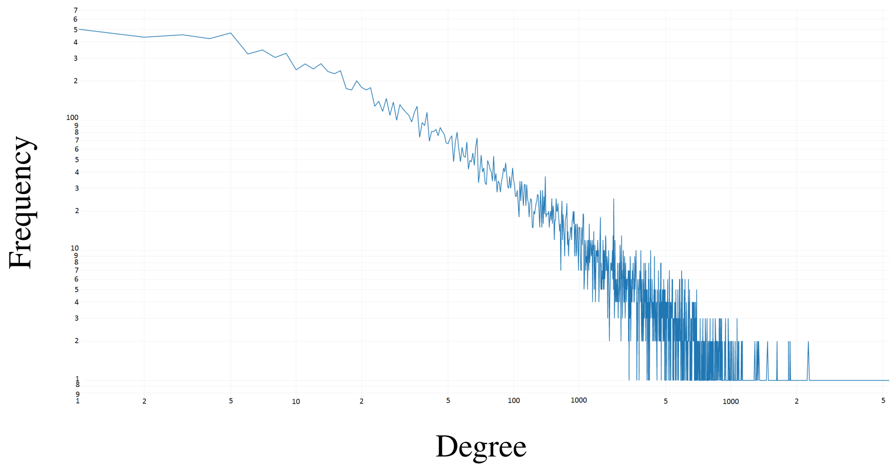


Figure 19: Log-Log plot of degree distribution of the projected lender-lender graph

The second type of data was obtained as a result of text analysis of the annual reports of US banks deposited at the Securities and Exchange Commission. All public US banks are required to report their annual report to the SEC. In their annual reports, banks usually mention the risks they face and the methods used to measure and manage them. I extracted the text parts (when available) where banks discuss their risk method adopted to manage market risk. The majority of banks mention either value-at-risk, sensitivity and scenario analysis, or both. Here are some examples:

> *Market risk arises from price changes in various markets. Market risk from foreign exchange and trading activities is monitored and controlled through established*

limits on positions and aggregate limits based on estimates of potential loss of earnings under assumptions about changes in market conditions.

- State Street Corporation, 1994 Annual Report

Market risk from foreign exchange and trading activities is controlled through established limits on aggregate and net open positions, sensitivity to changes in interest rates, and concentrations. These limits are supplemented by stop-loss thresholds. The Corporation uses a variety of risk management tools and techniques, including value at risk, to measure, monitor and control market risk.

- State Street Corporation 1997 Annual Report

Bank Name	Degree Centrality
Bank of America	5344
Deutsche Bank AG	5336
ABN AMRO Bank NV [RBS]	5317
Societe Generale SA	4983
Royal Bank of Scotland Plc [RBS]	4663

Table 3: Five Largest Degree Centralities in DealScan

Bank Name	Betweenness Centrality
Bank of America	3288722.5
Deutsche Bank AG	2907339.7
ABN AMRO Bank NV [RBS]	2804728
BNP Paribas SA	2403666.3
Citibank	2265226.2

Table 4: Five Largest Betweenness Centralities in DealScan

The data collected also includes the consulting company used by banks for preparing the annual report. The data shows that few companies dominate the auditing market for US banks such as Deloitte, KPMG, Ernst & Yong, Price Coopers, and ARTHUR ANDERSEN & CO.

In this model, I sustain that hubs in financial markets are a driving force behind the adoption of FRM. In order to prove the crucial role of hubs, I present next some analysis of market

Bank Name	Eigenvector Centrality
Bank of America	1
Deutsche Bank AG	0.75398648
Citibank	0.73367721
ABN AMRO Bank NV [RBS]	0.72972546
Bank of Nova Scotia	0.72752112

Table 5: Five Largest Eigenvector Centralities in DealScan

association rules using the apriori algorithm.

5.4 Examining the role of hubs in financial markets: association rules analysis

In this thesis, I sustain that hubs in financial markets play a substantial role in structuring and influencing the market. To give some justification of this claim this section reports on the results of a machine learning analysis using Association rule learning [1] that I applied to the DealScan data. The goal of applying this algorithm is to check whether there are participation rules in financial markets (e.g. if bank A is present then bank B is very likely to be present as well) and understand the role of hubs in the shaping of such rules. There are studies like [34] who already found evidence for trust relationships and memory in the Syndicated Loans Market; however, such studies did not consider the nature and role of hubs in the formation of such relationships. Next, I present the analysis and a discussion of the results.

5.4.1 Association Rule Learner applied to the syndicated loan market

Following the original definition by [1] the problem of association rule mining is defined as follows:

Let I=$\{i_1, i_2, \ldots, i_n\}$ be a set of *n* binary attributes called *items*. In my case these items are the lender banks in the DealScan database (17377 unique banks).

Let D=$\{t_1, t_2, \ldots, t_z\}$ be a set of transactions which would be the data input for the association rule learner. In my case, these transactions represent the facilities (263299 unique facilities). E.g. a facility can be given by bank A, bank B, and bank F in such case the transaction is (Bank A, Bank B, Bank F).

A rule is defined as an implication of the form:

$$X \Rightarrow Y, \quad where \quad X, Y \subseteq I.$$

Which could be interpreted as follows: the presence of X implies the presence of Y. A rule is composed by two different sets of items, X and Y, where X is called *antecedent* or and Y *consequent*. If X has only one element in it the it is called *item*, while if it contains more that one item then it is called *itemset*. Applying this logic to the DealScan data, a rule can be read as follows: The presence of Bank A in a loan implies the presence of Bank B, or the presence of Bank A and Bank B together in a loan implies the presence of Bank C.

In an association rule analysis, three concepts are essential. Let's use X to denote an item (or itemset), $X \Rightarrow Y$ to denote an association rule and T to represent a set of transactions from D. One important concept is the itemset *support*, which indicates the frequency with which an itemset X appears in the dataset. In other words, it is the number of transactions t in D which contain X.

$$\mathrm{supp}(X) = \frac{|\{t \in T; X \subseteq t\}|}{|T|}$$

The same antecedent can produce a number of different consequents. For this reason, another measure of the rule quality is how often that antecedent X appears in transactions that also contain Y. This is the rule *confidence*

$$\mathrm{conf}(X \Rightarrow Y) = \frac{\mathrm{supp}(X \cup Y)}{\mathrm{supp}(X})$$

One more quality measure - the rule *lift* - tells us how much precise a rule is, compared to just the a priori probability of consequent Y. In other words, it is the ratio of the observed support to that expected if X and Y were independent.

$$\mathrm{lift}(X \Rightarrow Y) = \frac{\mathrm{supp}(X \cup Y)}{\mathrm{supp}(X) \times \mathrm{supp}(Y)}$$

To examine the role of hubs on the structure of the syndicated loan markets, I first apply the association rule analysis to the entire dataset and analyse the results. After that I exclude the hubs from the dataset (using a simple criterion for deciding what a hub is) and then repeat the association rule analysis. Next, I present the results for both scenarios.

Column	min	mean	max	st.dev	skewness	kurtosis
Support (0-1)	0.01	0.0208	0.1135	0.0149	2.7425	10.1809

Table 6: Association model with min support 0.01. These statistics concern the itemsets (not the rules) with max length of 3. For example, a minimum support of 0.01 indicates that itemsets of length 1,2, or 3 are chosen if they appear at least in 1% of entire transactions set (facilities). itemsets are of the form (Bank A) or (Bank A, BankB) or (Bank A, Bank B, Bank C).

Column	min	mean	max	st.dev	skewness	kurtosis
Support	0.01	0.013	0.0203	0.0026	1.3342	1.199
Confidence	0.0905	0.2345	0.4953	0.8036	0.8097	0.7904
Lift	1.5455	4.0216	7.6832	1.3418	0.5728	-0.0983

Table 7: Summary statistics for association rules with min support 0.01. This means that these statistics concern the rules whose frequency is at least 1% among the entire set of rules that is obtained from association analysis.

5.4.2 Analysis and results for the entire dataset

After transforming the lender-lender graph into a transaction dataset, I ran the association rule analysis as follows: first, for computational reasons I chose the maximum itemset length to be 3. This means that the algorithm will try to find rules for which the length of the antecedent is at most 3. Second, the algorithm requires the choice of the minimum support level for itemsets and for this I chose the values 0.01, and 0.001. For values greater than 0.01, the result was an empty table (no rules).

With minimum support of 0.01 I get only ten items (I=10), with a maximum support of 0.11 and an average of 0.02 as shown in table 6. Calculating the association rules for this support level I get a total of 120 rules. Table 7 provides some of the descriptive statistics of the measures associated with these rules. As table 7 shows, the average support for the 120 rules is 0.013, meaning that, there are 120 association rules which on average are 1% frequent in the entire dataset. The average confidence was 23%, which means that the rules which I got are on average 23% of the time right. One of the original intentions behind this analysis is to treat rules as an indicator of trust or memory relationships between banks. If the participation of one bank, Bank A, in a facility produces with a high confidence level the consequent involvement of Bank B, then this could be attributed to the fact that a trust relationship exists between the two banks.

Column	min	mean	max	st.dev	skewness	kurtosis
Support (0-1)	0.001	0.0017	0.1135	0.0021	18.7418	579.9909

Table 8: Association model with min support 0.001. These statistics concern the itemsets (not the rules) with max length of 3. For example, a minimum support of 0.001 indicates that itemsets of length 1,2, or 3 are chosen if they appear at least in 0.1% of entire transactions set (facilities). itemsets are of the form (Bank A) or (Bank A, BankB) or (Bank A, Bank B, Bank C).

Column	min	mean	max	st.dev	skewness	kurtosis
Support	0.001	0.0015	0.0203	0.0008	06.8211	79.8988
Confidence	0.05	0.4731	0.9215	0.194	-0.3141	-0.7867
Lift	0.6139	17.4431	170.7222	13.4643	1.856	4.8585

Table 9: Summary statistics for association rules with min support 0.001. This means that these statistics concern the rules whose frequency is at least 0.1% among the entire set of rules that is obtained from association analysis.

The average lift was 4, meaning that on average we are four times more confident that the rules found are precise. This means that it is four times more likely to observe the participation of certain banks as a consequent of one rule compared to seeing the involvement of the same as a consequent in the entire dataset.

By decreasing further the minimum support to 0.001, I get 26088 items with average support of 0.0017 (table 8). Here it could be seen that the maximum support is equal to that of the previous support (0.11). This is because the only thing which was changed is the number of items with lower support.

Moving to the association rules table, with minimum support of 0.001 I get a total number of 84714 rules. The average support for rules here is 0.0015 where the maximum is the same as with the previous level of support (table 9). However, one can notice here that the average confidence is higher (=0.47). The higher confidence may be a further signal that trust or memory is evident when we reduce the minimum support to include lower support items. The reason could be the power law distribution of the lender-lender network where a significant number of banks has a low degree, and few have a very high degree (hubs). This means that, while hubs have many participations with other banks (they appear with very high frequency), the majority of banks appear with much less frequently (low frequency), and when the minimum support was reduced, the results produce much more items and much more rules.

Numerically I can reason about it in the following way: The total number of transactions (facilities) in the dataset is 263299. If we assume a minimum support level of 0.01, this means that I expect a bank to appear in at least 0.01*263299=2632 facilities, which is very unlikely (it would be enough to look at table 2 and we can see that the 80th quantile for participation frequency is 24!, meaning that 80% of banks appeared in a maximum of 24 loans (facilities)). On the other hand, reducing the minimum support to 0.001 means that we expect banks in this sector to appear in 0.001*263299 = 263 loans which is more reasonable, given that the DealScan dataset expands over 28 years and thus on average we expect our banks to participate every year in 263/28 = 9 loans. Another indicator of trust or market structure is the large number of rules that was obtained. If participation were random, we should not expect any association rules! (I pick my products from the supermarket randomly). If we look at table 9, we can notice that the average lift is much higher now than before, with an average value of 17. This may indicate that the precision of the association rules increased with lower support. Now we are sure that if A implies B, then on average this association is 17 times more likely compared to B being implied by some other item in the entire dataset. I decided not to lower more the support level for two reasons: first is the very high computational cost which is involved in reducing the minimum support and second is the fact that, as I said earlier, a frequency of participation in 0.001(minimum support)*263299(N of Transactions)/28(temporal length of the datasets in years) = 9 which is close to the first quartile (=8) in the degree distribution of lenders , so I assumed that it is an acceptable level.

Another way to look at the results is by examining both the antecedents and consequent columns to check for the frequency of times in which the same antecedent or consequent appeared in the rules. The results show that for min support level of 0.001, I get a number of 10348 unique antecedents and 283 unique consequents. This result suggests that there is more variety on the antecedent side than on the consequent side. The hypothesis here is that hubs

Pearson Correlation	Degree Centrality	Betweenness Centrality	Eigenvector Centrality
Degree Centrality	1	0.88	0.82
Betweenness Centrality	0.88	1	0.98
Eigenvector Centrality	0.82	0.98	1

Table 10: Pearson correlation between centrality measures

Spearman Correlation	Degree Centrality	Betweenness Centrality	Eigenvector Centrality
Degree Centrality	1	0.99	0.99
Betweenness Centrality	0.99	1	0.99
Eigenvector Centrality	0.99	0.99	1

Table 11: Spearman correlation between centrality measures

play a crucial role in shaping the structure of the market. Due to the existence of hubs, who are frequently present in many facilities, it is very likely that their participation gives rise to many associations between market participants. In order to check for the effect of hubs on the structure and associations in the syndicated market, I run the same Association Rule Analysis on the same dataset but excluding the hubs. Next, I discuss the results for this analysis.

5.4.3 Analysis and results after excluding the hubs

As I mentioned at the beginning of this analysis, I intend to capture or measure the role of hubs on market structure. To proceed with this analysis, I first analyze the correlation between the frequency of participation of a bank and acquiring a higher degree. To achieve these results, I calculate two measures:

1- The Pearson and Spearman Correlation between the Degree Centrality, Betweenness Centrality and the Eigenvector Centrality (tables 10 and 11).

2- The Pearson and Spearman correlation between the frequency of each lender in the dataset and its degree (tables 12 and 13). Afterwards, we move to discuss the output of the association rules analysis conducted by excluding the hubs from the dataset.

As shown in tables 10 and 11, there seems to be a high positive correlation between

Pearson Correlation	Degree Centrality	Frequency
Degree	1	0.780279
Frequency	0.780279	1

Table 12: Pearson Correlation between lender Degree and Lender Frequency

Spearman Correlation	Degree	Frequency
Degree	1	0.8047502
Frequency	0.8047502	1

Table 13: Spearman Correlation between lender Degree and Lender Frequency

all three measures of centrality. This could be interpreted as an indicator that if a lender is connected directly with a high number of other lenders (high degree), this lender would also be very important as a mediator between other lenders (high betweenness). These two measures are also positively correlated with a high Eigenvector centrality, which means that if a bank is connected to a large number of other lenders and it sits in many of the shortest paths of other lenders, this implies that the bank will have as neighbours those of high importance as well. This could be interpreted by referring to the almost stylized fact according to which the structure of financial markets is core-periphery. In a core periphery-structure, the biggest banks form a well-connected core, and the smaller banks form a loosely connected periphery that is mostly connected with the core.

Examining the correlation between degree and bank frequency (tables 12 and 13), it could be seen that the correlation between participation and degree is positive and high. With a value of 0.8, we can conclude that participating many times in this market is highly correlated with acquiring neighbours and thus becoming an important player. Putting together the results in table 10 and 11, the main idea emerges that the structure of the syndicated loan market is heavily dependent on the hubs who are present in the majority of flows.

Now going back to the association rules, I ran the association rule analysis on the same dataset, with minimum support of 0.001 and min confidence 0.05, but excluding the nodes which have a frequency (defined as the number of time a bank appeared in the dataset) greater than 10000. Next, I report the main results.

First, the total number of items moves down from 263299 to 246758 (a difference of 16541, 6%). Table 14 reports the statistics for the support count. It could be seen that the mean support is 0.0021, close enough to 0.0017 in the case with hubs (Table 8). However, if we look at the maximum support we can see a substantial reduction with respect to the case with hubs (0.039 vs. 0.1135). This is intuitive given that I excluded the hubs.

Second, including the association rules in the analysis, we get a lower number of rules equal to 11498 compared to 84714 rules when hubs are included (a reduction of 86%). This is a strong indicator that hubs in this market are, as initially assumed, a crucial element in forming the structure and relationships in this market. The table below reports some statistics for the rules. Here we can notice that the average support value is close to the case with hubs with the only difference being a lower maximum support equal to 0.009 compared to 0.02 with the hubs.

As for the confidence, the only difference is the average value which is lower (0.37 vs. 0.47 with hubs). As I showed earlier, the higher the confidence, the more we are sure that the rule is true. By removing the hubs from the system, we are less sure that the rules are true since in the syndicated market there is some sort of hierarchy where there is the arranger (usually the hub) who is responsible for structuring the loan and the participating banks. By removing the hubs we are less confident that the entire hierarchy of the market is represented. The last measure is the lift, which results higher on average if we exclude the hubs (29 vs. 17). This reduction could be due to the fact that by removing the hubs we get a reduction of 86% in the population of rules which, given the definition of lift, reduces the size of the population to which we compare the precision of the rule.

Finally, we can move to understand what happened to the number of unique antecedents and consequents when I removed the hubs from the dataset.
Results show that the number of unique antecedents is 1813 compared to 10348 in the case

Column	min	mean	max	st.dev	skewness	kurtosis
Support (0-1)	0.001	0.0021	0.0396	0.0.0027	7.139	63.5708

Table 14: Association model without hubs and with min support 0.001. These statistics concern the itemsets (not the rules) with max length of 3. For example, a minimum support of 0.001 indicates that itemsets of length 1,2, or 3 are chosen if they appear at least in 0.1% of entire transactions set (facilities). itemsets are of the form (Bank A) or (Bank A, BankB) or (Bank A, Bank B, Bank C).

Column	min	mean	max	st.dev	skewness	kurtosis
Support	0.001	0.0.0016	0.009	0.0007	3.1771	15.0758
Confidence	0.05	0.3794	0.9153	0.2235	0.2618	-1.0036
Lift	1.3294	29.1241	160.0439	21.0681	0.847	0.4347

Table 15: Summary statistics for association rules without hubs and with min support 0.001. This means that these statistics concern the rules whose frequency is at least 0.1% among the entire set of rules that is obtained from association analysis.

with the hubs (a decrease of 82%), while the number of unique consequents is 277 compared to 283 when hubs were included. Again here we can observe the importance of hubs in structuring and determining the rules of the market. If hubs do not exist, there are many relationships and associations (such as trust) which would not exist.

With this finding, I would conclude that trust in the syndicated loans market is heavily dependent on the hubs which, through their influence and power, can shape the bulk of market relationships. This conclusion is in line with interesting findings in network science. For example, [182] studied the evolution of cooperation in the framework of evolutionary game theory and found that when individuals interact through a network generated via preferential attachment, cooperation becomes the dominant trait; thus hubs play a very important in promoting cooperation. The same might apply for the analysis presented in this section: since hubs have experience, resources, and risk management capabilities, the participant banks would have more trust in the deal, and thus hubs encourage participation.

In the next section, I will turn to present and discuss the simulations procedure for the historical agent-based model for FRM.

5.5 The model

5.5.1 The network

Given the large size of the projected graph obtained from DealScan, I took a smaller network which is the lender-lender network for the year 1993. By 1993, financial markets almost completed the process of convergence to the actual financial risk paradigm which was dominated by Value at Risk and Scenario Analysis practices. Banks and regulators both accepted FRM as the leading paradigm for measuring and managing market risk. The sampled network has 905 nodes (banks) and 26412 unweighted and undirected edges. For the purpose of this model, the network is then converted to a directed network using similar logic as proposed in section 4.2. If a peripheral node is connected to a hub node, then the link is converted to a directed link from the peripheral node to the hub. If a peripheral node is linked to another peripheral node, then the link is excluded, and if a hub is connected to a hub, then a bi-directional edge is created between the two hubs. An edge from one bank to another indicates that the first bank observes the risk measurement solution of the second bank. I assume that banks are able to observe and understand the solutions of the other banks. Two simulation scenarios are considered, the first has as hubs the biggest 15 nodes (in terms of degree) while the second has as hubs the biggest 30. This choice is based on some facts from the history of FRM. For example, according to Till Guldimann, one of the main innovators who popularized VAR measures at JP Morgan, by the early 90s only the top 10-15 banks were developing risk management technologies [4]. The network with 30 hubs was constructed based on the number of participants in the G30 summit in 1993. These numbers are simply assumptions of what was the true number of banks working on risk management technologies. Figure 20 shows a (small) illustrative graph that depicts the network structure used in this model. Hub nodes are colored in blue while periphery nodes are in yellow. Is this network setting a realistic one? As I illustrated in section 3.2.5, in many situations it could result economically optimal for a system to organize itself as a core-periphery network where core nodes acquire information individually, while peripheral nodes acquire no information individually but form links and get all their

[4] http://finance-and-banking.blogspot.com/2007/06/blog-post.html

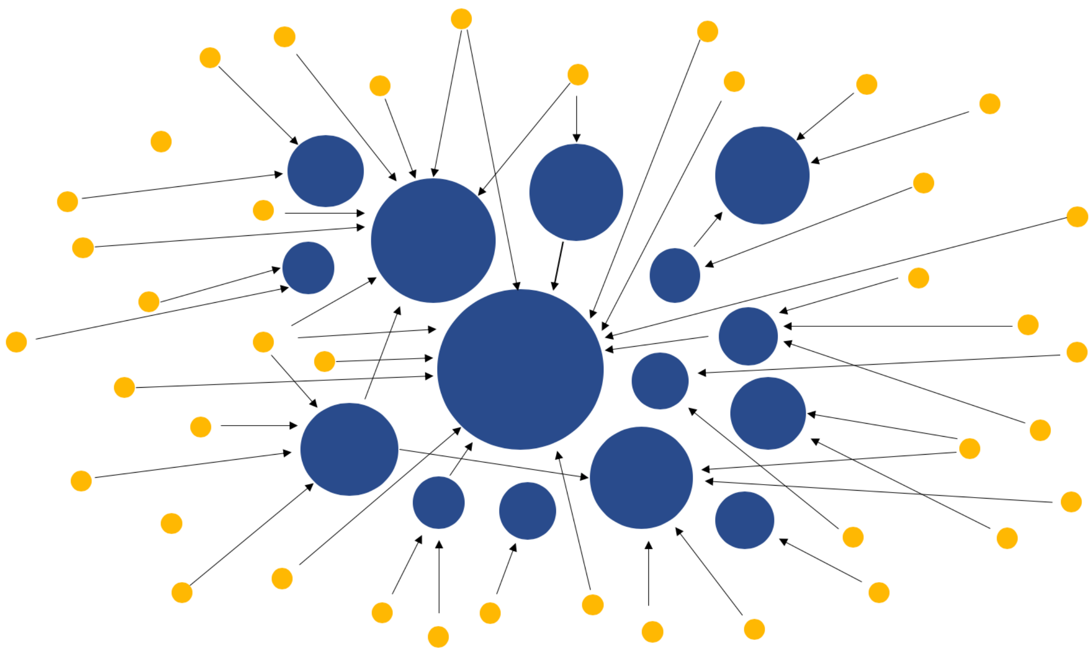


Figure 20: Interaction Structure in the financial network. I note that this is an illustrative graph of the original one. Blue nodes are the core and yellow are the periphery. The size of the core nodes represent their degree but this is not based on real data. Additionally, in the real data the core is fully connected but in the figure I excluded most links in the core for visualization reasons.

information from the core nodes [84]. The economic motives of such structure might derive from the fact that returns from the information acquiring by peripheral nodes is higher than the costs of individually obtaining information (exploration). Several factors might complicate the decision of a bank to develop a risk management technology. For example, there are usually complementary assets that are simultaneously developed with other technologies. In the case of risk management, these complementary elements could be software companies developing commercial implementations of risk technologies, or university programs who start teaching these techniques to finance students. Another factor is the need for regulatory approval when a bank develops a new risk management solution.

5.5.2 The problem and the solution space

The problem to be solved in this model consists of finding the risk measure that can predict future losses arising from market factors (hence market risk). The problem will be assumed to have two conflicting dimensions: *breadth* and *simplicity* as discussed in section 5.2. Increasing the breadth of a risk measure is likely to produce more accurate predictions but at the same time reduces its commercial applicability. On the other hand, increasing the simplicity of the risk measure makes it easy to implement in commercial space but at the risk of underestimating risk. The fitness value of a solution in the landscape would be the expected precision of its prediction ([87] adopted similar logic for operationalizing fitness). For the sake of simplicity, I assume that the landscape is given exogenously and agents arrive at the same fitness evaluation for similar solutions. The landscape used in the analysis is the same one in section 4.1, where one dimension is breadth and the other is simplicity.

5.5.3 Behavioral rules and research design

The search process is modelled as follows: first, hubs are assumed to be the only nodes which explore the fitness landscape. This assumption is based on the historical evidence discussed in section 5.2 that only 10 - 15 banks were involved in developing solutions for financial risk management. At each round, a hub observes a randomly chosen neighbouring hub (the data shows that all hubs are connected to each other), and if the fitness of the solution of the observed hub is higher, then the observing hub adopts the solution of the observed hub. If the fitness of the solution of the observed hub is lower or equal to the fitness of the current solution of the observing hub, then the hub adopts an independent search heuristic that splits the fitness landscape into four equal modules and choose four solutions from each. The first module include all solutions (i,j) where both i and j are the range 1-500, the second module has range 1-500 for i and 500-1000 for j, the third has range 500-1000 for i and 1-500 for j, and the last module has range 500-1000 for both i and j. Each module contains a total of 250,000 candidate solutions (see figure 21). The practice of decomposing a fitness landscape into smaller modules and searching modules in parallel is thought to be one of the most efficient heuristics that searching agents can adopt when facing complex economic problems (see [17; 126]).

After choosing four new solutions, an acceptance parameter is assigned to each solution as follows:

$$P(i^1_{new}, j^1_{new}) = 0.1(\frac{d[(i_{actual}, j_{actual}), (i^1_{new}, j^1_{new})]}{max(d)}) + 0.9(1 - exp(-\frac{fit(i^1_{new}, j^1_{new}) - fit(i_{actual}, j_{actual})}{T}) \quad (1)$$

$$P(i^2_{new}, j^2_{new}) = 0.1(\frac{d[(i_{actual}, j_{actual}), (i^2_{new}, j^2_{new})]}{max(d)}) + 0.9(1 - exp(-\frac{fit(i^2_{new}, j^2_{new}) - fit(i_{actual}, j_{actual})}{T}) \quad (2)$$

$$P(i^3_{new}, j^3_{new}) = 0.1(\frac{d[(i_{actual}, j_{actual}), (i^3_{new}, j^3_{new})]}{max(d)}) + 0.9(1 - exp(-\frac{fit(i^3_{new}, j^3_{new}) - fit(i_{actual}, j_{actual})}{T}) \quad (3)$$

$$P(i^4_{new}, j^4_{new}) = 0.1(\frac{d[(i_{actual}, j_{actual}), (i^4_{new}, j^4_{new})]}{max(d)} + 0.9(1 - exp(-\frac{fit(i^4_{new}, j^4_{new}) - fit(i_{actual}, j_{actual})}{T}) \quad (4)$$

Where (i_{actual}, j_{actual}) is the actual position in the fitness landscape, (i^k_{new}, j^m_{new}) denotes one of the four solutions chosen from the four modules, d() calculates the Euclidean distance between two solutions, max(d) is the maximum Euclidean distance in the landscape (d[(1,1),(1000,1000)], fit() is the fitness function, T is noise parameter equal to 200. A weight of 0.1 is given to the ratio between the distance between the actual and new positions (numerator) and the maximum possible distance (denominator). The reason I added this 0.1 weight is to account for what is called *dynamic capabilites* , which is a constraint on the ability of an

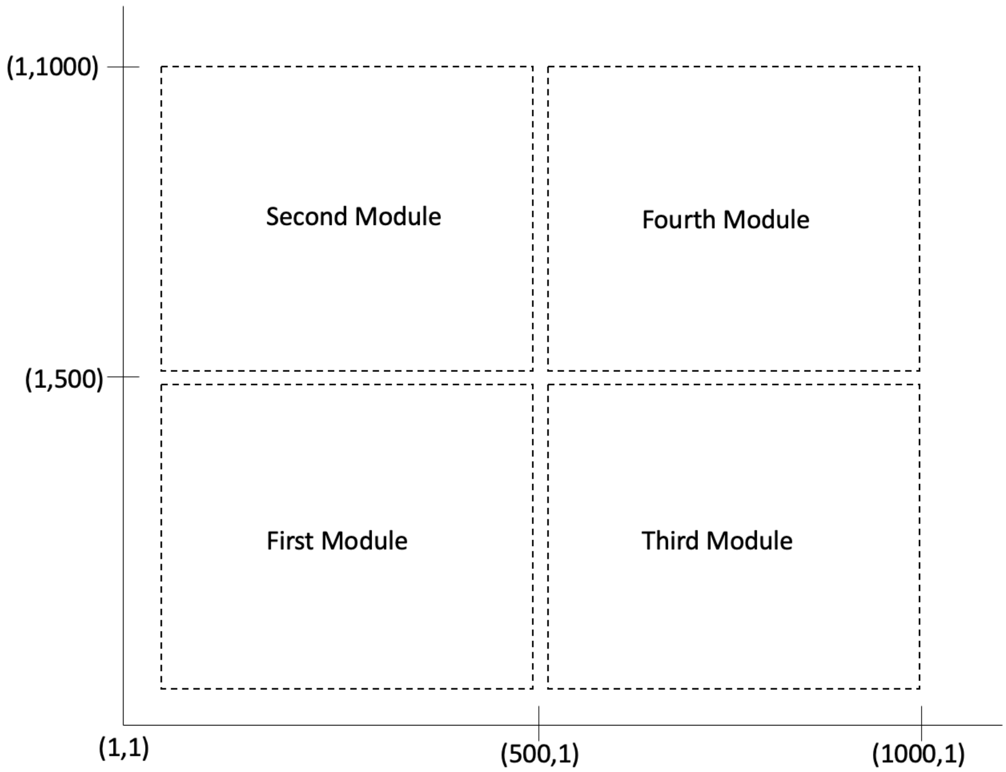


Figure 21: Modular Decomposition of the Fitness Landscape

organization to move in the landscape [204]. The other term is given a 0.9 weight and it is dependent on the difference in fitness of each solution, divided by a T=200 parameter which is included to add randomness to the probability of acceptance.

After evaluating these four parameters, I assume that a bank ranks them in decreasing order and chooses the solution with the highest acceptance parameter if its value is positive. Note that such an acceptance parameter might be negative if the fitness of the newly chosen solution is lower than the fitness of the actual solution. If all acceptance parameters are negative, then the bank does not choose any of the four solutions. When this happens, then the bank chooses a random solution from the fitness landscape and adopts it if it has higher fitness than the actual solution. If this second option results also in a solution with lower fitness than the actual one, then the bank remains in its position.

As for the search behaviour of the non-hub nodes, I adopted a simple heuristic: at each time step, each peripheral bank chooses one adjacent hub node at random and adopts its solution. Periphery nodes do not search on their own, they only observe the hubs.

At the end of each round, there will be the regulator who observes all the choices made by all banks, calculates the convergence in space (the fitness landscape) and with certain probability β accepts the convergence. In the next section, I illustrate how convergence is calculated. Throughout the thesis, I will call β the probability of convergence acceptance. Four values of β were used: 0.1,0.4,0.7, and 1. If the regulator accepts the convergence at round x, then banks will retain their positions in the fitness landscape at that round. If the regulator does not accept the convergence, then all banks will return to the positions there were in at round x-1, and start from there at round x+1.

The realizations of the model were repeated for 50 different sets of initial conditions, where each set contains a total of 905 (=number of nodes) positions in the fitness landscape. I note that none of the initial locations contains the global optima. For each set, I run one realization for 1000 rounds, and at the end, results are averaged for each round over the 50 realizations.

FRM-Fitness-Averages-eps-converted-to.pdf

Figure 22: Average fitness value for the networks with 15 hubs and regulator (upper left), 15 hubs without regulator (upper right), 30 hubs with regulator (lower left), and 30 hubs without regulator (lower right)

5.6 Statistical quantities

For the purpose of this model, I am concerned mainly in three quantities: average performance, network convergence in the landscape, and the average number of times the global optima was found. Denote with V(G) the set of all nodes in a network G, with $(p^1_{i,x}, p^2_{i,x})$ the position of a node i at round x, and with R the number of realizations performed, then the system level average performance for network G at round x is calculated as:

$$Avg(G,x) = \frac{\sum_1^R \sum_{\forall i \in V(G)} fit(p^1_{i,x}, p^2_{i,x})}{R * |V(G)|}$$

In this model, I am interested in observing the convergence of the banks in the solution space. At the end of each round, I calculate the total sum of all pairwise Euclidean distances between the positions of each pair of nodes. Let's call (p^i_n, p^j_n) the position of node n in the fitness landscape and (p^i_m, p^j_m) the position of another node m, then the total Euclidean distance of a network G (call it d(G)) is calculated as:

$$d(G) = \sum_{\forall m,n \in G, n \neq m} \sqrt{(p^i_m - p^i_n)^2 + (p^j_n - p^j_n)^2}$$

The third statistical quantities I am interested in is the number of times the banks found the global optima at any given round:

$$Glob(G,x) = |\{i | i \in G \, and \, fit(p^1_{i,x}, p^2_{i,x}) = 10000\}|$$

5.7 Discussion

In this section, I will discuss the results obtained so far for the simulations of the model proposed in the previous section. In figure 22, I report the average performance of the networks with 15 and 30 searching hubs. In the left column, results are reported for the simulations with regulator for the four convergence acceptance probabilities [0.1,0.4,0.7,1], while in the right column the same simulations are reported to show how the system would have performed in the absence of the regulator for each simulation. By focusing on the plots for the network with 15 hubs (upper row), it could be immediately noticed that the system achieves always high average fitness if the regulator is absent (right plot). This result can be attributed to two main factors: first, the search heuristic adopted by the hubs (landscape decomposition) is very efficient, therefore making it very likely that the system will achieve high fitness if left on its own (without the intervention of a regulator). Parallel search can be costly, and one might ask whether this is a realistic assumption. I would sustain (speculate) that this might be the case. In appendix B, I list the total amount in dollars of the investments made by the largest 50 banks in 1993 as reported by DealScan. The table shows that banks invest money in the order of tens of billions every year and this can be treated as an indicator that the financial sector did not lack the resources to explore the landscape of risk management solutions. Obtaining data or estimates of the costs involved in the development of risk management is one possible extension of this model. Second, given that periphery nodes do observe only hub nodes, this acts as an accelerating factor in the diffusion of solutions found by the hubs. This scenario is also dependent on the nature of the real fitness landscape of risk management solutions. If we assume that the real landscape has few peaks and is easy to explore via parallel and modular search, then it is very likely that the system with few hubs will be able to find the most optimal solutions and disseminates them through imitation. The history of FRM shows that financial institutions acted as if there were only a few optimal solutions to the problem of financial risk measurement and the evidence for that is the few options that were discussed whenever risk management was the object of policy discussion. As some authors show (see [103; 104]), most of the discussion revolved around tools like Value at Risk, scenario analysis, and stress testing. Authors in reference [103] claims that similar tools with the same underlying logic were used throughout the 20th century both by regulators and banks. This, however, could merely indicate that the landscape is rugged and further solutions are difficult to evaluate or test. This remains an extension to this study that deserves investigating.

Moving to the case where a regulator is present (upper left plot in figure 22), an evident trend emerges: the higher the probability that a regulator accepts network convergence, the higher is the average fitness achieved. This result shows that the presence of a regulator can act as a barrier towards an efficient exploration of the fitness landscape by the network of banks. This, however, presents the regulator with a trade-off between efficiency and risk, which is often the main tradeoff in financial regulation [163]. The efficiency can be represented by the average fitness of the system, while the risk lies in the fact that the system might converge to a uniform solution that reduces the diversity of risk models. If the financial system as a whole adopts the same risk measure, then this might create a system-level asymmetric risk profile where everyone will be affected by the same market factors. Back in 1997, Nassim Taleb [201] warned that the generalised use of measures like Value at Risk would lead to systematic risk especially because VAR measures tend to underestimate rare events. Ben Bernanke, former chairman of the Federal Reserve, criticized homogeneity in the banking system and sustained that '*a single firm may have an acceptable exposure to a particular type of risk that would be unacceptable if replicated across many firms*'.

When the number of searching hubs is 30 (figure 22 lower row), the results were very similar to the case of 15 banks. The only small difference could be noticed is that the difference in system performance with probabilities of acceptance 0.4,0.7 and 1 are less wide compared to the case with 15 nodes. This result could be explained by the fact that when 30 banks are exploring, there are always higher chances that the system will find the optimal solutions and therefore achieve higher fitness. By repeating the simulations on a wider range of core sizes (15, 30, 60, and 120), I found that there is a monotonic relationship between the average fitness and the core size: the larger is the size of the core, the better is the average performance. This result goes in line with what was found in section 4.5.3 where it was shown that less heterogeneous structures tend to achieve higher fitness. In a real case scenario, this would depend on cost-benefit analysis. Is it optimal for the entire system to invest in landscape exploration? This remains an interesting topic for further research

FRM-Distance-Averages-eps-converted-to.pdf

Figure 23: Average euclidean distance for the networks with 15 hubs and regulator (upper left), 15 hubs without regulator (upper right), 30 hubs with regulator (lower left), and 30 hubs without regulator (lower right)

To understand more in detail the dynamics of convergence, figure 23 plots the average Euclidean distance between all nodes in the network at each time step. With the presence of regulators, the system exhibits more convergence in the fitness landscape when the convergence acceptance probability increases (left column). If the regulator was absent, it could be shown that the system always exhibits a fast convergence in space (right column in figure 23). This behaviour indicates the trade-off again that regulators face between efficiency and risk. When the system exhibits convergence, this might act to increase efficiency as financial markets are known to adopt the solution that appears efficient. The efficiency in adopting a uniform risk model derives mainly from the fact that markets can speak the same risk language which facilitates trading and risk sharing between market agents. Regulators would be interested in having markets converge to a uniform solution because this way they can compare the risk profiles of different financial institutions and evaluate the system-level stability. However, if the risk model adopted by the system results to be seriously flawed or inadequate (especially when market conditions change), this might act to generate systemic risk that might jeopardize financial stability. Risk model homogeneity can be treated as a special form of systemic risk. Regulators have actually realized this category of risk in the 1990s by allowing banks to use their internal models to evaluate risk [193]. These internal models, however, are often constrained by some parameter assumptions that are set by regulators. The history of regulatory intervention (until these days) have always shown several episodes of negotiations between regulators and banks on how much freedom to give banks in formulating their internal risk models for capital requirements calculations. In a future extension of this model I will modify the fitness function to allow for dynamic fitness evaluation based on the number of adopters of a certain solution. The fitness landscape used in this study is static in that fitness values do not change with time. However, the fitness of specific solutions to a problem can be a *decreasing* function of the number of adopters of such solution. How is this possible?

FRM-Global-Found-Averages-eps-converted-to.pdf

Figure 24: Average number of times the global optima was found for the networks with 15 hubs and regulator (upper left), 15 hubs without regulator (upper right), 30 hubs with the regulator (lower left), and 30 hubs without regulator (lower right)

A naive way to think of it is by imagining that the adoption of a specific solution might encourage certain behaviours that change the conditions under which such a solution might work. Suppose that a bank adopts Value at Risk to measure market risk and assume that this measure gives very low probability to large and rare shocks. Now imagine that this measure unintentionally encourages risk-taking on the side of banks because of its reliance on normal market assumptions. This, of course, would not be a problem if we are limited to one or a few users (large shocks are still rare). However, if the entire market adopts this measure and everyone starts accumulating risks, then the likelihood of big shocks is no more low as predicted by Value at Risk and consequently, such measure would start giving low-quality predictions (thus its fitness decreases).

Finally, figure 24 shows how many times on average did the network (how many nodes) find the global optima. Surprisingly, in the absence of a regulator (right column), the entire system managed to find the global optima after circa 200 rounds. I note that none of the initial conditions includes global optimum. This high efficiency in finding the global optima without a regulator can be attributed to the effectiveness of search heuristic adopted by the hubs and the fact that peripheral nodes do observe only the solutions of the hubs. By decomposing the landscape into smaller areas and the exploration of each in parallel, the probability of locating the global optimum becomes much higher.

When the regulator is present (left column), different behaviour could be observed. For the network with 15 hubs (upper left), the system managed to find the global optima only when the probability of convergence acceptance is 1. If regulators were to accept network convergence with probability smaller than 1, then it is improbable that the system will be able to find the global optimum. The fluctuations shown in the figure are mainly due to the fact that peripheral nodes are randomly choosing the solutions of the hubs which they observe without making any fitness evaluation. The lack of fitness evaluation by peripheral nodes means that even if a peripheral node happens to be at the global optima in one round, this doesn't exclude that at the next round the node observes another hub and adopts its solution which might not be the global optima. This again can act as an intervention in the efficiency of markets by a public authority for the sake of achieving stability. This, however, raises an epistemological problem inline with what I discussed in section 2.3.1: To say that the system finds the global optima would be true only if the fitness landscape is objective and somehow known to the agents. Objective evaluations do not necessarily hold in real-life. A fitness landscape might be a subjective construct in the minds of agents and what could be a good solution for one agent might be a bad solution for another. It might also be that a fitness landscape is objective but agents are unable to know it given bounded rationality. In the case of FRM I assumed that fitness represents the precision of expected loss of a risk measure. However, this precision will only be known in the future. This remains an interesting question in the fitness landscape literature. Interestingly, when we consider the network with 30 hubs (lower left plot), we can notice that the system managed to find the global optima when the probability of convergence acceptance is less than 1. This is again due to the fact that there is twice the number of hubs exploring the fitness landscape and therefore the likelihood of finding the global optima is higher. After repearing the same analysis with core sizes of 15, 30, 60, and 120, it was found that there is a monotonic relationship between the number of nodes who found the global optima and the size of the core: the bigger is the core, the more nodes find the global optima.

6 Policy implications: how do network science, technological innovation, and policy-making interact?

Policy interventions in the economic sphere are so pervasive, and with the increasing reliance on network science to study complex economic phenomena like technological innovation, there is a corresponding need to address how policy-making might benefit from such studies and why. Authors of papers and studies usually include a short section at the end of their research recommending few policy implications; however, it is rare that authors understand the connection between science and policy. In this thesis, I will not list policy recommendations given the scope and explanatory nature of the models analysed. However, I will offer a brief and basic introduction to the process of policy making and how it might relate to network analysis like the ones in this thesis.

Generally speaking, there are three different angles from which policy-making can be scientifically approached. The first angle focuses on the descriptive or explanatory part of policy-making and can be summarized with the question *what does economic policy-making do*? A dominant theory in this regard is *public choice theory* which states that in politics there are separate interests and the pursuit of those interests mainly drives political behaviour. Crucially, many government interventions do not involve a search of self-interest, but instead, they are instrumental in the sense that they are intended to prevent market failures. Traditionally, the public choice theory is focused on equilibrium-oriented assumptions, comparative-static methodology, and assumes perfect or almost perfect information. However, adopting an evolutionary perspective, author of reference [219] sustained that politicians, as well as voters, are subject to bounded rationality and thus have limited knowledge. Having limited knowledge is likely to induce political agents to be highly selective in their learning. In this case, some information is expected to receive attention and influence agents perceptions and beliefs while other information will be ignored. Neglected information might be discovered sometime in the future and change agents beliefs and normative judgments, thus dynamically shifting attention. A shifting attention process might be subject to path dependency in the evolving norms, values, and ends [51].

Information selection and learning in policy-making happen often through communication channels. These channels might be a community or group where information dissemination occurs in a decentralised and face-to-face way, or through centralised means like the media. In many cases, communications processes are subject to *agenda-setting* effect [185], meaning that certain topics will be selected to be the main points of political discussion. The selective nature of policy-making is crucial if we want to understand the conditions under which a public opinion emerges. As the author of [218] explained, for some policy idea to attract sufficient public attention in a self-amplifying process, there is a need for a *critical mass* of people discussing and communicating that policy item within their social networks. This is the reason why interest groups are always interested in engaging in the communication and agenda-setting process. Building on these views, the process of policy making can be thought of as a *collective learning process*, where networks of politicians, academics and interest groups together discuss and set political agenda. An interesting example in this regards was offered in [107] who studied the political process that led to the adoption of patent laws in Italy in the 1960s. Hutter illustrated how social networks (or what the author called *conversation cycles)* that were established between lawyers, policymakers and interest groups had a significant effect on the public opinion formation and the final shape of the legislative measures taken.

In this thesis, I first showed that under certain conditions, short average path length networks outperform long average path length ones in collective search problems. Let's treat this as a piece of information; thus it might attract attention in conversation cycles or get neglected. Assume that it draws attention, then questions arise such as: how does this translate into policy problem? should policymakers encourage innovators to communicate or observe each other quickly? Is the evidence convincing? Is communication between innovators dependent on geography or geography doesn't matter? In the literature, results have disagreed on the impact of different types of networks on firm innovativeness (see [186]). Now if we consider my case study about Risk Management, the policy aspect was more evident. The story of FRM is a story of convergence to industry standard with the supervision and intervention of policymaking and the banking network exploring the space of possible solutions. Before the 1990s, developments in risk management were decentralised, and there was no unique reference for the financial system to use as a benchmark. Beginning in 1990, conversation cycles emerged to accelerate the process of convergence. Thes conversation cycles were materialised in the working group of senior managers and academics that was formed by JP Morgan Chairman Dennis Weatherstone based on a solicitation of Paul Volker in 1993, (see appendix A for a list of the participants). Knowing that few core banks are leading the innovation could help policymakers understand whether the system would achieve optimal results if left on its own. Or will hubs drive the system to a sub-optimal solution? Are hubs always the best for exploring complex technological landscapes? Aren't small banks more flexible and can study in areas that big banks would rarely approach?

The second approach to economic policymaking deals with the theoretical basis of *instrumental policy making* and revolves around the question *what economic policymaking could (try to) do*? The instrumental view deals with the practical and applied ways of dealing with policy problems. In this approach, values, ends, and motives of political agents are not part of the analysis. Crucially, instrumental policy analysis is dependent on the predictive power of the theories applied. This opens the door for unintended consequences due to uncertainty. For example, agents who are unfavourably affected by a policy measure might have adaptive incentives to find a way to circumvent such a policy. A good policy making is thus one that tries to understand the range of its unintended consequences before implementing it. Assume that the idea of short average path length innovation networks discussed in this thesis receives attention in policy conversation cycles, what could policymakers do about it? Can policy making change the network structure of agents in the marketplace? In general, the most common policy instruments for innovation are patent laws and R&D. With patent laws, agents have less incentive to observe and imitate each other because each innovation is legally protected. This might lead to the emergence of another type of innovation networks like alliance networks. Another unintended consequence is if alliance networks emerge and start acting like cartels, thus jeopardising market competition. It might also happen that networks operate to reinforce the institutional isomorphism mechanism which I discussed in section 3.2.8. Other policy instruments used to encourage inter-firm technological collaboration include science parks, industry clusters, support for business incubators, laws and regulations for contracting, intellectual property rights, and other.

The third angle is the *normative* angle which investigates the goals and legitimisation of economic policy, and can be characterised by the question *what ought economic policymaking do*? This is the level at which ends and values are discussed from a normative point of view. In this approach, the analysis focuses on distinguishing the goals that are legitimate to pursue and how to achieve them. Another critical issue is to differentiate between targets that are compatible and conflicting goals. Among the best examples that can be used to illustrate this question is technological innovation [219]. Innovation is often thought of as a good thing that will bring benefit to society. It is known however that technological innovation might also cause negative externalities, so what kind of technological innovation policy makers need to support? Now if we include networks in this discussion, then similar questions might arise: are networks good for innovation? Do networks always provide the best channel to solve technological problems? should policymakers focus on influential nodes or the peripheral ones? Can networks guarantee the necessary diversity need for the system to evolve? If networks are functional, how much regulatory intervention is needed to ensure a balance between market efficiency and system stability? I would encourage scholars to consider these issues with more detail.

7 Conclusion and future extensions

Besides being a primary and outstanding example for a complex economic problem, technological innovation is increasingly stylized as a 'collective' phenomenon of interactions between a network of innovating agents. To better understand the characteristics of the network structure needed to promote system-level innovation, I conducted the following analysis in this thesis: first, I analysed the role of network average path length, edge direction, and degree heterogeneity on the average performance of a network of innovating firms. All networks were located on a fitness landscape that is assumed to represent the space of solutions to a technological problem. Results for this first analysis showed that the system achieves higher performance in the short-medium term when networks are fast at circulating information about technological innovations. In the long run, however, some long average path length networks showed marginally higher performance in terms of average fitness and number of firms who found the global peak. It was also found that the average performance of the system results optimal when the probability of agents observing each other is neither too small nor too large, thus introducing a trade-off between autonomous and collective search. Finally, by introducing heterogeneity of the degree and directional observation, it was found that the network system performs much worse than in the case of undirected and degree-homogeneous networks. By increasing degree-heterogeneity, networks assume a more centralized structure primarily focused on the performance of the main hubs, which acts to constrain the search opportunities. The second analysis done in this thesis concerns a historical and empirical case study of Financial Risk Management which I modelled as a collective innovation phenomenon, involving a network of banks and a central regulator. The goal of this second analysis is to provide an empirical example of a situation where the network plays an important role in the development of technological innovation. The preliminary results obtained for the case study showed that the intervention of a regulator introduces a trade-off between efficiency (average fitness) and diversity of solutions, which is an important policy issue in financial markets. High efficiency comes at the cost of agent convergence in the space of solutions (fitness landscape), while high diversity comes with lower system performance.

As with any scientific study, the results in this thesis are strongly dependent on the assumptions of the models and several limitations are present. There are therefore critical potential extensions in both empirical testing and modelling. First, let's take the computational model presented in chapter 4. On the theoretical level, I purposely constructed this model with the minimum complexity in order to focus on the role of different network structures in supporting technological innovation. There is, therefore, space for extending the model. It would be useful to examine a more extensive array of network structures and their performance under different conditions. In reference [14], it was found that different network structures perform differently according to the learning strategy that agents adopt. This might raise an interesting question: is it the strategy that gives the network structure importance or the other way around? In other words, is it the nature of dynamics that we run on a network that gives significance to the network structure or it is instead the structure of the network that produces the dynamics? An interesting network structure that could be used for technological innovation is the modular (community) structure since innovating firms are known to form clusters [172; 25]. This thesis has been concerned with the system-level (or network-level) performance of firm networks, but it would also be of interest to examine the performance at the level of a single firm [90]. Another realistic extension is to assume that the network is not given a priori, but emerges endogenously from learning cost-benefit analysis as illustrated in section 3.2.4. Finally, interesting work could be done by assuming evolving networks of agents that change due to agent experience and learning [129]

Concerning the search rules, an implicit assumption in this model is that search process is the same for all firms in that it is modelled as a random process of trial and error plus a constant probability of observation. The main advantage of this assumption is its simplicity. However, in a real-world economic setting, agents are likely to be more strategic than just following simple incremental optimising paths and performing simple observation of their neighbours. Agents can indeed use sophisticated heuristics and/or complex adaptive strategies [79; 148; 126]. An important thing to note is that I started the simulations assuming the agents are randomly located on the fitness landscape; however, it might be more realistic to assume that agents start from near locations on the fitness landscape. Trying and comparing a variety of search strategies is an important extension of this model.

Another implicit assumption in the model is that the fitness landscape is both static and fixed. There are valid reasons to believe that fitness landscapes are neither static nor fixed and this creates potential to add a dimension to the model that incorporates dynamic complexity. Fitness landscapes are static if environmental conditions that influence the fitness of solutions remain stable over time. Phenomena whose context is rapidly changing will eventually result in the peaks and valleys moving up and down as time goes on. Furthermore, the fitness of solutions in the landscape can change as a result of the actions of an agent or a collection of agents acting and reacting to each other (endogenous factors), or as a result of external shocks like a new regulation that reduces the profitability of an investment or product (exogenous factors). The assumption of a fixed landscape can also be challenged, because the continuous emergence of novelty contributes to the expansion of fitness landscapes through time. Novelty creates new niches and opportunities that can be discovered.

Moving to the case study presented in section 5, some work still needs to be done on theoretical and empirical modelling. First, the model assumes no search costs both for the hubs who explore the landscape and for the peripheral nodes who observe and acquire information. A natural extension, therefore, would be to include search and observation costs in the model. As I show in the appendix, data is available on the amount of money that each bank invested, and this would be used as an indicator of resources. Second, links between banks are assumed to be unweighted, but in reality, banks appear together in the DealScan database with different frequency, and therefore it would make sense to consider weighted links in the simulations. Third, the regulator is assumed to accept convergence with a fixed probability, but a more realistic scenario would involve fitness evaluation and learning on the side of regulators so that they would allow convergence with changing probabilities according to factors such as fitness and diversity.

Finally, I would mention that among the most important limitations of this paper is the high computational costs and waiting time needed for running the simulations. Searching a vast fitness landscape like the one in used in this thesis is computationally demanding, especially if the network of agents is large as in the case study in chapter 5.

The results presented in this thesis are meant to illustrate (explain) the nature of the interaction between network structure and technological innovation on fitness landscapes. The models presented therefore are explanatory in nature and have no predictive power given the stage of the research. The next steps would be to work on predictive models that can be validated and statistically tested to provide robust and valid results for policymaking.

References

[1] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. Mining association rules between sets of items in large databases. In *Acm sigmod record*, volume 22, pages 207–216. ACM, 1993.

[2] Floortje Alkemade, Koen Frenken, Marko P Hekkert, and Malte Schwoon. A complex systems methodology to transition management. *Journal of Evolutionary Economics*, 19 (4):527, 2009.

[3] Robert C Allen. Collective invention. *Journal of Economic Behavior & Organization*, 4 (1):1–24, 1983.

[4] Thomas J Allen et al. Managing the flow of technology: Technology transfer and the dissemination of technological information within the r&d organization. *MIT Press Books*, 1, 1984.

[5] Lee Altenberg. Fitness landscapes: Nk landscapes. *Handbook of evolutionary computation*, pages 5–10, 1997.

[6] Philip Anderson. Perspective: Complexity theory and organization science. *Organization science*, 10(3):216–232, 1999.

[7] W Brian Arthur. Competing technologies, increasing returns, and lock-in by historical events. *The economic journal*, 99(394):116–131, 1989.

[8] W Brian Arthur. Why do things become more complex. *Scientific American*, 268(5): 144, 1993.

[9] W Brian Arthur. *The nature of technology: What it is and how it evolves*. Simon and Schuster, 2009.

[10] W Graham Astley. Administrative science as socially constructed truth. *Administrative Science Quarterly*, pages 497–513, 1985.

[11] Venkatesh Bala and Sanjeev Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.

[12] Abhijit V Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992.

[13] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[14] Daniel Barkoczi and Mirta Galesic. Social learning strategies modify the effect of network structure on group performance. *Nature communications*, 7, 2016.

[15] Daniel Barkoczi, Pantelis P Analytis, and Charley Wu. Collective search on rugged landscapes: A cross-environmental analysis. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, pages 918–923. Cognitive Science Society, 2016.

[16] John M Bates and Clive WJ Granger. The combination of forecasts. *Journal of the Operational Research Society*, 20(4):451–468, 1969.

[17] Eric D Beinhocker. Robust adaptive strategies. *Sloan management review*, 40(3):95, 1999.

[18] Ben Bernanke. Modern risk management and banking supervision. Technical report, Board of Governors of the Federal Reserve System (US), 2006.

[19] James Bessen and Eric Maskin. Sequential innovation, patents, and imitation. *The RAND Journal of Economics*, 40(4):611–635, 2009.

[20] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100 (5):992–1026, 1992.

[21] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives*, 12(3):151–170, 1998.

[22] Stephan Billinger, Nils Stieglitz, and Terry R Schumacher. Search on rugged landscapes: An experimental study. *Organization Science*, 25(1):93–108, 2013.

[23] Norman Blaikie. Approaches to social inquiry . malden, ma, 2007.

[24] Stephen P Borgatti and Rob Cross. A relational view of information seeking and learning in social networks. *Management science*, 49(4):432–445, 2003.

[25] Stefano Breschi and Franco Malerba. *Clusters, networks and innovation*. Oxford University Press, 2005.

[26] Stefano Breschi, Franco Malerba, and Luigi Orsenigo. Technological regimes and schumpeterian patterns of innovation. *The economic journal*, 110(463):388–410, 2000.

[27] Stefano Brusoni, Luigi Marengo, Andrea Prencipe, and Marco Valente. The value and costs of modularity: a problem-solving perspective. *European Management Review*, 4 (2):121–132, 2007.

[28] Gibson Burrell and Gareth Morgan. *Sociological paradigms and organisational analysis: Elements of the sociology of corporate life*. Routledge, 2017.

[29] Ronald S Burt. *Structural holes: The social structure of competition*. Harvard university press, 2009.

[30] Martin Campbell-Kelly, William Aspray, Daniel P Snowman, Susan R McKay, Wolfgang Christian, et al. Computer a history of the information machine. *Computers in Physics*, 11(3):256–257, 1997.

[31] John Cantwell, Alfonso Gambardella, and Ove Granstrand. *The economics and management of technological diversification*, volume 34. Routledge, 2004.

[32] John L Casti. *Connectivity, complexity and catastrophe in large-scale systems*, volume 7. John Wiley & Sons, 1979.

[33] Grazia Cecere, Nicoletta Corrocher, Cédric Gossart, and Muge Ozman. Lock-in and path dependence: an evolutionary approach to eco-innovations. *Journal of Evolutionary Economics*, 24(5):1037–1065, 2014.

[34] Claudia Champagne and Lawrence Kryzanowski. Are current syndicated loan alliances related to past alliances? *Journal of Banking & Finance*, 31(10):3145–3161, 2007.

[35] Sudheer Chava and Michael R Roberts. How does financing impact investment? the role of debt covenants. *The Journal of Finance*, 63(5):2085–2121, 2008.

[36] Shi-Jie Gary Chen and Enzhen Huang. A systematic approach for supply chain improvement using design structure matrix. *Journal of Intelligent Manufacturing*, 18(2): 285–299, 2007.

[37] Henry William Chesbrough. *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press, 2006.

[38] Susanne Chishti and Janos Barberis. *The FinTech book: the financial technology handbook for investors, entrepreneurs and visionaries*. John Wiley & Sons, 2016.

[39] Clayton M Christensen and Richard S Rosenbloom. Explaining the attacker's advantage: Technological paradigms, organizational dynamics, and the value network. *Research policy*, 24(2):233–257, 1995.

[40] Clayton M Christensen, Michael E Raynor, and Rory McDonald. What is disruptive innovation. *Harvard Business Review*, 93(12):44–53, 2015.

[41] Paul Cilliers. Why we cannot know complex things completely. *Emergence*, 4(1-2): 77–84, 2002.

[42] Aaron Clauset, Cosma Rohilla Shalizi, and Mark EJ Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

[43] Wesley M Cohen and Daniel A Levinthal. Innovation and learning: the two faces of r & d. *The economic journal*, 99(397):569–596, 1989.

[44] Wesley M Cohen and Daniel A Levinthal. Absorptive capacity: A new perspective on learning and innovation. *Administrative science quarterly*, pages 128–152, 1990.

[45] Arnold C Cooper and Dan Schendel. Strategic responses to technological threats. *Business horizons*, 19(1):61–69, 1976.

[46] Robin Cowan and Nicolas Jonard. The dynamics of collective invention. *Journal of Economic Behavior & Organization*, 52(4):513–532, 2003.

[47] A CPSS. Glossary of terms used in payments and settlement systems. *CPSS-Committee on Payment and*, 2003.

[48] Peter Csermely, András London, Ling-Yun Wu, and Brian Uzzi. Structure and dynamics of core/periphery networks. *Journal of Complex Networks*, 1(2):93–123, 2013.

[49] Linus Dahlander and David M Gann. How open is innovation? *Research policy*, 39(6): 699–709, 2010.

[50] Kristina B Dahlin and Dean M Behrens. When is an invention really radical?: Defining and measuring technological radicalness. *research policy*, 34(5):717–737, 2005.

[51] Paul A David. Path-dependence and predictability in dynamic systems with local network externalities: a paradigm for historical economics. *Technology and the Wealth of Nations*, pages 208–231, 1993.

[52] Chris DeBresson and Fernand Amesse. Networks of innovators: A review and introduction to the issue. *Research policy*, 20(5):363–379, 1991.

[53] Maxime Derex and Robert Boyd. Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences*, 113(11):2982–2987, 2016.

[54] Melinda Deutsch, Clive WJ Granger, and Timo Teräsvirta. The combination of forecasts using changing weights. *International Journal of Forecasting*, 10(1):47–57, 1994.

[55] Paul DiMaggio and Walter W Powell. The iron cage revisited: Collective rationality and institutional isomorphism in organizational fields. *American sociological review*, 48(2): 147–160, 1983.

[56] Mark Dodgson. *Technological collaboration in industry: strategy, policy and internationalization in innovation*. Routledge, 2018.

[57] Walter Dolde. The trajectory of corporate financial risk management. *Journal of Applied Corporate Finance*, 6(3):33–41, 1993.

[58] Giovanni Dosi. Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. *Research policy*, 11(3):147–162, 1982.

[59] Giovanni Dosi and Luigi Marengo. The dynamics of organizational structures and performances under diverging distributions of knowledge and different power structures. *Journal of Institutional Economics*, 11(3):535–559, 2015.

[60] Giovanni Dosi and Luigi Orsenigo. Coordination and transformation: an overview of structures, behaviours and change in evolutionary environments. *Technical change and economic theory*, pages 13–37, 1988.

[61] Giovanni Dosi, Christopher Freeman, Richard Nelson, Gerald Silverberg, Luc Soete, et al. *Technical change and economic theory*, volume 988. Pinter London, 1988.

[62] Giovanni Dosi, Giorgio Fagiolo, and Andrea Roventini. Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767, 2010.

[63] Nicholas Economides. Network externalities, complementarities, and invitations to enter. *European Journal of Political Economy*, 12(2):211–233, 1996.

[64] Steven D Eppinger and Tyson R Browning. *Design structure matrix methods and applications*. MIT press, 2012.

[65] Jan Fagerberg, David C Mowery, and Richard R Nelson. *The Oxford handbook of innovation*. Oxford university press, 2005.

[66] Giorgio Fagiolo, Luigi Marengo, and Marco Valente. Population learning in a model with random payoff landscapes and endogenous networks. *Computational Economics*, 24(4):383–408, 2005.

[67] J Doyne Farmer and François Lafond. How predictable is technological progress? *Research Policy*, 45(3):647–665, 2016.

[68] Lawrence F Feick and Linda L Price. The market maven: A diffuser of marketplace information. *The Journal of Marketing*, pages 83–97, 1987.

[69] Teppo Felin and Todd R Zenger. Closed or open innovation? problem solving and the governance choice. *Research Policy*, 43(5):914–925, 2014.

[70] Teppo Felin, Stuart Kauffman, Roger Koppl, and Giuseppe Longo. Economic opportunity and evolution: Beyond landscapes and bounded rationality. *Strategic Entrepreneurship Journal*, 8(4):269–282, 2014.

[71] Peter Field. *Modern risk management: a history*. Risk, 2003.

[72] John C Fisher and Robert H Pry. A simple substitution model of technological change. *Technological forecasting and social change*, 3:75–88, 1971.

[73] Lee Fleming and Olav Sorenson. Technology as a complex adaptive system: evidence from patent data. *Research policy*, 30(7):1019–1039, 2001.

[74] Lee Fleming, Charles King III, and Adam I Juda. Small worlds and regional innovation. *Organization Science*, 18(6):938–954, 2007.

[75] Christopher Freeman and Luc Soete. *The economics of industrial innovation*. Psychology Press, 1997.

[76] Christopher Freeman et al. Networks of innovators: a synthesis of research issues. *Research policy*, 20(5):499–514, 1991.

[77] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

[78] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.

[79] Koen Frenken. Fitness landscapes, heuristics and technological paradigms: a critique on random search models in evolutionary economics. In *AIP Conference Proceedings*, volume 573, pages 558–565. AIP, 2001.

[80] Koen Frenken. A fitness landscape approach to technological complexity, modularity, and vertical disintegration. *Structural Change and Economic Dynamics*, 17(3):288–305, 2006.

[81] Koen Frenken. Technological innovation and complexity theory. *Economics of Innovation and New Technology*, 15(2):137–155, 2006.

[82] Douglas Gale and Shachar Kariv. Bayesian learning in social networks. *Games and Economic Behavior*, 45(2):329–346, 2003.

[83] Andrea Galeotti. One-way flow networks: the role of heterogeneity. *Economic Theory*, 29(1):163–179, 2006.

[84] Andrea Galeotti and Sanjeev Goyal. The law of the few. *The American Economic Review*, 100(4):1468–1492, 2010.

[85] Giovanni Gavetti and Daniel Levinthal. Looking forward and looking backward: Cognitive and experiential search. *Administrative science quarterly*, 45(1):113–137, 2000.

[86] Sylvie Geisendorf. Searching nk fitness landscapes: On the trade off between speed and quality in complex problem solving. *Computational Economics*, 35(4):395–406, 2010.

[87] Sylvie Geisendorf. The impact of personal beliefs on climate change: the âĂIJbattle of perspectivesâĂİ revisited. *Journal of Evolutionary Economics*, 26(3):551–580, 2016.

[88] Gerd Gigerenzer and Reinhard Selten. *Bounded rationality: The adaptive toolbox*. MIT press, 2002.

[89] Richard Gilbert. Looking for mr. schumpeter: Where are we in the competition–innovation debate? *Innovation policy and the economy*, 6:159–215, 2006.

[90] Victor Gilsing, Bart Nooteboom, Wim Vanhaverbeke, Geert Duysters, and Ad van den Oord. Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research policy*, 37(10):1717–1731, 2008.

[91] Dennis A Gioia. Give it up! reflections on the interpreted world (a commentary on meckler and baillie). *Journal of Management Inquiry*, 12(3):285–292, 2003.

[92] Sanjeev Goyal. *Connections: an introduction to the economics of networks*. Princeton University Press, 2012.

[93] Sanjeev Goyal. Heterogeneity and networks. In *Handbook of Computational Economics*, volume 4, pages 687–712. Elsevier, 2018.

[94] Mark S Granovetter. The strength of weak ties. *American journal of sociology*, 78(6): 1360–1380, 1973.

[95] Henrich Greve, Tim Rowley, and Andrew Shipilov. *Network advantage: How to unlock value from your alliances and partnerships*. John Wiley & Sons, 2013.

[96] Jiancheng Guan and Na Liu. Exploitative and exploratory innovations in knowledge network and collaboration network: A patent analysis in the technological field of nano-energy. *Research policy*, 45(1):97–112, 2016.

[97] Till Guldimann. The story of riskmetrics. *Risk*, 13(1):56–58, 2000.

[98] Anil K Gupta, Ken G Smith, and Christina E Shalley. The interplay between exploration and exploitation. *Academy of management journal*, 49(4):693–706, 2006.

[99] Friedrich August Hayek. The use of knowledge in society. *The American economic review*, pages 519–530, 1945.

[100] Rebecca M Henderson and Kim B Clark. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative science quarterly*, pages 9–30, 1990.

[101] Diana Hicks. Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and corporate change*, 4(2):401–424, 1995.

[102] Eric von Hippel and Georg von Krogh. Open source software and the âĂIJprivate-collectiveâĂİ innovation model: Issues for organization science. *Organization science*, 14(2):209–223, 2003.

[103] Glyn A Holton. History of value-at-risk, 2002.

[104] Glyn A Holton. *Value-at-risk*. Acad. press, 2003.

[105] Werner Hölzl. Convergence of financial systems: towards an evolutionary perspective. *Journal of Institutional Economics*, 2(1):67–90, 2006.

[106] Thomas Parke Hughes. *Networks of power: electrification in Western society, 1880-1930*. JHU Press, 1993.

[107] Michael Hutter. Transaction cost and communication: A theory of institutional change, applied to the case of patent law. In *Law and economics and the economics of legal regulation*, pages 113–129. Springer, 1986.

[108] Matthew O Jackson. A survey of network formation models: stability and efficiency. *Group Formation in Economics: Networks, Clubs, and Coalitions*, pages 11–49, 2005.

[109] Matthew O Jackson. *Social and economic networks*. Princeton university press, 2010.

[110] Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996.

[111] Börje Johansson, Charlie Karlsson, and Lars Westin. *Patterns of a network economy*. Springer Science & Business Media, 2012.

[112] Peter Stewart Johnson. *Co-operative research in industry: an economic study*. Halsted Press, 1973.

[113] Phil Johnson and Joanne Duberley. *Understanding management research: An introduction to epistemology*. Sage, 2000.

[114] Steven Johnson. *Where good ideas come from: the seven patterns of innovation*. Penguin UK, 2011.

[115] Riitta Katila and Gautam Ahuja. Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of management journal*, 45 (6):1183–1194, 2002.

[116] Elihu Katz and Paul Felix Lazarsfeld. *Personal Influence, The part played by people in the flow of mass communications*. Transaction Publishers, 1966.

[117] Michael L Katz and Carl Shapiro. Technology adoption in the presence of network externalities. *Journal of political economy*, 94(4):822–841, 1986.

[118] SA Kauffman, Stefan Thurner, and Rudolf Hanel. The evolving web of future wealth. *Scientific American*, 2008.

[119] Stuart Kauffman and Simon Levin. Towards a general theory of adaptive walks on rugged landscapes. *Journal of theoretical Biology*, 128(1):11–45, 1987.

[120] Stuart Kauffman and William Macready. Technological evolution and adaptive organizations: Ideas from biology may find applications in economics. *Complexity*, 1(2): 26–43, 1995.

[121] Stuart Kauffman, José Lobo, and William G Macready. Optimal search on a technology landscape. *Journal of Economic Behavior & Organization*, 43(2):141–166, 2000.

[122] Stuart A Kauffman. *Investigations*. Oxford University Press, 2000.

[123] Stuart A Kauffman. *Reinventing the sacred: A new view of science, reason, and religion*. Basic Books, 2008.

[124] Stuart A Kauffman and Sonke Johnsen. Coevolution to the edge of chaos: coupled fitness landscapes, poised states, and coevolutionary avalanches. *Journal of theoretical biology*, 149(4):467–505, 1991.

[125] Kevin Kelly. *The inevitable: understanding the 12 technological forces that will shape our future*. Penguin, 2017.

[126] Tamer Khraisha. Complex economic problems and fitness landscapes: Assessment and methodological perspectives. *Structural Change and Economic Dynamics*, (in press) https://doi.org/10.1016/j.strueco.2019.01.002, 2019.

[127] Tamer Khraisha and Keren Arthur. Can we have a general theory of financial innovation processes? a conceptual review. *Financial Innovation*, 4(1):4, 2018.

[128] Scott Kirkpatrick, C Daniel Gelatt, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[129] Alan Kirman. The economy as an evolving network. *Journal of evolutionary economics*, 7(4):339–353, 1997.

[130] David Kreps and Kenneth Wallis. Advances in economics and econometrics: Theory and applications 3 volume paperback set. Technical report, Cambridge University Press, 1997.

[131] Paul Felix Lazarsfeld, Bernard Berelson, and Hazel Gaudet. *The people's choice: how the voter makes up his mind in a presidential campaign, by Paul F. Lazarsfeld [et al.]*. Columbia Univ. Press, 1965.

[132] David Lazer and Allan Friedman. The network structure of exploration and exploitation. *Administrative Science Quarterly*, 52(4):667–694, 2007.

[133] Aija Leiponen and Ina Drejer. What exactly are technological regimes?: Intra-industry heterogeneity in the organization of innovation activities. *Research Policy*, 36(8):1221–1238, 2007.

[134] Daniel Levinthal and James G March. A model of adaptive organizational search. *Journal of Economic Behavior & Organization*, 2(4):307–333, 1981.

[135] Daniel A Levinthal. Adaptation on rugged landscapes. *Management science*, 43(7): 934–950, 1997.

[136] Daniel A Levinthal and Massimo Warglien. Landscape design: Designing for local action in complex worlds. *Organization Science*, 10(3):342–357, 1999.

[137] Stanley J Liebowitz and Stephen E Margolis. *Winners, Losers & Microsoft; Competition and Antitrust in High Technology*. Independent Institute, 1999.

[138] Andrew W Lo. *Adaptive markets: Financial evolution at the speed of thought*. Princeton University Press, 2017.

[139] Henry Lopez-Vega, Fredrik Tell, and Wim Vanhaverbeke. Where and how to search? search paths in open innovation. *Research Policy*, 45(1):125–136, 2016.

[140] Vittorio Loreto, Vito DP Servedio, Steven H Strogatz, and Francesca Tria. Dynamics on expanding spaces: modeling the emergence of novelties. In *Creativity and universality in language*, pages 59–83. Springer, 2016.

[141] Franco Malerba and Luigi Orsenigo. Technological regimes and firm behaviour. In *Organization and strategy in the evolution of the enterprise*, pages 42–71. Springer, 1996.

[142] Franco Malerba, Richard Nelson, Luigi Orsenigo, and Sidney Winter. 'history-friendly'models of industry evolution: the computer industry. *Industrial and corporate change*, 8(1):3–40, 1999.

[143] Benoit Mandelbrot and Richard L Hudson. *The Misbehavior of Markets: A fractal view of financial turbulence*. Basic books, 2007.

[144] James G March. Exploration and exploitation in organizational learning. *Organization science*, 2(1):71–87, 1991.

[145] James G March and Herbert Alexander Simon. Organizations. 1958.

[146] Luigi Marengo and Giovanni Dosi. Division of labor, organizational coordination and market mechanisms in collective problem-solving. *Journal of Economic Behavior & Organization*, 58(2):303–326, 2005.

[147] Luigi Marengo and Marco Valente. Industry dynamics in complex product spaces: An evolutionary model. *Structural Change and Economic Dynamics*, 21(1):5–16, 2010.

[148] Luigi Marengo, Giovanni Dosi, Paolo Legrenzi, and Corrado Pasquali. The structure of problem-solving knowledge and the structure of organizations. *Industrial and Corporate Change*, 9(4):757–788, 2000.

[149] Donald G Marquis. The anatomy of successful innovations. *Innovation*, 1(7):28–37, 1969.

[150] Christopher Marshall and Michael Siegel. Value-at-risk: Implementing a risk measurement standard. 1996.

[151] Winter Mason and Duncan J Watts. Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, 109(3):764–769, 2012.

[152] James McNerney, J Doyne Farmer, Sidney Redner, and Jessika E Trancik. Role of design complexity in technology improvement. *Proceedings of the National Academy of Sciences*, 108(22):9008–9013, 2011.

[153] J Stanley Metcalfe. Evolutionary economics and technology policy. *The economic journal*, 104(425):931–944, 1994.

[154] David Mowery and Nathan Rosenberg. The influence of market demand upon innovation: a critical review of some recent empirical studies. *Research policy*, 8(2):102–153, 1979.

[155] Toshihiko Mukoyama. Innovation, imitation, and growth with cumulative technology. *Journal of Monetary Economics*, 50(2):361–380, 2003.

[156] Richard R Nelson. The simple economics of basic scientific research. *Journal of political economy*, 67(3):297–306, 1959.

[157] Richard R Nelson and Sidney G Winter. Evolutionary theorizing in economics. *Journal of Economic Perspectives*, 16(2):23–46, 2002.

[158] Jack A Nickerson and Todd R Zenger. A knowledge-based theory of the firmâĂŤthe problem-solving perspective. *Organization science*, 15(6):617–632, 2004.

[159] Bart Nooteboom. Innovation and diffusion in small firms: theory and evidence. *Small Business Economics*, 6(5):327–347, 1994.

[160] Gabriela Ochoa, Sébastien Verel, Fabio Daolio, and Marco Tomassini. Local optima networks: A new model of combinatorial fitness landscapes. In *Recent Advances in the Theory and Application of Fitness Landscapes*, pages 233–262. Springer, 2014.

[161] Kenichi Ohmae. The borderless world: Power and strategy in the interlinked economy. Technical report, HARPER PERENNIAL,, 1991.

[162] Margit Osterloh and Sandra Rota. Open source software developmentâĂŤjust another case of collective invention? *Research Policy*, 36(2):157–171, 2007.

[163] Tommaso Padoa-Schioppa. *Regulating finance: balancing freedom and risk*. Oxford University Press on Demand, 2004.

[164] Giuliano Andrea Pagani and Marco Aiello. The power grid as a complex network: a survey. *Physica A: Statistical Mechanics and its Applications*, 392(11):2688–2700, 2013.

[165] Ugo Pagano. Interlocking complementarities and institutional change. *Journal of Institutional Economics*, 7(3):373–392, 2011.

[166] Carlota Perez. *Technological revolutions and financial capital*. Edward Elgar Publishing, 2003.

[167] Linda D Peters, Wesley J Johnston, Andrew D Pressey, and Terry Kendrick. Collaboration and collective learning: networks as learning organisations. *Journal of Business & Industrial Marketing*, 25(6):478–484, 2010.

[168] Gary P Pisano. Knowledge, integration, and the locus of learning: An empirical analysis of process development. *Strategic management journal*, 15(S1):85–100, 1994.

[169] Erik Pitzer and Michael Affenzeller. A comprehensive survey on fitness landscape analysis. *Recent Advances in Intelligent Engineering Systems*, 378:161–191, 2012.

[170] Walter W Powell and Eric Giannella. Collective invention and inventor networks. *Handbook of the Economics of Innovation*, 1:575–605, 2010.

[171] Walter W Powell, Kenneth W Koput, and Laurel Smith-Doerr. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative science quarterly*, pages 116–145, 1996.

[172] Brigitte Preissl and Laura Solimene. *Dynamics of Clusters and Innovation*. Springer, 2003.

[173] Sebastian Raisch, Julian Birkinshaw, Gilbert Probst, and Michael L Tushman. Organizational ambidexterity: Balancing exploitation and exploration for sustained performance. *Organization science*, 20(4):685–695, 2009.

[174] Ray Reagans and Bill McEvily. Network structure and knowledge transfer: The effects of cohesion and range. *Administrative science quarterly*, 48(2):240–267, 2003.

[175] Maxine Robertson, Jacky Swan, and Sue Newell. The role of networks in the diffusion of technological innovation. *Journal of Management Studies*, 33(3):333–359, 1996.

[176] Paul L Robertson, Gian Luca Casali, and David Jacobson. Managing open incremental process innovation: absorptive capacity and distributed learning. *Research policy*, 41(5): 822–832, 2012.

[177] Everett M Rogers. *Diffusion of innovations*. Simon and Schuster, 2010.

[178] Nathan Rosenberg. *Perspectives on technology*. CUP Archive, 1976.

[179] Nathan Rosenberg. Why do firms do basic research (with their own money)? In *Studies On Science And The Innovation Process: Selected Works of Nathan Rosenberg*, pages 225–234. World Scientific, 2010.

[180] Nathan Rosenberg and Richard R Nelson. American universities and technical advance in industry. *Research policy*, 23(3):323–348, 1994.

[181] Lori Rosenkopf and Paul Almeida. Overcoming local search through alliances and mobility. *Management science*, 49(6):751–766, 2003.

[182] Francisco C Santos and Jorge M Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters*, 95(9):098104, 2005.

[183] Jari Saramäki, Mikko Kivelä, Jukka-Pekka Onnela, Kimmo Kaski, and Janos Kertesz. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E*, 75(2):027105, 2007.

[184] Jan Schapper, Helen De Cieri, and Julie Wolfram Cox. The ontological and epistemological dimensions of complex organizations. In *ANZAM 2005: Engaging the multiple contexts of management: convergence and divergence of management theory and practice: proceedings of the 19th ANZAM conference*. ANZAM, 2005.

[185] H Schnabel. Agenda-diffusion and innovation: A simulation model. *Journal of Evolutionary Economics*, pages 65–85, 1981.

[186] Thomas Schøtt and Kent Wickstrøm Jensen. FirmsâĂŹ innovation benefiting from networking and institutional support: A global analysis of national and firm effects. *Research Policy*, 45(6):1233–1246, 2016.

[187] Stephan Schrader. Informal technology transfer between firms: Cooperation through information trading. *Research policy*, 20(2):153–170, 1991.

[188] Joseph Alois Schumpeter. *The theory of economic developments: an inquiry into profits, capital, credit, interest, and the business cycle*. Harvard University Press, 1961.

[189] Thomas Shafee. *Evolvability of a viral protease: experimental evolution of catalysis, robustness and specificity*. PhD thesis, University of Cambridge, 2014.

[190] Gerald Silverberg. Adoption and diffusion of technology as a collective evolutionary process. In *Diffusion of technologies and social behavior*, pages 209–229. Springer, 1991.

[191] Gerald Silverberg and Bart Verspagen. A percolation model of innovation in complex technology spaces. *Journal of Economic Dynamics and Control*, 29(1):225–244, 2005.

[192] Herbert A Simon. The architecture of complexity. In *Facets of systems science*, pages 457–476. Springer, 1991.

[193] Andrea Sironi and Andrea Resti. *Risk management and shareholders' value in banking: from risk measurement models to capital allocation policies*, volume 417. John Wiley & Sons, 2007.

[194] Paolo Sironi. *FinTech innovation: from robo-advisors to goal based investing and gamification*. John Wiley & Sons, 2016.

[195] Ricard V Solée, Sergi Valverde, Marti Rosas Casals, Stuart A Kauffman, Doyne Farmer, and Niles Eldredge. The evolutionary ecology of technological innovations. *Complexity*, 18(4):15–27, 2013.

[196] Robert M Solow. Technical change and the aggregate production function. *The review of Economics and Statistics*, pages 312–320, 1957.

[197] Olav Sorenson, Jan W Rivkin, and Lee Fleming. Complexity, networks and knowledge flow. *Research policy*, 35(7):994–1017, 2006.

[198] David Strang and Michael W Macy. In search of excellence: Fads, success stories, and adaptive emulation. *American journal of sociology*, 107(1):147–182, 2001.

[199] Gabriel Szulanski. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic management journal*, 17(S2):27–43, 1996.

[200] Josef Taalbi. What drives innovation? evidence from economic history. *Research Policy*, 2017.

[201] Nassim Taleb et al. *Dynamic hedging: managing vanilla and exotic options*, volume 64. John Wiley & Sons, 1997.

[202] David J Teece. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research policy*, 15(6):285–305, 1986.

[203] David J Teece. Competition, cooperation, and innovation: Organizational arrangements for regimes of rapid technological progress. *Journal of economic behavior & organization*, 18(1):1–25, 1992.

[204] David J Teece, Gary Pisano, and Amy Shuen. Dynamic capabilities and strategic management. *Strategic management journal*, 18(7):509–533, 1997.

[205] Morris Teubal, Tamar Yinnon, and Ehud Zuscovitch. Networks and market creation. *Research Policy*, 20(5):381–392, 1991.

[206] Stefan Thurner, Rudolf Hanel, and Peter Klimek. *Introduction to the theory of complex systems*. Oxford University Press, 2018.

[207] Francesca Tria, Vittorio Loreto, Vito Domenico Pietro Servedio, and Steven H Strogatz. The dynamics of correlated novelties. *Scientific reports*, 4:5890, 2014.

[208] Mary Tripsas. Unraveling the process of creative destruction: Complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal*, 18 (S1):119–142, 1997.

[209] Andrew H Van de Ven, Marshall Scott Poole, et al. Explaining development and change in organizations. *Academy of management review*, 20(3):510–540, 1995.

[210] Jan Van den Ende and Wilfred Dolfsma. Technology-push, demand-pull and the shaping of technological paradigms-patterns in the development of computing technology. *Journal of Evolutionary Economics*, 15(1):83–99, 2005.

[211] Vesselin K Vassilev, Terence C Fogarty, and Julian F Miller. Information characteristics and the structure of landscapes. *Evolutionary computation*, 8(1):31–60, 2000.

[212] Bala Venkatesh and Sanjeev Goyal. Learning from neighbors. *Review of Economic Studies*, 65:595–621, 1998.

[213] Bart Verspagen. Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research. *Advances in Complex Systems*, 10(01):93–115, 2007.

[214] Eric Von Hippel. Cooperation between rivals: informal know-how trading. *Research policy*, 16(6):291–302, 1987.

[215] Andreas Wagner and William Rosen. Spaces of the possible: universal darwinism and the wall between technological and biological innovation. *Journal of The Royal Society Interface*, 11(97):20131190, 2014.

[216] Frank H Westhoff, Beth V Yarbrough, and Robert M Yarbrough. Complexity, organization, and stuart kauffman's the origins of order. *Journal of Economic Behavior & Organization*, 29(1):1–25, 1996.

[217] Paul Windrum, Giorgio Fagiolo, and Alessio Moneta. Empirical validation of agent-based models: Alternatives and prospects. *Journal of Artificial Societies and Social Simulation*, 10(2):8, 2007.

[218] Ulrich Witt. *The political economy of mass media societies*. Max-Planck-Inst. for Research into Economic Systems, 1996.

[219] Ulrich Witt. Economic policy making in evolutionary perspective. *Journal of Evolutionary Economics*, 13(2):77–94, 2003.

[220] Shaker A Zahra and Gerard George. Absorptive capacity: A review, reconceptualization, and extension. *Academy of management review*, 27(2):185–203, 2002.

[221] Jun Zhang, Mark S Ackerman, and Lada Adamic. Expertise networks in online communities: structure and algorithms. In *Proceedings of the 16th international conference on World Wide Web*, pages 221–230. ACM, 2007.

8 Appendix

A G30 Supporters

1. American International Group
2. Bank of America
3. Bank of Tokyo
4. Bankers Trust Company
5. Banque Indosuez
6. Banque Paribas
7. Barclays Bank
8. Chase Manhattan Bank
9. Chemical Bank
10. Citicorp
11. Commerzbank
12. Crédit Suisse
13. CS First Boston
14. Dai-ichi Kangyo Bank
15. Daiwa Securities
16. Deutsche Bank
17. Dresdner Bank
18. Fidelity Investments
19. First Chicago
20. Fuji Bank
21. Goldman Sachs
22. HSBC Holdings
23. Industrial Bank of Japan
24. J.P. Morgan
25. Kredietbank
26. Lloyds Bank
27. Merrill Lynch
28. Mitsubishi Bank
29. Morgan Stanley
30. NatWest Markets
31. Nikko Securities
32. Nomura Securities
33. O'Connor & Associates
34. Price Waterhouse
35. Royal Bank of Canada
36. S.G. Warburg Group
37. Sakura Bank
38. Salomon Brothers
39. Sanwa Bank
40. Société Générale
41. Standard Chartered
42. State Street Bank and Trust
43. Sumitomo Bank
44. Swiss Bank Corporation
45. Union Bank of Switzerland
46. Yamaichi Securities

B Amounts Invested by the Biggest Banks in 1993

	Bank Name	Amount Invested in 1993 in Bln Dollars
0	Bank of America	56.7408
1	Bank of Boston	44.5284
2	Chase Manhattan Bank	34.1471
3	Norwest Business Credit	32.5562
4	NationsBank	27.3572
5	Chemical Bank	26.5995
6	First Chicago	25.8042
7	NBD Bank NA	24.7670
8	Bank of Nova Scotia	24.3873
9	NationsBank of Tennessee	23.1395
10	First American National Bank Nashville	21.4407
11	NationsBank of Texas	20.8987
12	CIBC [Canadian Imperial Bank of Commerce]	18.3381
13	Bank of New York	17.9929
14	Boatmen's National Bank of St Louis	17.5128
15	Frost National Bank	16.1927
16	Morgan Guaranty Trust	16.0352
17	Le Credit Lyonnais SA [LCL]	15.0338
18	Heller Financial Inc	14.4798
19	Bank IV Oklahoma	13.7440
20	PNC Bank	13.6666
21	BNP Paribas [Ex-Banque Paribas]	13.6367
22	Citibank	13.3576
23	Comerica Bank	12.9459
24	Society National Bank	12.8809
25	Sanwa Business Credit Corp	12.6018
26	Societe Generale SA	12.2444
27	ABN AMRO Bank NV [RBS]	12.2359
28	Bank One Texas	12.2033
29	Wachovia Bank of Georgia	12.1108
30	Mellon Bank	11.8826
31	First Union National Bank of North Carolina	11.6368
32	CoreStates Bank	10.2287
33	Deposit Guaranty National Bank	10.0850
34	Credit Suisse AG	10.0412
35	Silicon Valley Bank	9.8257
36	Long-Term Credit Bank of Japan Ltd	9.6985
37	First Source Financial Inc	9.4762
38	Texas Commerce Bank	9.3234
39	LaSalle National Bank	9.1942
40	Fuji Bank Ltd	9.0198