



HAL
open science

The dynamics of technological knowledge

Jackie Krafft, Francesco Quatraro

► **To cite this version:**

Jackie Krafft, Francesco Quatraro. The dynamics of technological knowledge. Handbook on System Dynamics of Technological Change, Edward Elgar, pp.181-200, 2011. halshs-00727633

HAL Id: halshs-00727633

<https://shs.hal.science/halshs-00727633>

Submitted on 4 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Dynamics of Technological Knowledge: From Linearity to Recombination

Jackie Krafft^a and Francesco Quatraro^{a,b,c}

^a University of Nice “Sophia Antipolis”

^b BRICK, Collegio Carlo Alberto, Torino

^c International Centre for Economic Research (ICER), Torino

1 Introduction

Innovation and technological knowledge have long attracted the interest of scholars in economics. Most of the attention has been paid by the pioneers in the economics of innovation on the economic effects of the introduction of new technological knowledge as well as on the structural conditions better triggering innovative performances. This has paved the way to an empirically grounded research tradition which has initially considered knowledge as an homogeneous stock, as if it were the outcome of a quite uniform and fluid process of accumulation made possible by R&D investments, the same way as capital stock. This made it possible to include knowledge capital stock within an extended production function framework, as an additional input to labour and fixed capital (Griliches, 1979; Mansfield, 1980).

The focus therein was on the empirical assessment of the impact of technological knowledge on economic performances. Yet, very little was known about how new knowledge is brought about and, consequently, about how to provide a representation of knowledge that could be meaningful also from the epistemological viewpoint. Technology was mostly a black box, which begun to be explored in depth with a significant lapse of time. The idea progressively arose that knowledge was something more than the mere outcome of a linear accumulation process. Indeed such an idea was grounded on theoretical reflections on the nature of knowledge creation processes, with a particular emphasis on the concept of search and on the institutions involved in the production of new technologies (Nelson, 1982 and 1986; Nelson and Winter, 1982; Rosenberg, 1982).

Drawing on insightful intuitions of Schumpeter (1912 and 1942) and Usher (1954), an increasingly share of scholars in the economics of innovation has recently elaborated theoretical approaches wherein the process of knowledge production is viewed as the outcome of a recombination process, according to which innovations stem either from the combination of brand new components or from the combination of existing components in new ways (Weitzmann, 1998; Kauffman, 1993). These theoretical efforts are in turn complemented by a well-defined cognitive approach to innovation as well as by the increasing availability of historical accounts and sectoral studies on the dynamics

of technological knowledge (Vincenti, 1990; Nightingale, 1998; Katila and Ahuja, 2002; Fleming, 2001; van der Bergh, 2008).

Such framework has been largely used to build empirical studies aimed at investigating the dynamics of knowledge from the viewpoint of the complex systems approach. Knowledge was indeed seen as a set of elements connected by a network of relationships, the architecture of which affects its performances. However, despite the emergence of these new lines of inquiry in the economics of knowledge, only a few efforts can be found in literature attempting to analyze their empirical consequences, with respect to i) the identification of the relevant properties that better approximate the concept of recombinant knowledge, and hence provide a more sensible representation of knowledge on the one hand; ii) the operational translation of such properties, as well as the identification of the most appropriate analytical tools on the other hand. Moreover, such approaches are also characterized by an important theoretical limit, according to which the architecture of knowledge structure is stable over time, i.e. complexity exogenous rather than endogenous.

This chapter aims at providing an original review of the main theoretical approaches to technological knowledge, both implicit and explicit, and of their empirical counterparts in the field of economics of innovation. While there are in the literature interesting contributions aiming at assessing the relative goodness of the different proxies used in empirical analysis of innovation (see for example Kleinknecht et al., 2002), there is a lack of efforts explicitly directed towards synthesis of theoretical and empirical issues in a historical perspective.

In this direction, we will go through the most recent debates on the dynamics of knowledge by proposing new methodologies to identifying relevant properties of knowledge that are consistent with the recombinant knowledge concept and allow for its grafting in the complex system dynamics approach in a fairly different way from the extant literature.

In particular, such methodologies are well suited to reconcile two different aspects of the analysis of the complex dynamics of technology, that is the view of technology as an artefact and as an act (Arthur, 2009; Lane et al. 2009). Indeed, by proposing that knowledge is the outcome of a collective process of recombination, we may argue that technological knowledge itself is characterized by an internal structure emerging out of a complex dynamics that is strictly connected to the dynamics affecting the formation and evolution of technology coalitions (David and Keely, 2003).

The chapter is organized as follows. Section 2 provides an overview of the main different approaches to technological knowledge, both in empirical and theoretical terms. Section 3 lays down the basic ingredients of complex system dynamics and establishes the linkages with knowledge dynamics. In Section 4 we discuss the operational implications of knowledge understood as a complex system, by proposing a set of indicators that may fit this framework. Section 5 provides the conclusion and establish an agenda for future research.

2 Technological Knowledge: From Knowledge Capital Stock to Complex Knowledge

2.1 Knowledge capital stock and the linear model

The importance of creativity for the production of goods and wealth is not a recent discovery within economics. The earlier treatment can indeed be found already in Adam Smith's first four books of the *Wealth of Nations*. After more than a century, Alfred Marshall elaborated upon Adam Smith's contribution, by proposing a former systemic account of the role of knowledge in the production process. In particular, Marshall made it very clear both in *Industry and Trade* (1919) and in the *Principles of Economics* (1920) that knowledge is a key input in the production process and the main engine of economic growth.

Despite the venerable origins of the interest in technological knowledge within the field of economics, the former attempts to provide empirical accounts of the dynamics and the effects of innovation appeared only in the late 1950s. The studies by Griliches (1957) and Mansfield (1961) on the diffusion of innovation can be viewed as the earlier empirical efforts in this sense. However, very little was known at that time about knowledge and in particular about its production and exploitation. The earlier empirical works in which the word 'knowledge' appeared to refer to a factor affecting the production of firms can be dated back to the late 1970s. Zvi Griliches turned out to be a pioneer in the field again. In his 1979 paper indeed he proposed the famous extended production function, which paved the way to a pretty wide body of empirical investigations. In such paper the traditional production function was extended so as to include an additional explanatory variable, as follows:

$$Y_i = C_i^\alpha L_i^\beta K_i^\gamma \quad (1)$$

Where C is the fixed capital stock, L stands for labour services and K is the knowledge capital used by firm i . Strangely enough, the empirical literature has generated a great deal of confusion on this contribution, as it is usually taken as key reference in papers using the so-called 'knowledge production function' approach. We believe this is due to a basic misunderstanding. Indeed, Professor Griliches in his article made some step forward to give an empirical meaning to the K term. To this purpose he proposed the following relationship:

$$K = G[W(B)R, \nu] \quad (2)$$

Where R is R&D expenditures and ν is a set of unobserved disturbances. The term $W(B)$ is instead a lag polynomial describing the relative contribution of past and present R&D expenditures to the accumulated level of knowledge. Clearly, this representation is one more application of the distributed lag literature, which influenced Griliches to a great extent. Far from proposing a knowledge production function, this relationship simply was the formalization of the concept of knowledge capital stock, which the author

subsequently used in his 1980 paper on the US productivity slowdown (Griliches, 1980). In a nutshell, the 1979 paper offered the formal basis to the application of the permanent inventory method to calculate the knowledge stock starting from R&D expenditures, which are then considered as a flow measure.

The specification of knowledge capital also called for a proper account of the effects of knowledge spillovers, i.e. knowledge borrowed or stolen from other firms or industries that can equally affect productivity of the observed firm or industry. Knowledge spillovers have been accommodated in an extended production function at the firm level by including a proxy for the aggregate stock of knowledge available within the industry firm i operates:

$$Y_i = C_i^\alpha L_i^\beta K_i^\gamma K_a^\mu \quad (3)$$

Such equation enables to distinguish between the total effect of aggregate private knowledge and the total spillover effect. Since all private knowledge is supposed to spill over to some extent, the total effect of all private knowledge at the aggregate level is given by $\gamma+\mu$ (Griliches, 1979 and 1992).

On the basis of the argument elaborated so far, we may provide some insights about the possible theoretical underpinnings to the concept of knowledge capital stock. Indeed, we lack an explicit theoretical reasoning on technological knowledge leading to its operationalization in terms of knowledge capital stock. A quote from Griliches (1967) may be of some help here:

“For example, let investments affect the level of patenting with a lag whose generating function is given by $W_1(z)$, let these new inventions be embodied in new investment with a lag $W_2(z)$ and let new investment affect total factor productivity with a lag $W_3(z)$; then the total lag distribution of productivity behind investment is given by $W(T) = W_1(z)W_2(z)W_3(z)$ ” (Griliches, 1967: p. 20).

It is clear that the application of lag generating functions to investments measures so as to get a stock implies an underlying sequential process that start with R&D investments to yield a proxy of cumulated knowledge that in turn is supposed to show some effects on economic performances. In this direction, we believe it would not be that inappropriate saying that knowledge capital stock implies a vision of knowledge accumulation as an outcome of a linear process like this one: science precedes technology development, which then comes to be adopted by firms, and finally affects production efficiency.

After all, Vannevar Bush’s report to the US president had long been the main reference text to students of science and technology. Therefore it’s likely that the articulation of the linear model he proposed has influenced the way scholars from other fields looked at technological knowledge as well. Moreover, Kline and Rosenberg’s critique came only in the 1980s, and so did many of the works that opened up a new view on

knowledge and innovation providing the basis to the knowledge production function approach (Bush, 1945; Kline and Rosenberg, 1986; Balconi et al., 2009)¹.

2.2 Knowledge production function

The inclusion of knowledge capital stock within an extended production function approach allows economists to preserve the basic microeconomic assumptions about production sets out of which firms take their profit-maximizing choice. However, such approach assumes the existence of a separate R&D sector that is partly responsible of the change in the production technology, and hence of the shift of the production function (Nelson, 1980).

Because of this limitation, such a representation begun to be challenged mainly by evolutionary economists, who proposed to expand the view upon technological knowledge so as to account for it inherent compositeness. At the same time, scholars of science and technology started criticizing the linear model, by proposing an alternative view basically drawing upon systemic models of innovation based upon the interaction among different and yet complementary institutions involved in the complex business of knowledge production (Kline and Rosenberg, 1986; Gibbons et al. 1992).

A couple of Dick Nelson's contributions in the early 1980s provided a clear statement of the problems with the concept of knowledge capital stock, along with the theorization of a more articulated concept of knowledge, understood as a set of capabilities guiding the search processes undertaken by organizations performing R&D. Such capabilities may be themselves the outcome of R&D activities, and are likely to improve over time due to dynamic increasing returns stemming from learning by doing dynamics (Nelson, 1980 and 1982).

In this sense, such contributions may be viewed as pioneering in the attempt of opening the black box of technological knowledge so as to explicitly improve upon Griliches' and Mansfield's former operationalizations. Moreover, they also proposed a more realistic view in which science and technology are far from being sharply differentiated. There are a number of institutions producing knowledge, some of them are public while some others are private, and it is not possible to identify a one to one mapping from science to public institutions or from applied technology to private business firms. Scholars must acknowledge that different kinds of organizations take part in the process of knowledge production, like firms, research labs and universities (Nelson, 1982 and 1986).

This set of arguments has been well received mostly in the literature dealing with knowledge production at the aggregate level. In particular the literature on regional systems of innovation provided a fertile ground to develop the implications of this new

¹ We do not intend to go into the debate on the virtues and drawbacks of the linear model. The work by Balconi et al. (2009) provides an excellent synthesis in this direction.

view (Cooke, 1996; Cooke et al., 1997). Regional economists translated the idea that knowledge is the result of the interaction of a number of complementary inputs provided by different research institutions, into the concept of knowledge production function. The differences with the concept of knowledge capital stock are clear. Knowledge is no longer the mere result of cumulated R&D spending subject to decreasing returns. The knowledge production function provides a mapping from knowledge inputs to knowledge outputs that appears as follows:

$$\log(K_t) = \alpha + \beta \log(R_t) + \gamma \log(U_t) + \delta \log(Z_t) + \varepsilon \quad (3)$$

Where K stands for a measure of knowledge output, say patents, R stands for the industry R&D and U represents the university research, while Z includes a proxy for the concentration of a given type of activity (Acs et al., 2002; Fritsch, 2002). Equation (3) represents a production function, the arguments of which enter a multiplicative relationship, and hence are seen as complementary rather than substitute. The coefficients are in turn the elasticities of knowledge output to knowledge inputs.

On a fairly similar ground, the localized technological change approach has stressed that the dynamics of knowledge production are characterized by the joint utilization of internal and external knowledge, both tacit and codified. Mechanisms of learning, socialization and recombination are considered as crucial in a context characterized by the production of knowledge by means of knowledge itself (Antonelli, 1999).

The knowledge production function approach represents an improvement both from the theoretical and the empirical viewpoint, with respect to the concept of knowledge capital stock. It allows to gaining a better understanding of the interactive dynamics leading to the production of technological knowledge, by accounting for possible dynamic increasing returns stemming from learning dynamics as well as knowledge externalities. However, knowledge on the left hand side of the equation still is conceived as an homogeneous stock, and little is said about the intrinsic heterogeneity of knowledge base. In other words such representation still lacks proper cognitive models of knowledge production.

2.3 Complex knowledge and NK models

The development of the knowledge production approach inevitably leaves with a basic question as to what are the micro-founded mechanisms underlying knowledge production. In this respect, the interest in the cognitive mechanisms leading to production of new technological knowledge has recently emerged in the field of economics of innovation. This strand of analysis has moved from key concepts brought forward by Schumpeter (1912 and 1942) and Usher (1954), and then elaborated upon the models proposed within evolutionary economics (Nelson and Winter, 1982).

In his seminal works, Schumpeter proposed to view innovation as the outcome of a recombination process. Most of innovations brought about in the economic system stem from the combinations of existing elements in new and previously untried ways. Such

innovations appear to be mainly as incremental. Radical innovations stem instead from the combination of existing components with brand new ones.

The contributions by Weitzman (1996 and 1998) represent the former, and very impressive, attempt to draw upon such assumptions. His recombinant growth approach provides a sophisticated analytical framework grafting a micro-founded theory of knowledge production within an endogenous growth model. The production of knowledge is seen as the outcome of an intentional effort aimed at reconfiguring existing within a genuine cumulative perspective. However, there is no particular focus on the constraints that the combination of different ideas may represent, especially when these ideas are technologically distant. The only limiting factor seems to be the bounded processing capacity of economic agents.

The recombinant knowledge approach is based on the following assumptions. The creation of new knowledge is represented as a search process across a set of alternative components that can be combined one another. However, within this framework a crucial role is played by the cognitive mechanisms underlying the search process aimed at exploring the knowledge space so as to identify the pieces that might possibly be combined together. The set of potentially combinable pieces turns out to be a subset of the whole knowledge space. Search is supposed to be local rather than global, while the degree of localness appears to be the outcome of cognitive, social and technological influences. The ability to engage in a search process within spaces that are distant from the original starting point is likely to generate breakthroughs stemming from the combination of brand new components (Nightingale, 1998; Fleming, 2001).

Incidentally, such an approach also enables to better qualify the distinction between exploration and exploitation formerly articulated by March (1991). Most of the research in organization studies has usually seen search processes as ranging between two poles of a one-dimensional continuum, i.e. exploration and exploitation. The view of knowledge as an outcome of a recombination activity allows the introduction of two nested dimensions, defined according to degree to which agents decide to rely either on exploration or exploitation or on a combination of both. To this purpose concepts like search depth and search scope have been introduced. The former refers to degree to which agents intend to draw upon their prior knowledge, while the latter refers to the degree to which agent intend to rely on the exploration of new areas in the knowledge space (Katila and Ahuja, 2002).

Recombination occurs only after agents have put much effort in searching within the knowledge space. This strand of literature posits that knowledge so obtained is complex, meaning that it comprises many elements that interact richly (Simon, 1966; Kauffman, 1993). This has paved to way to an increasing number of empirical works based on the NK model proposed by Kauffman, according to which the search process is conducted across a rugged landscape, where pieces of knowledge are located and which provides the context within which technologies interact.

The bulk of the focus is on the concept of interdependence among the pieces that are combined together, while complexity is defined as the relationship between the number of components and the degree of interdependence (Fleming and Sorenson, 2001;

Sorenson et al., 2006). Following the intuition on the importance of patent citations contained in the seminal paper by Manuel Trajtenberg (1990), the empirical implementation of the interdependence concept is based on the deployment of the information contained in patent documents, i.e. technological classes and citations to other patents. In particular, interdependence is considered as a powerful explanatory variable building upon the technological classes the patent is assigned to. The interdependence of a patent l is obtained in two steps. First of all one has to calculate the ease of recombination for each subclass i (E_i), defined as the count of subclasses $j \neq i$ previously combined with class i weighted by total number of patents assigned to class i :

$$E_i = \frac{\sum_{j \neq i} l_j}{\sum l_i} \quad (4)$$

Then one can calculate the degree of interdependence of patent l (K_l) by inverting its average ease of recombination:

$$K_l = \frac{\sum_{i \in l} E_i}{\sum_{i \in l} l_i} \quad (5)$$

This empirical approach allows for evaluating the relative probability of recombination of each technological class observed in the patent sample, and then to assign an average recombination score to a patent. The basic idea is that the more combinable are the classes contained within a patent, the lower the degree of interdependence, as the technology is susceptible to be developed in a larger number of directions. On the contrary, should the classes be hardly combinable, then a relatively low number of possible combinations is possible, for which the technology turns out to show a high degree of interdependence. Such measure of interdependence is in turn expected to explain differentials in usefulness of inventions as proxied by the flow of citations received by patents over time.

Such framework clearly has the merit to push the economic discussion about technological knowledge beyond the conventional vision considering it as a sort of black box. It sheds light on the possibility to further qualify knowledge as proxied by patents, by better exploiting the information contained in patent documents. Moreover, it provides a former and innovative link between knowledge and complexity.

However, the notion of complexity used therein seems to be constrained to a generic definition of an object the elements of which are characterized by a high degree of interaction. As an implication the empirical effort does not go beyond the count of classes and of patents assigned to classes. The NK models fail to identify knowledge as an emergent property of an adaptive complex system, characterized by an architecture that can influence the actions at the micro and meso levels as well as be influenced as a result of what happens at lower layers. This requires first to make it explicit a concept of knowledge structure and then to explore the different tools made available by different methodological approaches.

Summing up, the grafting of complexity theory into economic sciences has proved to be particularly fertile, especially for what concerns the economics of knowledge and innovation. The explicit reference to the NK-model by the recombinant knowledge literature provides a clear example in this respect.

Most NK-models are however affected by a severe limit, which constrains their usefulness. The complex system is characterized by a set of elements and the connections amongst them. The configuration of the linkages connecting the elements of the system is likely to affect agents' performances. The main problem here is that the architecture of the system is often considered as table over time rather than evolving (Frenken, 2006). This amounts to consider the degree of complexity of the system as exogenous, defined *ex ante*. The contribution by Fleming and Sorenson discussed above presents exactly this limitation, which makes it unsuitable to the analysis of the evolutionary and path dependent dynamics of technological change.

3 Endogenous Complexity and Technological Knowledge

The main issue to be considered now is that the architecture of a complex system may well change over time, and so may the structure of *epistatic relationships*. This may occur either due to a change in the relative weight of some elements in the system, these elements switching from a non-influential to an influential position, or by means of introduction of new elements within the system. This is in turn likely to alter the existing structure of relationships. Within this context, the *pleiotropy* represents the number of elements in the system that are affected by the appearance of new elements. It is clear that the higher the pleiotropy, the greater the change in the architecture of the system that the inclusion of new elements may engender.

The model of constructional selection by Altenberg (1994 and 1995) represents one of the few attempts to cope with the issue of changing architectures of complex systems. As noted by Frenken (2005 and 2006), such class of models is well suited to investigate the evolution of technologies considered as artefacts made of interdependent elements (Lane and Maxfield, 2005).

The viewpoint of endogenous complexity makes the analysis of knowledge dynamics particularly appealing and challenging. *Knowledge* can indeed be represented as an *emergent property stemming from multi-layered complex dynamics* (see Figure 1). Knowledge is indeed the result of a collective effort of individuals who interact with one another, sharing their bits of knowledge by means of intentional acts of communication (Antonelli, 2008; Saviotti, 2007). In other words, the adoption of an endogenous complexity made possible by the recombination approach allows for the combination of the view on technology as an artefact with the view of technology as an act, i.e. as the product of collective actions involving agents with converging incentives and aligned interests (Arthur, 2009; Lane et al., 2009).

INSERT FIGURE 1 ABOUT HERE

The structure of the network of relationships amongst innovating agents represents therefore a crucial factor able to shape the ultimate outcome of knowledge production processes. Constructional selection matters, in that new institutions entering the network need first of all to choose with which incumbents they want to be linked with. The concept of preferential attachment applies to this situation. In a wide number of contexts, the new nodes in a network generally end up to link with those 'old' nodes already characterized by a large number of connections (Barabasi and Albert, 1999). As a consequence, the entrance of new actors in the network is likely to reshape the relative weight of nodes, and hence modify the structure and the balance of relationships.

Collective knowledge so produced stems from the combination of bits of knowledge dispersed among innovating agents. Creativity refers to the ability of agents to combining together these small bits of knowledge so as to produce an original piece of technological knowledge. This in turn may be thought about as a collection of bits of knowledge linked one another. The knowledge base of a firm can be therefore imagined as a network in which the nodes are the small bits of knowledge and the links represent their actual combination in specific tokens. Knowledge in this sense turns out to be an emergent property of complex dynamics featuring the interdependent elements of the system, i.e. the bits of knowledge.

This is a quite unexplored consequence of the collective character of knowledge production, which provides further richness to its dynamics. Such complex system may be represented as network the nodes of which are the smaller units of knowledge while the edges stand for their actual combination. Hence the knowledge base is characterized by a structure with its own architecture. This in turn may evolve over time, as an effect of the introduction of new small bits of knowledge and the consequent change in the relative weight of the nodes within the network. Indeed, like in the networks of innovators, new nodes will be attached to some existing nodes, the centrality of which will be altered. Learning dynamics and absorptive capacity represent a channel through which the topology of knowledge structure affects search behaviour at the level of agents networks. Indeed, agents move across the technology landscape in regions that are quite close to the area of their actual competences. Technological change is localized as an effect of the interactions between the complex dynamics at the knowledge and the agents' level. However the topology of knowledge structure is in turn shaped by the choices made by innovating agents as to which bits of knowledge combine together. A self-sustained process is likely to emerge, according to which the knowledge creation process tends more and more towards a local attractor in which they are locked in (see the chapter by Colombelli and von Tunzelmann in this book).

This dynamics indeed makes preferential attachment work also at the knowledge level. Agents' search behaviour is indeed constrained by the topology of the knowledge structure. In this direction, those small bits of knowledge which have grown in importance are likely to exert a much stronger influence. This process is rooted in historical time, according to which the gradual sorting out of knowledge bits which have proved not to be so fertile, leaves the floor to few and more fertile bits. New bits of

knowledge entering the knowledge base later on are likely to be linked to these few pillars.

Preferential attachment introduces a great deal of path dependence in system dynamics of technological knowledge. It amounts to articulate the concept of persistence beyond the rate of introduction of innovations, so as to apply it to the centrality of the specific smaller bits of knowledge which make the structure of the knowledge base.

Still, while this self-enforcing process is likely to trap the search process within a bounded area, the dynamics of technological communication at the agents' level as well as the capabilities to cope with search in areas that are far away from the competences of innovating agents are likely to introduce discontinuities in the evolutionary pattern. This amounts to introduce a wide variety of new bits of knowledge which are loosely related with those already existing in the knowledge base, so as to give rise to radically new combinations. The process of evolution, fed by learning dynamics and cumulativeness, leads to the gradual selection of the best combinations, which grow in centrality and hence begin to constrain agents' search behaviour.

Knowledge sharing and technological communication ensure therefore the emergence of new variety, which is more likely to occur in transition phases. At this stage a wide range of alternatives are viable, and multiple local attractors are likely to emerge from mutual influences between complex dynamics at the knowledge and the agents' layers.

4 Some Operational Methodologies

The outcome of considering endogenous complexity in technological knowledge is that the layout of knowledge structure appears to be both an outcome and a determinant of agents' search. This deserves further careful attention and more in depth analysis. In what follows we propose two alternatives methodologies which have been recently introduced, and are equally suitable to improve our empirical ability to measure the various facets of the evolution of the knowledge base, including the occurrence of path dependency and persistence phases as well as the emergence of variety and discontinuity phases.

4.1 Measures based on co-occurrence matrixes

The purpose of this first methodology consists in the exploration of the evolution of the properties of the knowledge base, with particular emphasis on the issues of variety, similarity and complementarity.

- 1) Variety can be measured by using the information entropy index. Entropy measures the degree of disorder or randomness of the system, so that systems characterized by high entropy is also be characterized by a high degree of uncertainty (Saviotti, 1988).

The information entropy has some interesting properties, and especially a property of multidimensional extension (Frenken and Nuvolari, 2004). Consider a pair of events (X_i, Y_j) , and the probability of co-occurrence of both of them p_{ij} . A two dimensional total variety (TV) measure can be expressed as follows:

$$TV \equiv H(X, Y) = \sum_l \sum_j p_{lj} \log_2 \left(\frac{1}{p_{lj}} \right) \quad (6)$$

If one considers p_{lj} to be the probability that two technological classes l and j co-occur within the same patent, then the measure of multidimensional entropy focuses on the variety of co-occurrences of technological classes within regional patents applications.

Moreover, the total index can be decomposed in a “within” and a “between” part anytime the events to be investigated can be aggregated in a smaller numbers of subsets. Within-entropy measures the average degree of disorder or variety within the subsets, while between-entropy focuses on the subsets measuring the variety across them. Frenken et al. (2007) refer to between- and within- group entropy respectively as unrelated and related variety.

It can be easily shown that the decomposition theorem holds also for the multidimensional case. Hence if one allows $l \in S_g$ and $j \in S_z$ ($g = 1, \dots, G; z = 1, \dots, Z$), we can rewrite $H(X, Y)$ as follows:

$$TV = H_Q + \sum_{g=1}^G \sum_{z=1}^Z P_{gz} H_{gz} \quad (7)$$

Where the first term of the right-hand-side is the between-entropy and the second term is the (weighted) within-entropy. In particular:

$$UTV = H_Q = - \sum_{g=1}^G \sum_{z=1}^Z P_{gz} \log_2 P_{gz} \quad (8)$$

$$RTV = \sum_{g=1}^G \sum_{z=1}^Z P_{gz} H_{gz} \quad (9)$$

$$P_{gz} = \sum_{l \in S_g} \sum_{j \in S_z} p_{lj}$$

$$H_{gz} = - \sum_{l \in S_g} \sum_{j \in S_z} p_{lj} \log_2 \left(\frac{1}{p_{lj}} \right)$$

We can therefore refer to between- and within-entropy respectively as *unrelated technological variety (UTV)* and *related technological variety (RTV)*, while total information entropy is referred to as *general technological variety*.

- 2) The similarity amongst different types of knowledge can be captured by a measure of cognitive distance. A useful index of distance can be derived from the measure of *technological proximity* originally proposed by Jaffe (1986 and 1989), who investigated the proximity of firms' technological portfolios. Subsequently Breschi et al. (2003) adapted the index in order to measure the proximity, or relatedness, between two technologies. The idea is that each firm is characterized by a vector V of the k technologies that occur in its patents. Knowledge similarity can first be calculated for a pair of technologies l and j as the angular separation or uncentred correlation of the vectors V_{lk} and V_{jk} . The similarity of technologies l and j can then be defined as follows:

$$S_{lj} = \frac{\sum_k V_{lk} V_{jk}}{\sqrt{\sum_k V_{lk}^2} \sqrt{\sum_k V_{jk}^2}} \quad (10)$$

The idea underlying the calculation of this index is that two technologies j and l are similar to the extent that they co-occur with a third technology k . The cognitive distance between j and l is the complement of their index of the similarity:

$$d_{lj} = 1 - S_{lj} \quad (11)$$

Once the index is calculated for all possible pairs, it needs to be aggregated at the industry level to obtain a synthetic index of technological distance. This can be done in two steps. First of all one can compute the weighted average distance of technology l , i.e. the average distance of l from all other technologies.

$$WAD_l = \frac{\sum_{j \neq l} d_{lj} P_{jt}}{\sum_{j \neq l} P_{jt}} \quad (12)$$

Where P_j is the number of patents in which the technology j is observed. Now the average cognitive distance at time t is obtained as follows:

$$CD_t = \frac{\sum_l WAD_l P_{lt}}{\sum_l P_{lt}} \quad (13)$$

Complementarity: typically a firm needs to combine, or integrate, many different pieces of knowledge to produce a marketable output. In order to be competitive a firm not only needs to learn new 'external' knowledge. It also needs to learn how to combine it with other, new and old, pieces of knowledge. We can say that a knowledge base, in which different pieces of knowledge are well combined, or integrated, is a coherent knowledge base. Such technologies are by definition complementary in that they are jointly required to obtain a given outcome. We can now turn to calculate the *coherence* (R) of the knowledge base, defined as the average relatedness of any technology randomly chosen within the sector with respect to any other technology (Nesta and Saviotti, 2005 and 2006; Nesta, 2008). To yield the knowledge coherence index, a number of steps are required however. In what follows we describe how to obtain the index at the sector

level. First of all, one should calculate the weighted average relatedness WAR_l of technology l with respect to all other technologies present within the sector. Such a measure builds upon the measure of technological relatedness τ_{lj} (see Nesta and Saviotti, 2005). Following Teece et al. (1994), WAR_l is defined as the degree to which technology l is related to all other technologies $j \in l$ in the sector, weighted by patent count P_{jt} :

$$WAR_l = \frac{\sum_{j \neq l} \tau_{lj} P_{jt}}{\sum_{j \neq l} P_{jt}} \quad (14)$$

Finally the coherence of knowledge base within the sector is defined as weighted average of the WAR_{lt} measure:

$$R = \frac{\sum_{l,t} WAR_{lt} P_{lt}}{\sum_{l,t} P_{lt}} \quad (15)$$

It is worth stressing that such index implemented by analysing co-occurrences of technological classes within patent applications, measures the degree to which the services rendered by the co-occurring technologies are complementary one another. The relatedness measure τ_{lj} indicates indeed that the utilization of technology l implies that of technology j in order to perform specific functions that are not reducible to their independent use.

4.2 Measures based on social network analysis

The starting point of this second methodology is to consider that a *network* may be defined as a graph made of nodes that are tied each other by one or more types of interdependency. Relationships among nodes are expressed by arcs, which in turn may be directed or undirected. Two nodes that are connected by a line are said to be *adjacent* to one another. Adjacency is therefore the graphical expression of the fact that two nodes are directly related or connected to one another. The points to which a particular point is adjacent are termed its *neighbourhood*. Points may be directly connected by a line, or they may be indirectly connected through a sequence of lines. It may be thought as a 'walk' in which each point and each line are distinct. This is called *path*. The *length* of path is measured by the number of lines that constitute it. The *distance* between two points is the shortest path (the geodesic) that connects them.

One of the most widely used measures to describe a network is the *density*. It describes the general level of linkage among the points in a graph. The density of a network is therefore defined as the total number of actual lines, expressed as a proportion of the maximum possible number of lines:

$$\Delta = \frac{l}{n(n-1)/2} \quad (16)$$

A network is complete when all the nodes are adjacent, and the measure of density attempts to summarize the overall distribution of lines in order to assess how far the network is from completion. Density depends upon two other important parameters of the network, i.e. the inclusiveness and the sum of the degree of its points. *Inclusiveness* can be defined as the share of network nodes that are not isolated, i.e. the share of nodes that are connected to at least another node. For example, in a network of 20 nodes with 5 isolated nodes the inclusiveness is 0.75. The more inclusive the graph, the more dense the network will be.

However some nodes will be more connected than other ones. The *degree* of a node is an important measure of centrality that refers to the total number of other points in its neighbourhood. Formally one can represent the degree by the following equation:

$$D(v) = \sum_{s \in V, s \neq v} x_{v,s} \quad (17)$$

This measure is obviously biased by the network size. Therefore it is useful to use a standardized measure, which consists in dividing the degree measure by its maximum value as follows:

$$NI(v) = \frac{D(v)}{n-1} \quad (18)$$

The higher the degree of the connected points in the network, the higher will be the density. For this reason the calculation of density needs to take into account both measures. It should compare the actual number lines present in the graph with the total number of lines that the graph would show if it were complete.

While the density describes the network as a whole, the measures of *centrality* refer to the relevance of the nodes belonging to the network. A point is locally central if it has a large number of connections with other points in its immediate environments, i.e. other points in its neighbourhood. Global centrality refers instead to the prominence of the node with respect to the overall structure of the network. Measures of global and of local centrality have a different meaning.

Measures of global centrality are expressed in terms of the distance among various points. Two of these measures, i.e. closeness and betweenness, are particularly important. The simplest notion of *closeness* is that calculated from the ‘sum distance’, the sum of geodesic distances to all other points in the graph (Sabidussi, 1966). After having calculated the matrix of distances among the nodes of the network, the sum distance is the row or column marginal value. A point with a low sum distance is close to a large number of other points, and so closeness can be seen as the reciprocal of the sum distance. Formally it can be expressed as follows:

$$C(v) = \frac{1}{\sum_{s \in V, s \neq v} d(v,s)} \quad (19)$$

Where the denominator represents the sum of the geodesic distance of the vertex v to all other points.

The *betweenness* measures the extent to which a particular point lies ‘between’ the other points in the graph: a point with a relatively low degree may play an important intermediary role and so be very central to the network (Freeman, 1979). The betweenness of a node measures how much it can play the part of a broker or gatekeeper in the network. Freeman’s approach is built upon the concept of local dependency. A point is dependent upon another if the paths which connect it to the other points pass through this point. Formally, let G be a graph with n vertices, then the betweenness is calculated as follows:

$$B(v) = \sum_{\substack{s \neq v \\ s \neq t}} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (20)$$

Where σ_{st} is the number of shortest geodesic paths from s to t , and $\sigma_{st}(v)$ is the number of shortest geodesic paths from s to t passing through a vertex v .

The centrality measures discussed above, allow us to characterize each single network node. However, it is also possible to calculate the sector averages for all of the three indexes. In this direction, one must consider that each node corresponds to a technological class observed with a specific relative frequency, which must be taken into account when averaging out the centrality measures. We can then propose weighted average centrality measures as follows. Let $Z(v)$ be one of the three centrality measures referred to the generic node v , the weighted average centrality at time t is:

$$\bar{Z} = Z \times \frac{P_v}{\sum P} \quad (21)$$

Where P_v is the number of patents in which the technology v is observed.

5 Conclusion: Avenues for Future Research

The chapter was intended to provide an original and creative review of the literature on the dynamics of technological knowledge. Table 1 provides a synthesis and a taxonomy of the different approaches to technological knowledge, as well as of their theoretical underpinnings and empirical consequences.

INSERT TABLE 1 ABOUT HERE

We argue that among the new developments on the theme, the investigation on endogenous complexity in technological knowledge is certainly the most promising advance. First it provides an accurate representation of how knowledge is created and diffused at the analytical level, and second it also benefits of an empirical value since it can be expressed by a wide range of indicators and measures. In particular, we claimed

that such framework has a great potential in that it provides both the theoretical and empirical grounds to carry out an interdependent analysis of technology as an act and as an artefact. In this direction, the structure of technological knowledge is represented as a network the architecture of which is in turn influenced by the architecture of the network of innovation, and vice versa. It follows a never ending process of mutual influences that keeps the system constantly out of equilibrium (see the chapter by Antonelli in this book).

The notion of coalitions for innovation gains momentum in this context (David and Keely, 2002). They can be regarded as the product of spontaneous order, yet their emergence can be guided and designed by means of the intentional intervention of policy makers as well as corporate strategies. The successful introduction of an innovation may be regarded as the result of a hegemonic coalition, that is a coalition that has been able to design a group of complementary agents, coordinate their incentives and integrate their competences so as to achieve hegemony in a given technological space. The design of coalitions for innovation is therefore likely to exert a great deal of influence on the direction of technology evolution, and hence on future developments of the knowledge space. Within non-ergodic systems, this is likely to favour the lock-in engendered by path dependent dynamics, unless the structures of the two nested networks change so much that a new hegemonic coalition emerge able to introduce a discontinuity in the technology space.

The implications of such approach are far reaching. One of the major domains of application so far has been the analysis of the technological basis of knowledge of firms, characterized by patent portfolios (see Nesta, 2008). Further applications have been recently proposed in empirical studies dealing with the evolutionary patterns of development of knowledge intensive sectors, especially focused in the identification of the introduction of discontinuities and the periodicity of random screening and organized search stages (Krafft, Quatraro and Saviotti, 2009 and 2011; Antonelli, Krafft and Quatraro, 2010).

A non exhaustive list of potential applications can be elaborated, and each element in this list can be considered as a major avenue of research to be explored in the future:

- industrial dynamics and evolution: the fact that, in an industry, knowledge can either come from a recombination of existing knowledge or from the creation of new knowledge has an impact on industrial evolution. Incumbents may play the role of efficient recombination of existing knowledge, but very often may also rely on new entrant firms on the creation of new knowledge. Depending on the share of combination of existing knowledge versus creation of entirely new knowledge, incumbents or new entrants may act as leaders in the industry.

- networks: in most industries, networks occur among firms, and appear more and more as a stable form of industrial organization. Endogenous complex knowledge allows an accurate mapping of the formation of networks, and their transformation over time. Moreover depending on preferential attachment characteristics of the agents within the network, it is possible to identify the centrality of some actors in the network at some

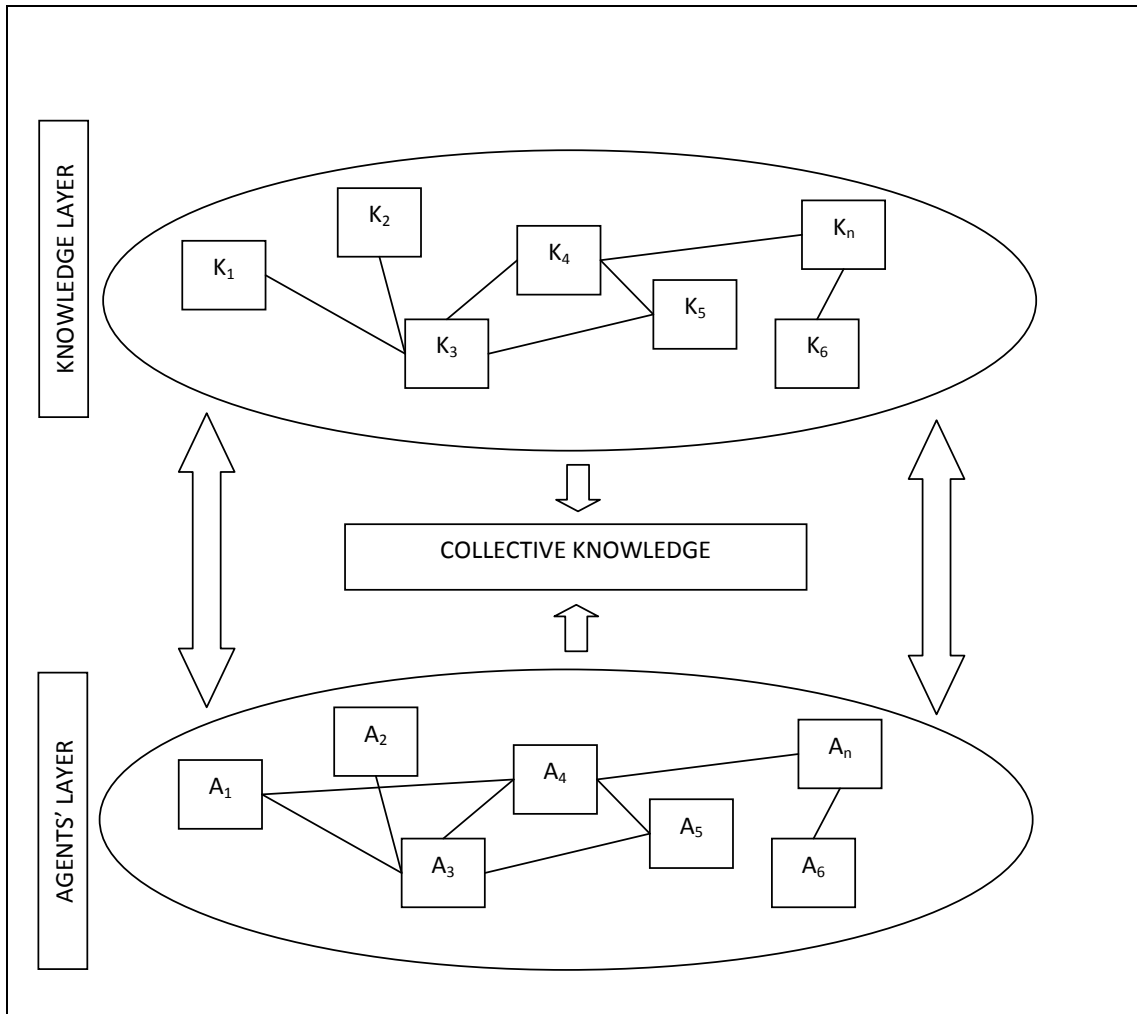
point in time, and to predict how it may change over time with the entrance of new actors.

- geographical issues: the recent debates on knowledge cities, or the more traditional ones on learning regions, also can have a new echo based on the use of the analysis of complex knowledge. On this theme, the approach can provide new quantitative results on the importance of geography in the creation and recombination of knowledge. Especially it is possible to assess quantitatively how new actors bringing new pieces of knowledge may aggregate other actors already installed or not, and eventually how these new actors may gain over time some weight (or centrality) over older ones, shaping thus the technological characteristics of a region.

Table 1 – Taxonomy of the different approaches to technological knowledge

Early 1980s	Extended production function	Knowledge capital stock as explanatory variable	Homogeneous good	Linear model	Linear mode and top down process R&D and specialized institutions of knowledge Large, vertically related companies Internal financial markets
Late 1980s Early 1990s	Knowledge production function	Knowledge capital stock as dependent variable	Homogeneous good	Systemic interactions	Learning effects and bottom up process Markets for knowledge and strong IPR regimes Small, specialized firms Venture capital and IPOs
Late 1990s Early 2000s	Exogeneous complexity	Citations and ease of recombination	Heterogeneous good	Emergent property of a given architecture	Search conducted across a rugged landscape Explain differentials in usefulness of inventions
Early 2000s	Endogeneous complexit	Technological classes and knowledge structure	Heterogeneous good	Emergent property of a changing architecture	Knowledge discontinuities and search strategies Stable innovation networks with large and small firms retain and reinvest and long term investors strategies

Figure 1 - Multi-layered complex dynamics of knowledge



6 References

- Acs, Z.J., Anselin, L. and Varga, A., 2002, Patents and innovation counts as measures of regional production of new knowledge, *Research Policy*, 31, 1069-1085.
- Altenberg, L., 1994, Evolving better representations through selective genome growth. In Schaffer, J.D., Schwefel, H.P. and Kitano H. (eds.) *Proceedings of the IEEE World Congress on Computational Intelligence*. Piscataway, NJ: IEEE, pp. 182–187.
- Altenberg, L., 1995, Genome growth and the evolution of the genotype–phenotype map. In Banzhaf W. and Eckman F.H. (ed.) *Evolution and Biocomputation*. Berlin & Heidelberg: Springer-Verlag, pp. 205–259.
- Antonelli, C. 2008, *Localized technological change: Towards the economics of complexity*, London, Routledge.
- Antonelli, C., 1999, *The microdynamics of technological change*, London, Routledge.
- Antonelli, C., Krafft, J., Quatraro, F., 2010, Recombinant knowledge and growth: The case of ICTs, *Structural Change and Economic Dynamics*, 21, 50-69.
- Balconi, M., Brusoni, S. and Orsenigo, L., 2009, In defense of the linear model: An essay, *Research Policy*, in press.
- Barabasi, A.L. and Albert, R., 1999, Emergence of scaling in random networks, *Science*, 286, 509-512.
- Breschi, S., Lissoni, F., and Malerba, F., 2003, Knowledge relatedness in firm technological diversification, *Research Policy*, 32, 69-97.
- Bush, V., 1945, *Science the endless frontier. Report to the president*, Washington DC, United States Government Printing Office.
- Cooke, P., M. Uranga and G. Etzebarria, 1997, Regional innovation systems: institutional and organisational dimensions, *Research Policy* (1997), 26, 475-491
- Cooke, P., 1996, The New Wave of Regional Innovation Networks: Analysis, Characteristics and Strategy, *Small Business Economics*, 8:1-13 (1996).
- David, P.A. and Keely, L., 2003, The endogenous formation of scientific research coalitions, *Economics of Innovation and New Technology* 12, 93–116.
- Fleming, L. and Sorenson, O., 2001, Technology as a complex adaptive system: Evidence from patent data, *Research Policy*, 30, 1019-1039.
- Fleming, L., 2001, Recombinant uncertainty in technological search, *Management Science*, 47, 117-132.
- Freeman, L. C., 1979, Centrality in social networks: Conceptual clarification, *Social Networks*, 1(3), 215-239.
- Frenken, K., 2005, *Innovation, Evolution and Complexity Theory*. Cheltenham: Edward Elgar,
- Frenken, K., 2006, Technological innovation and complexity theory, *Economics of Innovation and New Technology* 15, 137-155.
- Frenken, K., Nuvolari, A., 2004, [Entropy Statistics as a Framework to Analyse Technological Evolution](#), in J. Foster and W. Hözl (eds), *Applied Evolutionary Economics and Complex Systems*. Edward Elgar: Cheltenham, U.K. and Northampton, Mass.
- Frenken, K., von Oort, F., Verburg, T., 2007, [Related Variety, Unrelated Variety and Regional Economic Growth](#), *Regional Studies*, 41(5), 685-97.

- Fritsch, M., 2002, Measuring the quality of regional innovation systems: A knowledge production function approach, *International Review of Regional Science*, 25, 86-101.
- Gibbons, M., Limoges, C., Nowtny, H, Schwartzman, S., Scott, P. and Trow, M., 1984, *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*, London, Sage.
- Griliches, Z., 1967, Distributed Lags: A Survey, *Econometrica*, 35, 16-49.
- Griliches, Z., 1957, Hybrid corn: An exploration in the economics of technological change. *Econometrica*, 25(4), 501–522; reprinted in Fox, K. A. and Johnson, D. G. (Eds.) (1969). *AEA Readings in Agricultural Economics*,
- Griliches, Z., 1979, Issues in Assessing the Contribution of Research and Development to Productivity Growth, *The Bell Journal of Economics*, 10 (1), 92-116.
- Griliches, Z., 1992, The search for R&D spillovers. *The Scandinavian Journal of Economics*, 94(Supplement),
- Griliches, Z., 1980, R&D and the Productivity Slowdown, *American Economic Review*, 70, 343-348.
- Jaffe, A., 1986, [Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value](#), *American Economic Review*, 76(5), 984-1001.
- Jaffe, A., 1989, [Real Effects of Academic Research](#), *American Economic Review*, 79(5), 957-70.
- Katila, R. and Ahuja, G., 2002, Something old, something new: A longitudinal study of search behaviour and new product introduction, *Academy of Management Journal*, 45, 1183-1194.
- Krafft J., Quatraro F. and Saviotti P.P., 2011, The Evolution of Knowledge Base in Knowledge-Intensive Sectors: Social Network Analysis of Biotechnology, *Economics of Innovation and New Technology*, forthcoming, 20(2/3).
- Krafft, J., Quatraro F. and Saviotti P.P., 2009, Evolution of the Knowledge Base in Knowledge Intensive Sectors, revised and resubmitted, *Industrial and Corporate Change*.
- Kauffman, S., 1993, *Origins of order: Self-Organization and selection in evolution*, Oxford University Press, Oxford.
- Kleinknecht, A.H., van Montfort, K., Brouwer, E. 2002, The non-trivial choice between innovation indicators, *Economics of Innovation and New Technology* 11, 109 - 121.
- Kline, S.J. and Rosenberg, N., 1986, An overview on innovation, in Landau, R. and Rosenberg, N., 1986, *The Positive Sum Strategy*, Washington DC, National Academy Press.
- Lane, D., and Maxfield, R., 2005, Ontological uncertainty and innovation, *Journal of Evolutionary Economics*, 15(1), 3-50.
- Mansfield, E., 1961, Technical change and the rate of imitation. *Econometrica*, 29, 4 (1961), pp. 741–766.
- Mansfield, E., 1980, Basic research and productivity increase in manufacturing, *American Economic Review*, 70, 863-73.
- March, J., 1991, Exploration and exploitation in organizational learning, *Organization Science*, 2(1), 71-87.
- Nelson, R.R. and Winter, S., 1982, *An Evolutionary Theory of Economic Change*, Cambridge, Harvard University Press.

- Nelson, R.R., 1980, Production sets, Technological Knowledge, and R&D: Fragile and overworked constructs for the analysis of productivity growth, *American Economic Review*, 70, 62-67.
- Nelson, R.R., 1982, The role of knowledge in R&D efficiency, *Quarterly Journal of Economics*, 97, 453-470.
- Nelson, R.R., 1986, Institutions supporting technical advance in industry, *American Economic Review*, 76, 186-189.
- Nesta, L., 2008, Knowledge and productivity in the world's largest manufacturing corporations, *Journal of Economic Behavior and Organization*, 67, 886-902.
- Nesta, L., and Saviotti, P.P., 2005, [Coherence of the Knowledge Base and the Firm's Innovative Performance: Evidence from the U.S. Pharmaceutical Industry](#), *Journal of Industrial Economics*, 53(1), 123-42
- Nesta, L., and Saviotti, P.P., 2006, [Firm Knowledge and Market Value in Biotechnology](#), *Industrial and Corporate Change*, 15(4), 625-52.
- Nightingale, P., 1998, A cognitive model of innovation, *Research Policy*, 27, 689-709.
- Orsenigo, L., Pammolli, F. and Riccaboni, M., 2001, Technological change and network dynamics. Lessons from the pharmaceutical industry, *Research Policy*, 30, 485-508.
- pp. 221–246; and in Hooper, J. W. and Nerlove, M. (Eds.) (1970). Selected Readings in Econometrics. MIT Press, Cambridge, MA, pp. 231–253. Also reprinted in Wolff, E. (Ed.) (1997). The Economics of Productivity,
- Rosenberg, N., 1982, *Inside the Black Box: Technology and Economics*, Cambridge, Cambridge University Press.
- Sabidussi, G., 1966, The centrality index of a graph. *Psychometrika*, 31 (4), 581-603.
- Saviotti, P. P., 1988, Information, variety and entropy in technoeconomic development, *Research Policy*, 17(2), 89-103.
- Saviotti, P.P., 2007, On the dynamics of generation and utilisation of knowledge: The local character of knowledge, *Structural Change and Economic Dynamics*, 18, 387-408.
- Schumpeter, J., 1912, *The Theory of Economic Development*, Harvard University Press, Cambridge.
- Schumpeter, J., 1942, *Capitalism, Socialism and Democracy*, Harper and Row, New York.
- Simon, H., 1966, *The Sciences of the Artificial*. MIT Press, Cambridge, MA.
- Sorenson, O., Rivkin, J.W. and Fleeming, L., 2006, Complexity, network and knowledge flows, *Research Policy*, 35, 994-1017.
- Teece, D. et al., 1994, [Understanding Corporate Coherence: Theory and Evidence](#), *Journal of Economic Behavior and Organization*, 23(1), 1-30.
- Trajtenberg, M., 1990, A Penny for Your Quotes: Patent Citations and the Value of Innovations, *The Rand Journal of Economics*, 21, 172-187.
- Usher, A., 1954, *History of Mechanical Inventions*, Cambridge: Harvard University Press.
- Van der Bergh, J.C.J.M., 2008, Optimal diversity: Increasing returns versus recombinant innovation, *Journal of Economic Behaviour and Organization*, 68, 565-580.
- Vincenti, W., 1990, *What Engineers Know and How They Know It*, Baltimore and London: John Hopkins university Press.

- Weitzman, M. L., 1996, Hybridizing growth theory, *American Economic Review*, 86, 207-212.
- Weitzman, M. L., 1998, Recombinant growth, *Quarterly Journal of Economics*, 113, 331-360.