The Dynamics of Collective Invention*

Robin Cowan[a] Nicolas Jonard[a,b]

[a]MERIT, University of Maastricht, P.O. Box 616
6200 MD Maastricht, The Netherlands
[b]BETA, CNRS, 61 Ave. de la Forêt Noire
67000 Strasbourg, France

May 25, 2000

Abstract

This paper models the phenomenon of collective invention. Collective invention exists when the apparently irrational disclosure of information among competing entities creates a positive feedback that allows for high innovation rates and fast knowledge accumulation. Conditions for collective invention exist today in institutions such as the virtual communities (LINUX, or open source software groups e.g.) enabled by the Internet. We develop a formal model that accounts for the dynamics of knowledge and collective invention, and examine how the architecture of the network of agents influences patterns and rate of innovation. We find that the communication network structure has as strong influence on system performance. The small world structure stands out: the aggregate knowledge level grows fastest, but in addition the distribution of knowledge across agents is most unequal when compared to other network structures. Spatial correlation in knowledge levels exists and is affected by network architecture, but does not display small world properties however.

JEL Classification: L00, O31, O38, R10.
Keywords: Collective Invention, Knowledge, Networks, Small Worlds.

*We thank Jean-Benoît Zimmermann for many helpful comments.

1 Introduction

Early economists of innovation and technological change (see e.g. Nelson [17] and [18]; Nordhaus [19]) have identified three locations in which invention takes place: in non-profit institutions like universities; in profit-seeking firms that purposefully undertake costly exploration of their technological environment through research and development; and in the mind of individual inventors. Based on historical examination of the British blast furnaces industry in the nineteenth century, Allen [1] proposes a fourth location, namely in a collection of firms whose interactions would 'collective inventions'. In Allen's view there are two central features of collective invention: firms release to their competitors information about the design and efficiency of new plants or technologies; and individual firms devote few resources explicitly to the discovery of new knowledge. Thus the key to understanding an episode of collective invention is in the exchange and free circulation of knowledge and information within groups rather than in the inventive efforts of particular firms or individuals.

This paper models the phenomenon of collective invention. We examine an economy in which agents share knowledge, and with the knowledge received from each other improve their own knowledge levels. The motivating examples concern groups of firms operating in an era of rapid technical change and market growth, whose information-sharing activities lead to further rapid technical change. Since communication of knowledge is vital to this process, the network structure over which this communication takes place is an obvious area of concern. The focus here is explicitly on the relationship between the architecture of this network and the aggregate performance of the system. We begin by discussing several historical and modern examples, following which we present a formal model and the results of a dynamic simulation of it.

Cleveland (UK) was a site of collective invention in the mid-nineteenth century (Allen [1]). Owners of blast furnaces shared information, publishing it and presenting it in trade association meetings, regarding the technical and economic properties of their recent furnaces. This produced the discovery of a positive relationship between furnace height and production levels. Fast rates of innovation and rapid productivity growth in blast furnace operation were the result. A large number of other historical examples have also been documented. McGaw [15] shows that the mechanization of paper manufacture in the Berkshire (New England) area in the early 1800s is another case of collective invention. Mill owners formed a community by sharing business and technical information, and "... experienced paper makers transmitted a variety of information through their informal communications networks." (McGaw, p. 137). A major part of this information had to do with their own recent innovations, how they were contrived or acquired, and how they performed. The continuous mechanization taking place in this area both demanded and was built upon frequent oral, informal transmission of just this type of technical information. This again yielded large productivity gains. Lamoureaux [13] gives further examples of information sharing among inventors and innovators in the 18th and 19th century US.

While Allen conjectures that since 1900, and the rise of the industrial R&D laboratory, collective invention has become less important, modern examples do exist. The industry or trade association convention, which includes presentations of technical results of research is a form of broadcasting of information that comprises one part of collective invention. Von Hippel [26] documents knowledge trading among engineers in competing mini-mill firms in the US steel industry. Here innovation is not strictly collective under Allen's strong definition, in that firms do perform formal R&D, but the rapid and free distribution of knowledge is clearly an important input into the innovation process. Similar practices have been shown to exist in aerospace and wafer board manufacture (Von Hippel [26], p. 83). In the biotech industry managers acknowledge that significant amounts of information are acquired by their engineers through informal contacts with colleagues in other firms and institutions (Powell et al. [21], p. 120). The dramatic pace of innovation in Silicon Valley is frequently attributed to the rapid, relatively unrestrained, open diffusion of innovation (Saxenian [24] or Hyde [12]) and certainly the Italian industrial district is also well described as a case of collective invention (Russo [23]).[1]

There is today a particularly striking example in which collective invention is at work, generating tremendous amounts of knowledge and wealth. Here, discovering knowledge is relatively inexpensive and broadcasting information is the most common behaviour. This place is the Internet, and more specifically the world wide web. It contains many examples of information broadcast which diffuses and creates knowledge. A first observation here is that consumers now routinely go to the Internet for information which increases efficiency in the production of utility, for this activity demands not only physical goods but also knowledge.[2] Many types of knowledge that would be useful in this regard are broadcast via the Internet. Even such a simple example as the decision about which movie to see can be improved by accessing one of the many movie review pages on the WWW. These pages allow movie-goers to post their opinions about movies they have seen. Looking at these postings, which are unpaid provisions of useful information, allows one to make a better-informed choice about where to spend one's movie-going budget. It is true that other movie-goers are not competitors, but they are providing useful information at no cost to the user, and presumably, are finding useful information provided freely by others. While this certainly is a relatively inconsequential example, there are a huge number of similar examples, which by their large number amount to a significant case of information

[1] On the general issue of the importance of competitors as sources of technical information, see Schrader [25] and Appleyard [2].

[2] For a survey of the literature on household production see Gronau [10].

provision. (Needless to say, the academic seminar is such a case.)

A second—and to some extent more economically profound—modern example concerns the development of public domain software. Two famous cases are the Free Software Foundation (Gnu project), and the Internet LINUX community.[3] Users of this software are free to develop their own modifications and to integrate them, under very mild constraints, in the community's product. Such innovations are costless, at least for people who enjoy developing computer skills and interacting with other computer wizards. Two aspects of the LINUX developer community are worth noting. The first is that there is a threshold level of technical skill necessary before one can reasonably join. That is, I must be reasonably adept before I will be able to offer anything to the community and, further, to be able to benefit from its current development activities.[4] Second, while it is possible to be relatively passive, acting merely as a recipient of information (essentially free riding), reputation effects are extremely important in communities like the LINUX community. Postings by different participants are paid differing amounts of attention by the community. In addition, merely passive participants will be excluded from any interactions that are personalized, whether bilateral exchanges between two developers or fora that are organized by any sort of invitation. Thus the 'inner group', or presumably those at the cutting edge, advancing the software most rapidly, will have strong similarity constraints on the levels of their expertise. A final observation is that the community self-organizes into sub-groups—collections of agents interested in the same sub-piece of the system. Within these sub-groups similarity will be even greater, as all members will be specialists of the same aspect of the system. Hence we would expect that when viewed at an aggregate level, successful communities display heterogeneity and complementarity in knowledge. But within this heterogeneous population there are sub-groups within which agents' knowledge is very similar. This seems to be the way the LINUX community has self-organized during recent years.

When collective invention is taking place, the key to innovative performance is in the transmission of information. Broadcasting new knowledge, giving it away freely, seems a vital part of the process.[5] But most of this broadcasting is not entirely global—broadcasts are only to a subset of potentially interested agents. When this is the case, the structure of the network over which transmission of information takes place may be vitally important to the performance of the industry.

To embed these phenomena in a formal model is the objective of this paper. In

[3] See Ghosh [8] for a discussion of LINUX and the free provision of information on the Internet.

[4] In this case, it has become possible for non-innovating users to benefit from past development, now that commercial installation packages are available. But these packages do not contain the latest developments of the system, and so are relatively static.

[5] As debates around the turn of the century in the British non-ferrous metals industry indicate, 'freely' does not mean that the broadcaster never gains. It takes place in a context in which all (local) agents are doing the same thing, and are in a sense contributing to or creating a public good that makes all better off.

the model knowledge diffuses through broadcast among groups of agents. In such a framework, one agent's mistakes or discoveries will benefit to the people he interacts with, and innovations in effect take place as a result of the broadcasting of knowledge. But pieces of knowledge generated this way can be beneficial only to those agents who are at least partly capable of integrating them into their own private knowledge endowment. Thus there is a threshold value for agents' dissimilarity beyond which no knowledge transmission is possible, meaning that if i and j are too dissimilar they cannot learn from each other.

For this economy, we measure aggregate performance as the mean knowledge level over all agents. The tuning parameter is the degree of spatial randomness in the links between agents through which knowledge can pass. At one extreme we therefore have an entirely local network—every agent is connected to his n nearest neighbours, whereas at the other extreme agents are connected randomly to, on average, n other agents, located anywhere in the space. We examine the space of network structures between these extremes and find that one region of the space stands out: the 'small world' network structure as defined formally by Watts and Strogatz [28] and Watts [27]. The small world is a structure in which there is a strong degree of local cohesiveness but also a small fraction of 'long-distance' links which allow knowledge to be circulated among 'distant' parts of the economy. Hence the small world appears to be a good candidate target for policy intervention. But in this portion of the space of possible networks, not only the mean knowledge level over all agents, but also the dispersion of knowledge among agents are maximized. Thus having identified this region the policy maker would be left with a dilemma between efficiency in terms of how much knowledge diffuses through the system, and equity in the allocation of this knowledge.

2 The Model

We begin by giving a schematic description of the model, before turning to the network architecture and the details of the dynamics of knowledge distribution, acquisition and creation.

2.1 A schematic description of the model

In our economy, many agents are located on a graph, each agent having direct connections with a small number of other agents. Each agent has a knowledge endowment in the form of a real-valued vector. At random times, an agent is selected and broadcasts his knowledge.[6] Knowledge is broadcast to any agent who satisfies two criteria: he has a direct link to the broadcasting agent, and his knowledge endowment is not

[6] The nature of knowledge, and its relation to data and information are both very complex. We avoid entirely these complexities here, and consider "knowledge, data, or information that is able to be broadcast and received by other agents".

too dissimilar. Knowledge is received and assimilated by any agent who can use it. Formally, if i broadcasts to j, then in any knowledge dimension in which i dominates j, j's knowledge increases. Agents broadcast sequentially and randomly, so that they are acting in a changing world. Agent i might have broadcast to j, but before either i or j broadcast again, one of them might have received information from other agents and in so doing have created more opportunities for useful broadcasting. The extent to which knowledge is assimilated by a recipient is a key parameter we also vary. Recipients can assimilate part of the knowledge (when absorptive capacity is less than perfect) or they can super-assimilate, in a regime of collective invention. We examine below the effect that this parameter has on the relation between network structure and knowledge growth.

2.2 Knowledge interaction

Each agent is characterized by a knowledge vector which evolves over time as the agent receives information broadcast by other agents. Formally, let $V_i(t) = (V_{i,k}(t); k = 1, ..., K)$ denote agent i's knowledge endowment at time t. Agent i broadcasts to every $j \in \Gamma(i)$ (equivalently $i \in \Gamma(j)$ as the graph is non-directed), if dissimilarity between i and j is low enough. By dissimilarity, we mean the relative distance between i and j in terms of knowledge, which we write as

$$\Delta(i,j) = \frac{\|V_i - V_j\|}{\|V_i\|} = \sum_k \left(\frac{V_{i,k} - V_{j,k}}{V_{i,k}} \right)^2 \tag{1}$$

For each agent $j \in \Gamma(i)$, provided $\Delta(i,j) < \theta \in (0,1)$, i makes his knowledge available to j. For every knowledge category, k, when i broadcasts, j's knowledge increases according to

$$V_{j,k}(t+1) = V_{j,k}(t) + \max\{0, \alpha [V_{i,k}(t) - V_{j,k}(t)]\}, \text{for all } k = 1, \ldots, K \tag{2}$$

without any consequent loss of knowledge to agent i. The parameter α captures an important aspect of knowledge diffusion and transfer. In some cases, knowledge is only partly assimilable. This notion has been examined as an issue of absorptive capacity by Cohen and Levinthal [4], or Cowan and Foray [6]. When $\alpha < 1$ broadcasting results in partial acquisition of knowledge by the recipient, as well as a partial diminution of the distance between broadcaster and recipient. In a regime of collective invention, however, knowledge is characterized as super-additive, i.e. $\alpha > 1$. Here, unobserved (by the analyst) complementarities in the knowledge stocks of i and j imply that when j receives i's knowledge he is able to improve upon it, innovating by combining his knowledge with the knowledge newly acquired.[7]

[7] Note that α could be intepretated as parameterizing the tacitness of knowledge. If $\alpha < 1$ even in the absence of a dissimilarity constraint the failure to absorb all available knowledge can arise because codified, broadcast knowledge needs to be interpreted and this interpretation intimately involves tacit knowledge, which the receiving firm is unlikely to have. When $\alpha > 1$, the dissimilarity

2.3 The interaction structure

Consider N agents existing on an undirected connected graph $\mathcal{G}(I,\Gamma)$, where $I = \{1,\ldots,N\}$ is the set of vertices (agents) and $\Gamma = \{\Gamma(i), i \in I\}$ the list of connections (the vertices to which each vertex is connected). Formally $\Gamma(i) = \{j \in I \setminus \{i\} \mid d(i,j) = 1\}$, where $d(i,j)$ is the length of the shortest path (geodesic) from vertex i to vertex j. Only agents separated by one edge can interact, and when i broadcasts, only those agents in $\Gamma(i)$ are potential recipients.

Our interest in network architecture differs from that seen in the majority of the literature in that we do not vary the density of the network. The family of graphs we consider here contains a constant number $n \cdot N/2$ of edges . Our concern is with the degree of regularity in the structure. The heuristic by which we examine it can be seen by the 're-wiring' procedure we employ.[8] The space of the graph is a periodic one dimensional lattice—agents are located at fixed intervals around a circle. At one extreme of the space of network structures we have a regular structure: each agent is connected to his n nearest neighbours (n is restricted to being even to avoid introducing spatial asymmetries.) At the other extreme, each agent is connected to, on average, n agents located at random on the lattice. To interpolate, we use the following algorithm. Create the regular lattice structure. With probability p re-wire each edge of the graph. That is, sequentially examine each edge of the graph; with probability p disconnect one of its vertices, and connect it to a vertex chosen uniformly at random. In the algorithm we ensure both that vertices are not self-connected by this procedure, and that there are no duplications, i.e. no two vertices are connected more than once. For large graphs, this procedure ensures that the connectivity is preserved and that the average number of edges per vertex is constant at n. By this algorithm we tune the degree of randomness in the graph with parameter $p \in [0,1]$, hence the label $\mathcal{G}(I,n,p)$ for graphs in this family.

Figure 1 shows three illustrative configurations with increasing disorder as p is increased, for $N = 16$ and $n = 4$. Figure 1 suggests that for small but non-zero values of p, the graph is highly clustered like a regular graph but, as we shall see, shares some features of almost-random graphs.

2.3.1 Two structural properties

The structural properties of $\mathcal{G}(I,n,p)$ graphs can be captured by the concepts of average path length and average cliquishness. To illustrate, in friendship networks, the path length is the number of friendships in the shortest chain connecting two agents, whereas cliquishness reflects the extent to which friends of one agent are also

constraints performs the same function. Agents with similar codified knowledge are likely to have similar, if not the same, tacit knowledge. See Cowan et al. [5] for a discussion of codification and tacitness.

[8] This is the re-wiring procedure employed by Watts and Strogatz [28], in their seminal work on small worlds.

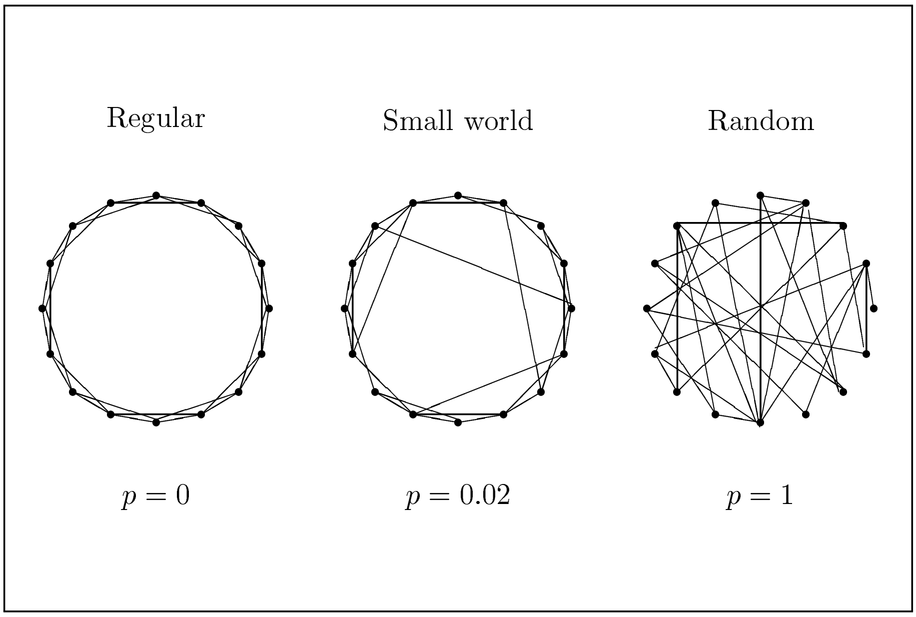

Figure 1: The transition from a locally ordered structure to a random one and the small world phenomenon for $\mathcal{G}(I, n, p)$ graphs

friends of each other. Formally, defining $d(i, j)$ as the length of the shortest path between i and j, the average path length $\mathcal{L}(p)$ is

$$\mathcal{L}(p) = \frac{1}{N} \sum_{i \in I} \sum_{j \neq i} \frac{d(i, j)}{N - 1} \tag{3}$$

and average cliquishness $\mathcal{C}(p)$ is given by

$$\mathcal{C}(p) = \frac{1}{N} \sum_{i \in I} \sum_{j, l \in \Gamma(i)} \frac{X(j, l)}{|\Gamma(i)| \left(|\Gamma(i)| - 1\right)/2}, \tag{4}$$

where $X(j, l) = 1$ if $j \in \Gamma(l)$ and $X(j, l) = 0$ otherwise. The evolution of path length and clique size with p is depicted on figure 2, for a graph of $N = 500$ vertices, each vertex having on average $n = 10$ connections. For the sake of clarity, we plot the normalized values $\mathcal{L}(p)/\mathcal{L}(0)$ and $\mathcal{C}(p)/\mathcal{C}(0)$. The upper curve (thin black) in figure 2 is the normalized average cliquishness index $\mathcal{C}(p)/\mathcal{C}(0)$ for $p \in [0, 1]$. It remains almost constant when p is reasonably small and falls slowly for large values of p. By contrast, average path length (thick grey) as measured by $\mathcal{L}(p)/\mathcal{L}(0)$ falls quickly for very small p values and flattens out near 0.01. As emphasized by Watts and Strogatz, there is a non-negligible interval for p over which $\mathcal{L}(p) \simeq \mathcal{L}(1)$ yet $\mathcal{C}(p) \gg \mathcal{C}(1)$. This interval, in which high cliquishness and low path length coexist, constitutes the small world region.

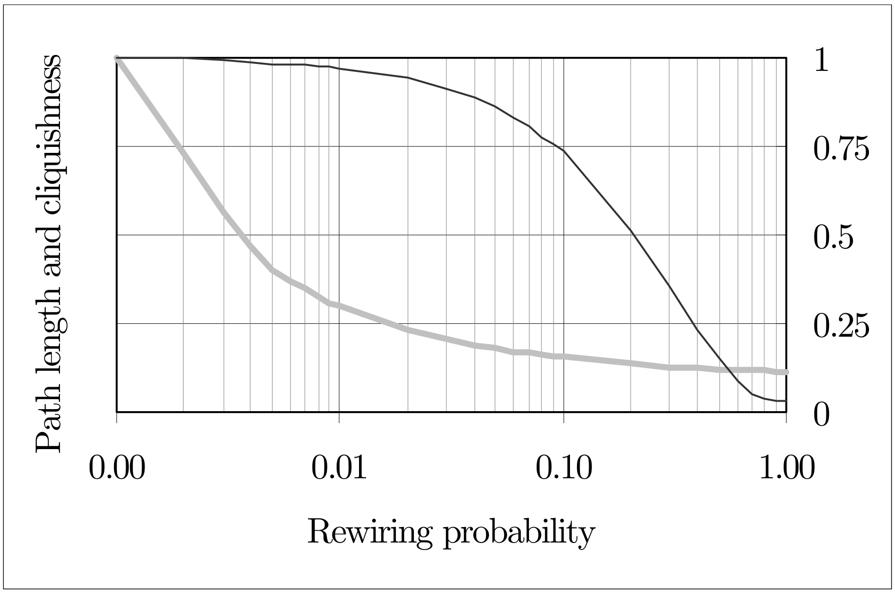

Figure 2: Average cliquishness and average path length as functions of p

2.4 Small worlds

Watts and Strogatz define a small world network as a structure in which there is a coincidence of cliquishness and short average path lengths between agents. Roughly speaking this occurs when a regular structure is only slightly disturbed (a small value of p in the re-wiring algorithm described above). This creates an architecture that is highly regular (preserving overlapping and therefore cliquishness) but with a small number of short-cuts across the graph, which significantly lowers average path length. Watts and Strogatz conjecture that the small world structure is common in natural networks, and give three examples.[9] This is consistent with findings in sociology, especially Milgram's six degrees of separation or the work by Granovetter [9] on the importance of indirect links, that suggest small world structures for social networks. Cowan and Jonard [7] find similar small world properties in three knowledge networks,[10] and provide a simulation study of a model in which the network structure is similar to that of the model described here, but in which knowledge is diffused through bilateral barter exchanges. In that model the small world structure is significant: average long run knowledge levels are higher in small worlds than in either more regular or more random networks. Curiously, the variance in knowledge levels

[9] They provide evidence showing that the networks of film co-stars, the power grid in the western U.S. and the neural network of the worm *Caenorhabditis Elegans* all have the small world structural features of high cliquishness and short average path lengths.

[10] These networks are a network of innovating Dutch firms; the network of research institutes involved in one activity of the EU's TSER programme; and a network of innovating firms participating in the BRITE/EURAM programme.

across agents is also highest in the small worlds.

3 Simulation of the broadcast model

We are interested generally in the relationship between the structure of the network across which knowledge diffuses and the distribution power of the innovation system. It is natural therefore to examine the evolution of knowledge levels in this economy. We can do this by simulating the economy and relating long run knowledge levels to the value of p in our re-wiring algorithm. Of particular interest is whether the model displays small world properties. Define agent i's average knowledge level at time t as $\overline{\mu}_i(t) = \sum_k V_{i,k}(t)/K$. The average level of knowledge in the economy at time t is

$$\overline{\mu}(t) = \frac{1}{N}\sum_{i \in I} \overline{\mu}_i(t) \tag{5}$$

and the variance in knowledge allocation is

$$\sigma^2(t) = \frac{1}{N}\sum_{i \in I} \overline{\mu}_i^2(t) - \overline{\mu}^2(t). \tag{6}$$

We consider a third measure, having to do with the degree of spatial order in a system of locally interconnected components. To check whether our broadcast economy generates spatially auto-correlated knowledge allocations, we compute the Moran coefficient defined as follows (see McGrew and Monroe [16]). Recall $d(i,j)$ is the geodesic between i and j. Dropping time indexes, the coefficient is written

$$\mathcal{S} = \frac{1}{\sigma^2}\sum_{i \in I}\sum_{j \neq i} w_{i,j}\left(\overline{\mu}_i - \overline{\mu}\right)\left(\overline{\mu}_j - \overline{\mu}\right), \tag{7}$$

where

$$w_{i,j} = \frac{1/d(i,j)}{\sum_{i \in I}\sum_{j \neq i} 1/d(i,j)}. \tag{8}$$

Were all weights to equal one, index $\mathcal{S}$ would be the correlation coefficient, as it would reduce to the ratio of the covariance to the variance. So spatial correlation is simply a weighted correlation coefficient, where weights are given by inverse distances.

We consider an economy with $N = 500$ agents, each having, on average $n = 10$ direct connections to other agents. Within a single simulation run the network structure is fixed, so each possible economy is represented by a graph from the family $\mathcal{G}(I, 10, p)$ with $I = \{1, \ldots, 500\}$.[11] Each agent is endowed with a knowledge vector of length 5, with each element initialized randomly from a uniform distribution between 1 and 10. In simulation time, in each period one randomly selected agent broadcasts his knowledge. The knowledge is received by all those who are directly connected

[11] The issue of the endogeneity of the network structure is not addressed here. This is typical in this literature. Two notable exceptions, in which endogenous network strutures are the focus of the analysis, are Bala and Goyal [3] and Plouraboue et al. [20].

to him and are sufficiently similar. Formally, the broadcast of agent i is received by the set of agents $\{j \in \Gamma(i) \mid \Delta(i,j) < \theta = 0.2\}$. We consider a large number T of periods, to ensure that each agent takes part in a reasonably large number of knowledge transmissions, both as sender and receiver. With $T = 30,000$, the average agent broadcasts 60 times.

We are interested in the effects of two parameters. The first is p, which determines the network structure. This we vary from 0 to 1 in units convenient to display on a log scale. The second is α, which determines the extent to which knowledge can be assimilated. We vary α from 0.95 to 1.2. Values of $\alpha < 1$ indicate a regime in which tacit knowledge is vitally important, and knowledge received is only absorbed with considerable effort. Thus the effects of receiving a broadcast are a partial absorption by the recipient and thus a diminution of the dissimilarity of sender and receiver. When $\alpha > 1$ we are in a regime in which knowledge is super-additive. The recipient of knowledge is able to leap-frog the sender. This corresponds to Allen's description of collective invention, in which knowledge is relatively simple to absorb and easy to improve upon.

4 Results

The first set of results we examine has to do with the efficiency of knowledge creation and diffusion. We then turn to the spatial behaviour of the system.

4.1 Efficiency and equity in knowledge diffusion

We examine the efficiency of the network structure in terms of long run average knowledge levels. The statistic we use is the average knowledge level in the economy $\overline{\mu}(T)$. The parameter space can be partitioned into two distinct regions corresponding to $\alpha le 1$ and $\alpha > 1$. In the first, knew knowledge is not created, and changes in aggregate knowledge level arise purely through diffusion. In the second there is both creation and diffusion. We consider them separately

4.1.1 Diffusion

Consider first the case in which absorptive capacity is less than perfect—agents are able to assimilate only part of the knowledge with which they are presented. This corresponds to the region of the parameter space where $\alpha < 1$. Here, a purely diffusive mechanism is at work—knowledge is diffused, but not generated. In this regime, network structure has no apparent effect on long run knowledge levels. All values of p produce the same long run state of the economy. The effect of changing average path length between agents is to change the speed of convergence — shorter path lengths imply faster convergence.

Contrasting results concerning this region of the parameter space are presented in Cowan and Jonard [7], who show that a small-world architecture dominates other forms of organization. This result is obtained in a model in which knowledge is diffused not by broadcast but by barter trades among agents. Broadcasting, as opposed to barter, eliminates the need for a double coincidence of wants for knowledge diffusion. Hence eliminating this requirement eliminates the advantage of cliquishness, which lies in the fact that in a cliquish world a failure of the double coincidence of wants does not have dramatic consequences for there are many other possible paths along which a piece of knowledge can travel.

4.1.2 Creation and Diffusion

In the second region of the parameter space, where $\alpha > 1$, we get a joint process of knowledge creation and diffusion. Each agent incorporates the knowledge he receives into his existing stock and becomes more knowledgeable, hence achieving higher efficiency than before transmission.

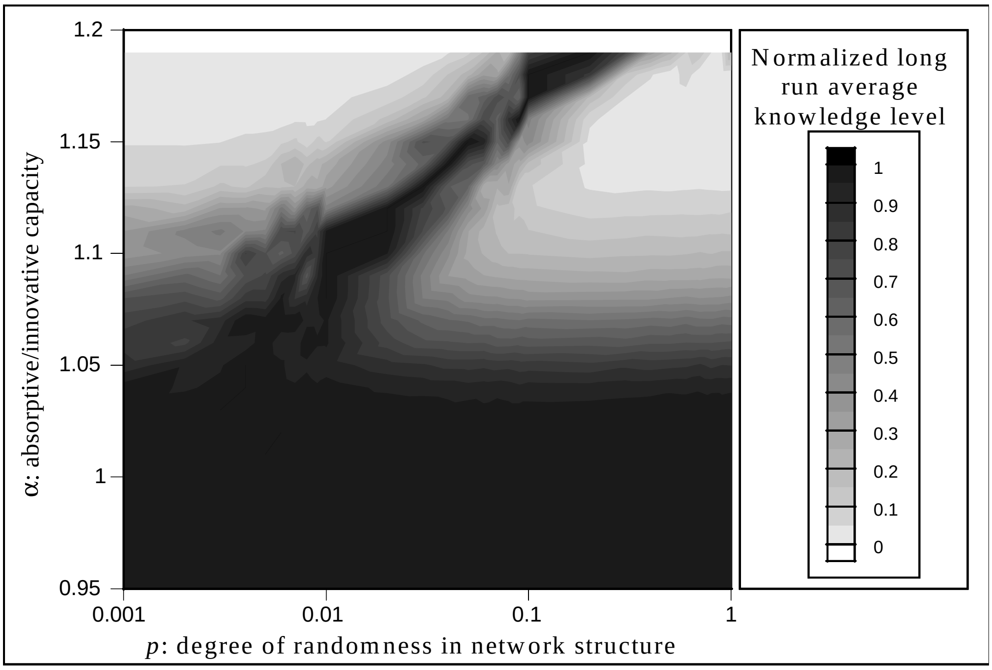

Figure 3: Knowledge levels in the (p, α)-space

Figure 3 depicts the relationship between the overall knowledge level and both the re-wiring probability p and the absorptive/innovative capacity α. We plot normalized rather than absolute values, to preserve legibility.[12] For α lower than 1.05, no clear-cut distinction exists between ordered and random graphs. With finer graining, we

[12] What we plot is actually $\overline{\mu}(T)$ divided, for each α-value, by the maximal value it takes for $p \in (0,1)$, i.e. $\overline{\mu}(\alpha, p) / \max_{p'} \overline{\mu}(\alpha, p')$.

see evidence of a smooth change starting at $\alpha = 1$. By contrast, when α exceeds approximately 1.05, clear patterns emerge even with coarse graining of the parameter space and a sharp efficiency peak obtains in the small world. At the same time, the knowledge-maximizing p value (degree of randomness) increases monotonically with α, while remaining confined to the small-world region, i.e., the area between $p = 0.01$ and $p = 0.1$. The maximizing p-value always lies in this region, but it is worth mentioning that most of the mass of the distribution itself is also located between 0.01 and 0.1. Outside this region, for every value of α, long run average knowledge levels are much lower than the maximum. Hence a small-world architecture, i.e., one in which there is at the same a significant amount of local order and a few random connections, achieves the best overall performance in the broadcast economy we consider.

Several remarks are in order here. First, in our experiments we ignored α values greater than $1+\theta$. Recall an agent who receives information not only incorporates all of it but improves upon it (by $\alpha-1$). This implies that after receiving a transmission, the knowledge level of the recipient is higher than that of the sender. If $\alpha > 1+\theta$ the new knowledge level of the recipient would be so much greater than that of the sender that the sender now fails to pass the dissimilarity threshold vis-à-vis the former recipient. That is, if I send you knowledge, you improve so much that I cannot recieve from you. This seem entirely unreasonable, so we have excluded this possibility by restricting α to be less than $1+\theta$.

In order to explain the mechanism underlying the general results, we consider the effects of cliquishness and path length on system performance. As an initial step in that explanation, consider the role of dissimilarity. In the knowledge re-combination world we consider here, the innovative jump a recipient makes increases with his lag vis-à-vis the sender as long as their distance in terms of knowledge endowments stays below θ. (see equation 1). To grasp the long term implications of this, consider knowledge category k in which i dominates j, i.e. $V_{i,k}(0) > V_{j,k}(0)$. If i and j repeatedly interact, then after the first broadcast one has

$$V_{j,k}(1) = V_{i,k}(0) + (\alpha-1)\left[V_{i,k}(0) - V_{j,k}(0)\right], \tag{9}$$

and after j broadcasts back to i one gets

$$V_{i,k}(2) = V_{i,k}(0) + (\alpha-1)\left[V_{i,k}(0) - V_{j,k}(0)\right]\{1 + (\alpha-1)\}. \tag{10}$$

It can be readily seen that as long as $\alpha \in [1,2)$, both quantities converge to the same value of

$$V_{i,k}(0) + \left[V_{i,k}(0) - V_{j,k}(0)\right](\alpha-1)\sum_{t=0}^{\infty}(\alpha-1)^t, \tag{11}$$

which yields

$$\lim_{t\to\infty} V_{i,k}(t) = \lim_{t\to\infty} V_{j,k}(t) = \frac{V_{i,k}(0) - (\alpha-1)V_{j,k}(0)}{2-\alpha}. \tag{12}$$

Hence if we focus on two agents repeatedly broadcasting to each other, their knowledge level converges to a value that is an increasing function of both the initial level of the more knowledgeable agent and the initial gap between the two agents. Of course the larger α the higher the limit, but this was intuitively more clear. So big gaps in this model lead to high long-run levels, unless initial gaps are so large that they preclude transmission.

As mentioned earlier, there are two dimensions when knowledge growth is considered, namely creation and diffusion. Though in the model knowledge creation and knowledge diffusion are distilled into a single episode, it is intuitive that cliquishness mainly influences knowledge creation and path length drives knowledge diffusion. We shall now examine the importance of these two forces in knowledge growth.

The regular world is one in which an agent's potential recipients are recipients of each other. So when a piece of knowledge is released, a self-reinforcing local mechanism is at work, and produces fast localized knowledge growth within an agent's neighborhood. This is the knowledge generation part of the dynamics, which mainly takes place at the intra-clique level. It entails unequal knowledge growth across clusters of agents. By and large, this produces strong inequalities between groups that advance quickly and those that lag behind. However, fast advancing groups can achieve very high results as there is no homogenizing tendencies provoked by other parts of the graph. In Figure 4 we summarize long-run knowledge states in three network structures, and the regular world is the first of these graphs.

At the other extreme we have the random world in which in only a few transmissions knowledge is passed through the whole population due to short travel distances. In this situation there is a strong tendency for knowledge levels to homogenize, and we have seen that this is bad for knowledge creation, as homogeneity hurts more than it helps (see equation 12). Knowledge diffusion is efficient (hence the homogenization) but disruptive in that it squeezes the dispersion of people in terms of knowledge, thereby leaving little chance for knowledge creation. The random world is the lowest part of Figure 4, and homogeneity is patent.

Between these extremes there is a trade off between knowledge creation and knowledge diffusion, and the small world turns out to be the knowledge-growth maximizing architecture. It shares most of the advantages of the two extremes. It connects distant and possibly heterogenous parts of the graph, creating the possibility for large innovative jumps. Shortcuts bring together parts of the graph in which, at the same time, local reinforcement is at work. There is a tendency to homogenize overall knowledge levels but without weakening too much local cumulativeness, thus we end up with levels that are larger than in a homogenous random graph, but relatively homogeneous even at the global level. Homogenization is stronger as soon as shortcuts are introduced, though as long as there are not too many the local reinforcement

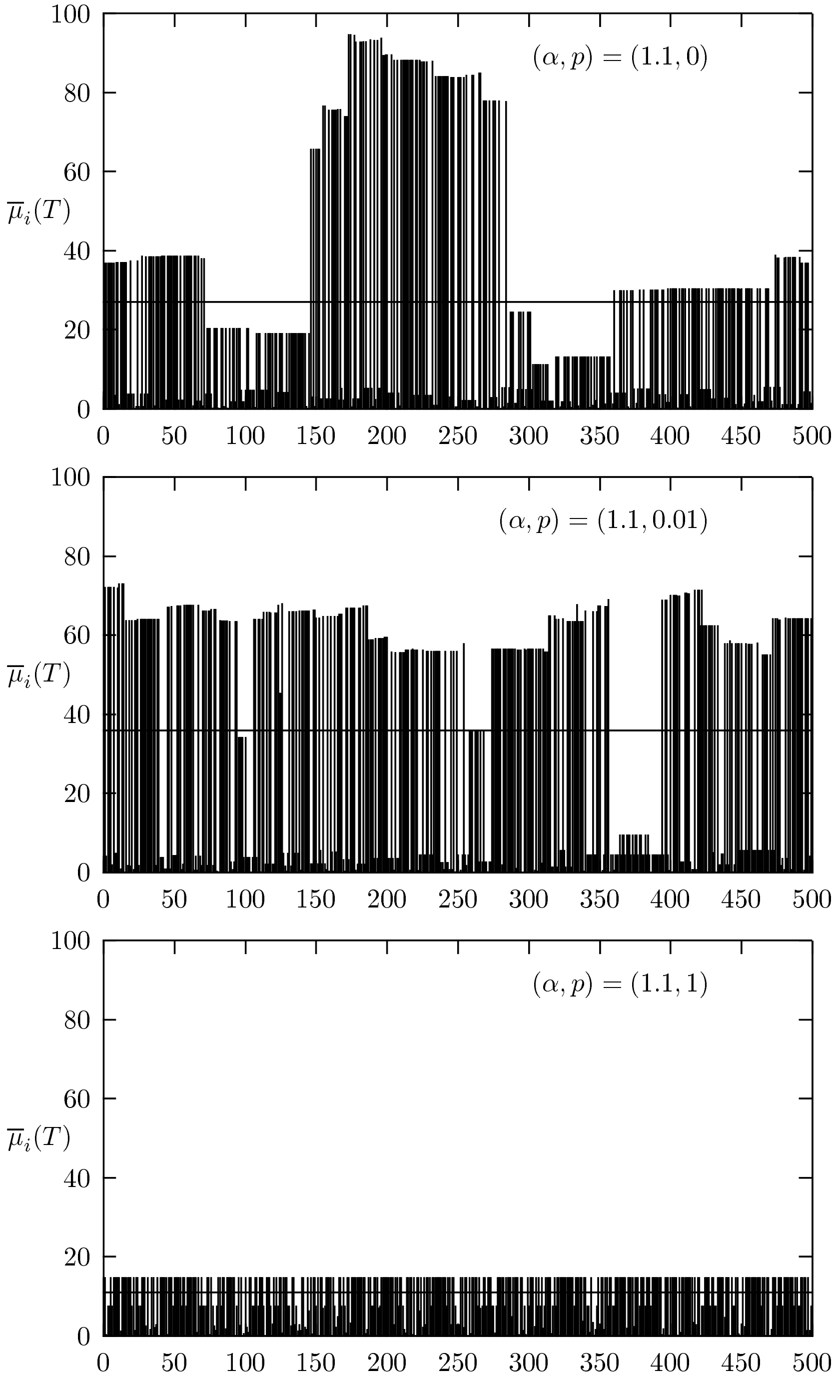

Figure 4: The average individual long-run knowledge levels in three typical simulations runs for $\alpha = 1.1$; population averages are shown as horizontal lines

stemming from large degrees of cliquishness is preserved. This is shown in the middle part of Figure 4. Corresponding average values are represented in Figure 4 by horizontal lines, and we have $\overline{\mu}(0,T) = 26$, $\overline{\mu}(0.01,T) = 36$ and $\overline{\mu}(1,T) = 11$.

A second comment one can make is that there is an apparent trend in the optimal number of shortcuts. This arises in the marginal contribution of a random connection as α is varied. When a shortcut is added, there is at the same time a loss from decreasing cliquishness—harming local accumulation—and benefits from decreasing path length—making feasible the possibility of connecting highly knowledgeable parts of the graph.

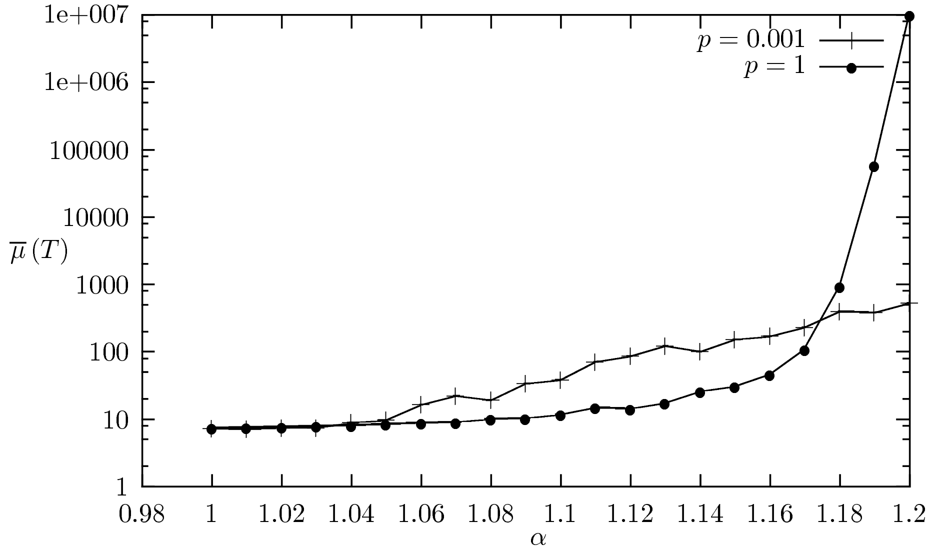

Figure 5: The population average long-run knowledge level in two polar cases, as a function of α

As α increases these two effects grow in importance, but what matters is actually their relative importance. If the marginal gain from decreasing path length increases faster than the marginal cost of decreasing cliquishness, then (if the solution is interior) the optimal number of shortcuts (value of p) will increase with α. This cannot be checked directly since the two effects are not separable within the context of the model. We can perform an indirect check in the following way, however. In a regular world, path lengths are long, cliquishness is at its maximum. Thus the main driver of growth in knowledge levels will be the local reinforcement that arises from cliquishness. By contrast, in the random word cliquishness is minimized, as is path length, so growth in average knowledge levels will be driven by the fact that diffusion is rapid. Plotting long run knowledge levels in these two extreme worlds against different levels of α will give an indication of how the two effects (increasing cliquishness and decreasing path length) change in magnitude with α. This we do in Figure 5. The benefits of cliquishness increase slowly but steadily with α; the benefits of short paths rises rapidly when α becomes relatively large. This suggests that effects of decreasing path length does indeed increase faster than the costs of decreasing cliquishness. Figure 5 suggests that at some point, i.e., for large enough α-values, the benefits from shorter paths exceed the losses from reduced cliquishness.

We turn now to the issue of equity in the distribution of knowledge across agents. Figure 6 shows the relationship between the economy-wide knowledge variance as a function of the re-wiring probability p and the absorptive/innovative capacity α. Again we present normalized values.

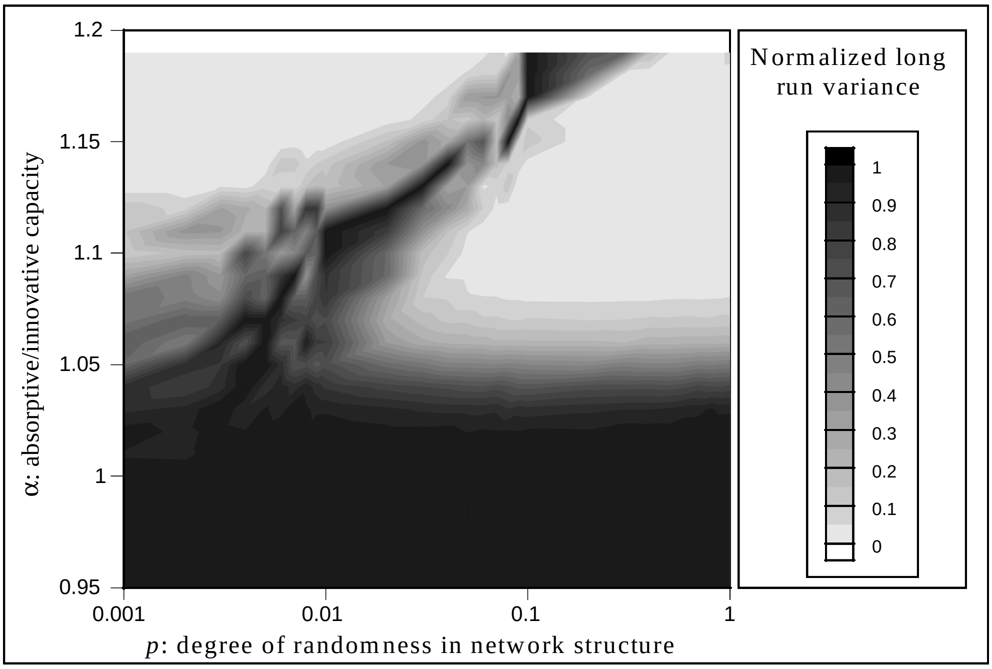

Figure 6: Long run heterogeneity in knowledge allocation in the (p, α)-space

Surprisingly, Figure 3 and Figure 6 are very similar (though dispersion is lower in Figure 6), with highest inter-agent differences observed for a degree of randomness monotonically increasing with α while remaining confined to the area between $p = 0.01$ and $p = 0.1$. We therefore come to the somewhat uncomfortable conclusion that the small world region is at the same time generates the best overall performance in terms of how much knowledge is generated by the system, and the worst overall performance when homogeneity of allocation is considered desirable. Figure4 indicates that the source of this variance is that some agents get left behind. Roughly the same number of agents get left behind in each world, and stay close to their initial knowledge levels, but in the small world, the agents who do advance advance rapidly and far, thus creating a large gap between themselves and those left behind.

4.2 The spatial allocation of knowledge

Spatial correlation of knowledge levels can be considered either in geographical space on in the space of the network itself. In general, a positive spatial correlation exists if agents 'near each other' have comparable knowledge vectors. By contrast, negative correlations obtain when knowledgeable agents are the neighbors of laggards and no clustering of knowledge appears. Correlation in the geographic space means taking as the distance between nodes i and j the simple absolute difference $|i - j|$ (with the adequate modulo). In that sense, node i is very close to nodes $i \pm 1$, slightly less close to nodes $i \pm 2$ and so on. A priori, since knowledge generation and diffusion takes place over the network, there is no reason in general to expect spatial correlation in geographic space. For small values of p however, geographic space has a very similar topology to network space (recall that the spaces differ in roughly p percent of the edges). Thus, if there are non-trivial correlations for small p values in the network space, there should be echoes of them in geographic space.

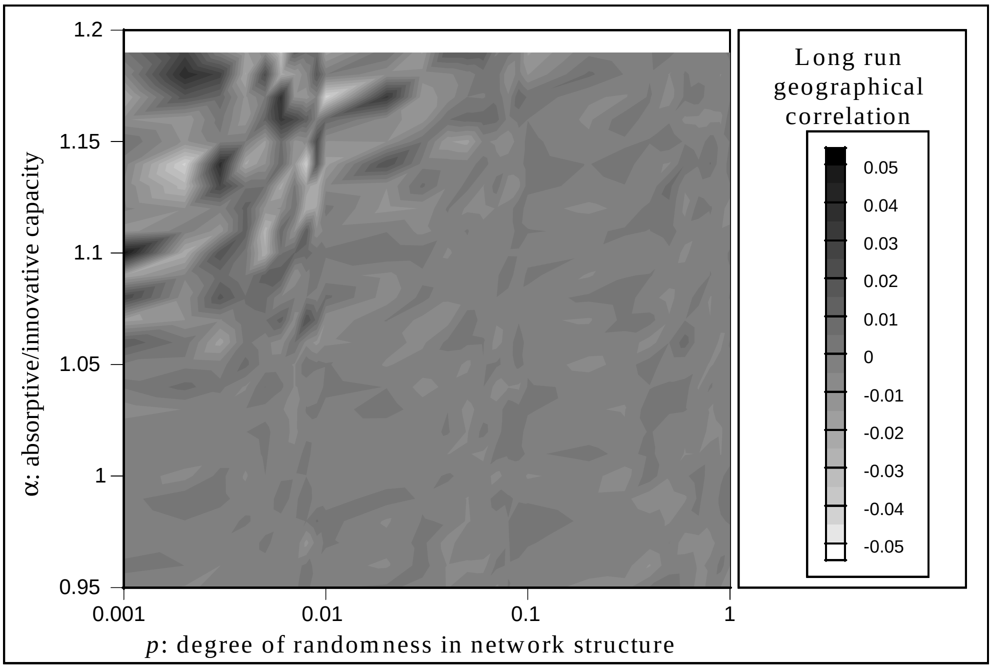

Figure 7: Spatial correlation in the geographical space

Figure 7 shows the geographical spatial correlation as a function of p and α. As expected, there is virtually no spatial correlation in this space. All the values are small (in absolute value) and in general not statistically significantly different from zero. There is one region that stands out in contrast. For large α and small p, many of the correlations, while small, do differ significantly from zero.[13] On the time scale

[13] This was determined using a standard two-tailed t-test.

we consider, though, this relationship between p, α and spatial correlation contains considerable randomness.

Figure 8 gives the Moran coefficient $\mathcal{S}$ as a function of the degree of randomness p and the absorptive/innovative capacity α, when the network metric is considered. A very different picture obtains in the network space, with three distinct regions. There is a wide band that goes from the lower left corner of the (p, α)-space (low absorptive capacity in a regular world) to the upper right corner (significant super-additivity in a random world) where knowledge is more or less randomly allocated among agents. There is a region in which the world is random and absorptive capacities are imperfect, in which a (small) negative correlation is observed. Finally, positive spatial correlation obtains when the network of agents' relationships is cliquish and innovation rates are high, as we would expect.

The correlation in network space for small p values was indeed echoed in geographic space. The echo is very weak however. Even a small discrepancy between the space in which knowledge moves and the space in which correlations are measured creates a very different impression on the existence and strength of clustering.

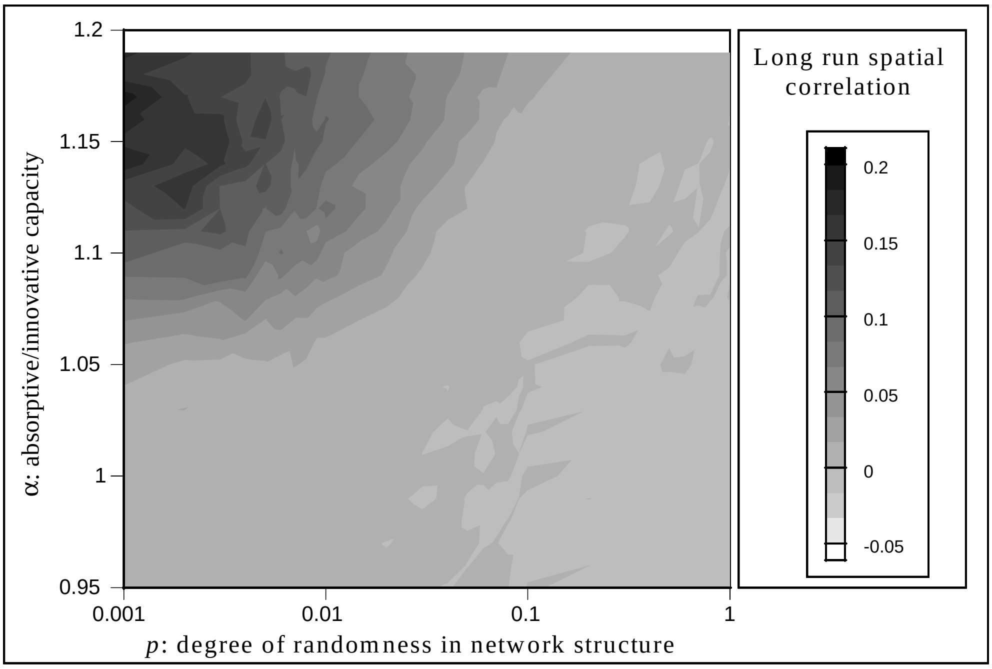

Figure 8: Spatial correlation in the network space

It is clear from the results presented in Figures 7 and 8 and the discussion, that there is no small work effect on spatial correlation or knowledge clustering. Cliquishness is a clear-cut source of knowledge clustering and introducing shortcuts to possibly distant parts of the graph only tends to impede clustering.

5 Conclusion

Collective invention is a historically important phenomenon which has occurred in different eras and locales. Common to those episodes is that this organization of innovation generates rapid technical advance. One of the important features of collective invention is the sharing of information among a broad, though typically localized, group of agents. It is this feature that we have emphasized here, showing how the results of this sharing is affected by the network structure over which communication takes place. Unless technical opportunities are extremely large, and thus innovation is very straightforward, small-world network structures produce fastest knowledge growth rates. Cliquishness is in general a good thing, but the ability to bring knowledge into the clique from outside has a vital role. As technological opportunities grow, though, there can be too much cliquishness—because innovation involves capital investments that depreciate over time, it can happen that with a rapid innovation rate and strong super-additivity in knowledge, an agent can be left behind by members of his clique, as they use knowledge he has generated but which has created for him a temporary lock-in due to his capital investments. In this sort of situation a very wide variety in knowledge sources is important for an agent to keep in the race.

Collective invention is more than a historical curiosity. Today we see examples of innovation which, while different in some respects from the historical examples, could certainly be referred to as collective. The most salient examples exist in software development, and in particular that mediated by the Internet. LINUX and the Free Software Foundation are the two most obvious case. Here, and in LINUX in particular, technical change has been very rapid, and done collectively, through massive information sharing.[14] When bugs appear or new tasks are described, any programmer (in the world) interested can contribute.[15] This is true only because the basic code, and all bug fixes and additions are published freely. That is to say, they are broadcast. Every member of the community benefits from the work of other members. The community is not monolithic, though. If the overall project is large, it is typically divided into sub-projects, each having a "devoted core" of participants. Within these sub-projects people work on similar things, so we see a degree of homogeneity at a micro level, but considerable heterogeneity at the aggregate level. This dichotomy between micro and macro descriptions of the system suggest that the communication network may be important in creating the rapid advances that we observe.[16] They suggest further that the kinds of results presented in this paper

[14] One famous example of the rapidity of this type of software development concerns an Internet security bug which affected all systems. LINUX corrected the bug long before any commercial system (Lang [14] quoting Planete Internet, 1997, p. 60).

[15] See `http://www.cosource.com` for a typical task list in an open source software development project. For more general information about how this style of software development proceeds, see `http://www.openresources.com`

[16] That LINUX is not merely a one-off phenomenon is shown by Raymond [22] who describes a successful deliberate attempt to mimic the LINUX development organization for a different software

are pertinent to understanding the modern version of collective innovation.

The key feature that makes today's collaborative software generation similar to the collective invention of the past is the sharing of technical information among a group of user/developers. Each user/developer uses the advances made by the others both in their use and in their development. One distinction between the current phenomenon and past episodes appears to be that historically, the group among whom information was shared was circumscribed, whereas today it is not. Then, the circumscription was geographic—steel makers in Cleveland; or paper makers in Berkshire. Programmers working on LINUX are located all around the globe. There is a sense in which the broadcasting is indeed global. Typically the results of innovations or improvements (bug fixes or added features) are posted on bulletin boards accessible by anyone with an Internet connection. But there are two very important ways in which the information broadcasting is not global. First, many important communications (and one could argue that these are the most important from the point of view of advancing the technology) take place via mailing lists. A mailing list is a form of localized broadcasting (viz. "narrow-casting") and distribution of information. This is non-global from the distributor's point of view. Only a subset of the population gets the messages. Second, from the receivers' point of view distribution is local as well. Receipt through mailing lists is obviously local, but even the bulletin board distribution system is made local through the fact that potential recipients do not look at every possible bulletin board. There is a distinction between potential accessibility (anyone can look at this bulletin board) and actual access (only some people do so). Typically, participants in these communities check a select few bulletin boards or news groups frequently. This creates a localization of information flow even through bulletin boards, and creates cliques, neighbourhoods or communities of "readers of bulletin board X". And certainly, information passes between groups when an agent is a member of both groups X and Y, reads something on X and posts it on Y. Thus even though the Internet is potentially a global information diffusion tool, the analysis presented here, in which actual distribution is in the first instance local, remains relevant.

Indeed the analysis here may gain relevance because of the growth of the Internet. Even commercial software firms are seeing the potential available in these communities, or communities of communities. In 1998 the software world was not particularly surprised by the decision of Netscape to distribute its browser for free. But it was "stunned" by the decision to publish the source code (Hamerly et al. [11]). Netscape made this decision after seeing and reading about the kinds of software development that were taking place in this way.[17] What the Internet may have done is to create

package.

[17] Raymond [22] (section 13) quotes E-mail from Eric Hahn, Executive Vice President and Chief Technology Officer at Netscape stating that the Netscape's thinking was inspired by writings about this style of software development.

a large enough community of "collaborators" to make the gains from collaboration outweigh the costs perceived in loss of proprietary control (which may, of course, include loss of market power, technological secrets, trade secrets and so on). There are strong indications that this is so in the development of computer software. What remains to be discovered is whether the development of other technologies can benefit in a similar way.

References

[1] R. Allen. Collective invention. *Journal of Economic Behaviour and Organization*, 4:1–24, 1983.

[2] M. Appleyard. How does knowledge flow? inter firm patterns in the semiconductor industry. In *The Competitive Semiconductor Manufacturing Human Resources Project: Second Interim Report*, University of California at Berkeley, 1999. Institute of Industrial Relations. Available at http://socrates.berkeley.edu /iir/research/csm-hr/chap10.

[3] V. Bala and S. Goyal. Learning from neighbours. *Review of Economic Studies*, 65:595–621, 1998.

[4] W. Cohen and D. Levinthal. Innovation and learning: The two faces of research and development. *The Economic Journal*, 99:569–596, 1989.

[5] R. Cowan, P. David, and D. Foray. The explicit economics of knowledge codification. Technical report, 2000. Paper prepared for the European Contract "TIPIK: Technology and Infrastructure Policy in the Knowledge-Based Economy", programme TSER, DG XII (1998/2000).

[6] R. Cowan and D. Foray. Quandaries in the economics of dual technologies and spillovers from military to civilian research and development. *Research Policy*, 24:851–868, 1997.

[7] R. Cowan and N. Jonard. Network structure and the diffusion of knowledge. Technical Report 99028, MERIT/Maastricht University, Maastricht, The Netherlands, 1999. submitted to Economic Journal.

[8] R. Ghosh. Cooking pot markets: An economic model for the trade in free goods and services on the internet. *First Monday*, 1998. Available at http://www. firstmonday.org/ issues/issue3 3/ghosh/index.html.

[9] M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78:1360–1380, 1973.

[10] R. Gronau. Home production: A survey. In O. Ashenfelter and R. Layard, editors, *Handbook of Labor Economics*. North-Holland, 1986.

[11] J. Hamerly, T. Paquin, and S. Walton. Freeing the source: The story of mozilla. In C. DiBona, S. Ockman, and M. Stone, editors, *Open Sources: Voices from the Open Source Software Revolution*. O'Reilly, 1999.

[12] A. Hyde. The wealth of shared information: Silicon valley's High-Velocity labor market, endogenous economic growth, and the law of trade secrets. Technical report, Rutgers University, 1998.

[13] N. R. Lamoureaux. Accounting for capitalism in early american history: Farmers, merchants, manufacturers, and their economic worlds. Technical report, UCLA, 1999. Available at http://www.sscnet.ucla.edu/history/ activities/ usccpapers/lamoreaux.html.

[14] B. Lang. Free resources and independent technology in the information sectors. 1997. Available at http://www.openresources.com/documents/free-resources-independent.

[15] J. A. McGaw. *Most Wonderful Machine: Mechanization and social Change in Berkshire Paper Making, 1801-1885*. Princeton University Press, 1987.

[16] C. McGrew and J. Monroe. *An Introduction to Statistical Problem-Solving in Geography*. W. C. Brown Publishers, 1993.

[17] R. R. Nelson. The simple economics of basic scientific research. *Journal of Political Economy*, 67:297–306, 1959.

[18] R. R. Nelson. The link between science and invention: The case of the transistor. In *The Rate and Direction of Inventive Activity*. National Bureau of Economic Research, Princeton University Press, 1962.

[19] W. D. Nordhaus. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. MIT Press, 1969.

[20] F. Plouraboue, A. Steyer, and J.-B. Zimmermann. Learning induced criticality in consumers adoption pattern: A neural network approach. *Economics of Innovation and New Technology*, 6:73–90, 1998.

[21] W. Powell, K. W. Koput, and L. Smith-Doerr. Inter-Organizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 41:116–145, 1996.

[22] E. S. Raymond. The cathedral and the bazaar. 1999. Available at http://www.openresources.com/documents/cathedral-bazaar.

[23] P. Russo. Technical change and the industrial district. *Research Policy*, 14, 1985.

[24] A. Saxenian. Regional advantage: Culture and competition in silicon valley and route 128. mimeo, 1994.

[25] S. Schrader. Informal technology transfer between firms: Cooperation through information trading. *Research Policy*, 20:153–170, 1991.

[26] E. Von Hippel. *The Sources of Innovation*. 1998.

[27] D. Watts. Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, 105:493–527, 1999.

[28] D. Watts and S. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:400–403, 1998.