Knowledge Creation, Knowledge Diffusion and Network Structure

Robin Cowan[1,3] and Nicolas Jonard[2,3]

[1] MERIT, University of Maastricht,
P.O. Box 616, 6200 MD Maastricht, The Netherlands
[2] CNRS, BETA, Université Louis Pasteur,
61 Ave. de la Forêt Noire, 67000 Strasbourg, France
[3] International Institute of Infonomics (IIoI),
P.O. Box 32, 6333 PS Heerlen, The Netherlands

Abstract. This paper models knowledge creation and diffusion as processes involving many agents located on a network. Knowledge diffusion takes place when an agent broadcasts his knowledge to the agents to whom he is directly connected. Knowledge creation arises when agents receive new knowledge which is combined with their existing knowledge stocks. Thus both creation and diffusion are network-dependent activities. This paper examines the relationship between network architecture and aggregate knowledge levels. We find that knowledge growth is fastest in a "small world". That is, when the underlying network structure is relatively cliquish (dense at a local level) yet has short paths. This corresponds to a locally-connected graph which includes a few long-distance connections or shortcuts.

1 Introduction

One reason to reject the representative agent model is the observation that economic agents are not identical. This simple fact creates serious problems for the underpinnings of the standard model.[1] Agents are heterogeneous in many ways, but an important one stems from the fact that any agent in a large population interacts directly with only a very small number of other agents. Thus a potentially important source of agent heterogeneity stems from the "neighbourhood" in which an agent operates. To represent this feature of economic interactions, we can model the population of agents as being located on a network. This gives a natural structure which captures the fact that an agent's direct connections are few relative to the total population, and thereby distinguishes one agent from another.

In this paper we examine the relationship between the architecture of the network of agents and its aggregate performance. The issue here is rooted in the economics of innovation and growth. It is commonplace since the growth accounting work of Solow in the 1950s that technological change is central to economic growth. There are two aspects to technical change: knowledge creation and knowledge diffusion. Early concepts of knowledge treated it as

[1] On this, see Kirman (1992).

a public good, non-rivalrous, and non-exclusive; while expensive to create, it is very cheap to duplicate, and thus trivial to disseminate or diffuse. (See for example Arrow 1962; Nelson 1959.) Recently, though, this view is considered incomplete. While it may nicely capture certain aspects of codified knowledge, there is much knowledge that does not have these properties. In particular, work on tacit knowledge (Cowan and Foray 1997; von Hippel 1998) and absorptive capacity (Cohen and Levintahl 1989), has emphasized that knowledge diffusion is not a trivial activity. Indeed, empirical studies of knowledge flows using patent data have shown that the ease of knowledge flow is negatively related to the distance over which it travels, implying that knowledge is not freely available to the entire population once it has been created. (See the works of Jaffe and his collaborators, e.g. Jaffe et al. 1993.)

The model in this paper captures both of these notions. Knowledge, when created, is not globally available. It is transmitted through face-to-face interactions. Further, because agents are located on a network, the requirement of face-to-face interaction creates a natural notion of distance — the number of interactions needed to pass the knowledge from originator to final recipient. The central issue in this paper is the relationship between network structure or architecture and the ability of the system to create and diffuse knowledge rapidly. We model knowledge creation and diffusion taking place within a population of agents located on a network, and examine the growth of aggregate knowledge levels. Diffusion of knowledge clearly increases aggregate knowledge levels simply through agents acquiring some existing knowledge. But recent work on innovation as the recombination of existing ideas suggests another benefit from diffusion. As an agent receives knowledge or information he is able to integrate it with his existing stock, and create new knowledge. Knowledge diffuses by an agent broadcasting (or perhaps more accurately narrow-casting) his knowledge to those he is directly connected with. Rebroadcasting diffuses the knowledge throughout the economy. In this framework one agent's mistakes or discoveries will benefit those with whom he interacts, and innovations take place as a result of this broadcasting as recipients re-combine the new knowledge with their existing knowledge. But knowledge diffused this way can only be beneficial to those agents who are at least partly capable of understanding and integrating it. Thus there is a threshold value for dissimilarity in agents' knowledge levels below which no transmission is possible — 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 parameter we use to characterize network architecture is the degree of spatial regularity in the inter-agent connections through which knowledge flows. At one extreme of the space of networks there is a regular structure in which every agent is connected to his n nearest neighbours, whereas at the other extreme we have an irregular network in which each agent is connected to, on average, n agents located at random in the network. We are dealing with networks that are identically local in the sense that their density is low and constant, but they have different structural features which, we will see, influence collective properties in a non-trivial way. Our interest is in the space of networks between these two extremes, and we find that one region stands out: "small world" networks generate the fastest knowledge growth. As defined formally by Watts and Strogatz (1998), the small world combines a strong degree of local cohesiveness with a small fraction of long distance links permitting knowledge to be circulated rapidly among distant parts of the network.

2 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 represented by a real-valued vector. At random times an agent broadcasts his knowledge.[2] Knowledge is broadcast to all those agents with whom the sender has a direct connection. If the knowledge level of the potential recipient is not too dissimilar from that of the broadcaster, knowledge is received and assimilated. Formally, if i broadcasts to j, then in any knowledge category in which i exceeds j, j's knowledge increases. In some cases, agents have imperfect absorptive capacity, and only assimilate part of what they receive. In others, they are "super-absorptive": they absorb everything they are sent and are able to use it to create new knowledge, increasing their knowledge levels further.

2.1 Knowledge interaction

Each agent is characterized by a knowledge vector v_i which evolves over time as the agent receives information broadcast by other agents. Formally, let $v_{i,k}^t$ denote agent i's knowledge endowment in category $k \in \{1, \ldots, K\}$ 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

$$\Delta(i,j) = \max\left\{r, \frac{1}{r}\right\} - 1, \tag{1}$$

where $r = |v_i| / |v_j|$ and $|\cdot|$ is the standard Euclidean norm of a vector. For each agent $j \in \Gamma(i)$, provided $\Delta(i,j) < \theta \in (0, \infty)$, i makes his knowledge available to j.[3] For every knowledge category, k, when i broadcasts, j's

[2] The nature of knowledge, and its relation to data and information is complex. we avoid these complexities except to the extent that we recognize the importance of and difficulty in transmitting tacit knowledge.

[3] As an illustration, if $|v_j| = 1$ and $\theta = 0.2$ then agent j can learn from agent i provided $|v_i| \in [0.833, 1.2]$.

knowledge increases according to

$$v_{j,k}^{t+1} = v_{j,k}^{t} + \alpha \cdot \max\left\{0, v_{i,k}^{t} - v_{j,k}^{t}\right\}, \tag{2}$$

for all k without any consequent loss of knowledge to agent i. The parameter α captures an important aspect of knowledge diffusion and transfer. When $\alpha < 1$ knowledge is only partly assimilable. This notion has been examined as an issue of absorptive capacity by Cohen and Levinthal (1989), or Cowan and Foray (1997). Hence 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" (Allen 1983), 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.[4]

2.2 The interaction structure

Consider N agents existing on an undirected connected graph $G(I, \Gamma)$, 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 following heuristic (or 're-wiring' procedure) is employed: 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.[5] For large graphs, this procedure ensures that the

[4] 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 constraint performs the same function. Agents with similar codified knowledge are likely to have similar, if not the same, tacit knowledge. See Cowan et al. (2000) for a discussion of codification and tacitness.

[5] This is the re-wiring procedure employed by Watts and Strogatz (1998), in their seminal work on small worlds.

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 a single parameter $p \in [0,1]$, hence the label $G(I,n,p)$ for graphs in this family.

The structural properties of $G(I,n,p)$-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 friends of each other. Recall $d(i,j)$ is the length of the shortest path between i and j. The average path length $L(p)$ is

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

and average cliquishness $C(p)$ is given by

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

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 1, for a graph of $N = 500$ vertices, each vertex having on average $n = 10$ connections. (The graph depicts average values over 50 replications.) For the sake of clarity, we plot the normalized values $L(p)/L(0)$ and $C(p)/C(0)$.

The upper curve (thick black) in Figure 1 is the normalized average cliquishness index $C(p)/C(0)$ 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 (thin black) as measured by $L(p)/L(0)$ 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 $L(p) \simeq L(1)$ yet $C(p) \gg C(1)$. This interval, in which high cliquishness and low path length coexist, constitutes the small world region.

3 Numerical experiments

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 $\mu_i^t = \frac{1}{K} \sum_k v_{i,k}^t$. The average level of knowledge in the economy at time t is

$$\mu^t = \frac{1}{N} \sum_{i \in I} \mu_i^t \tag{5}$$

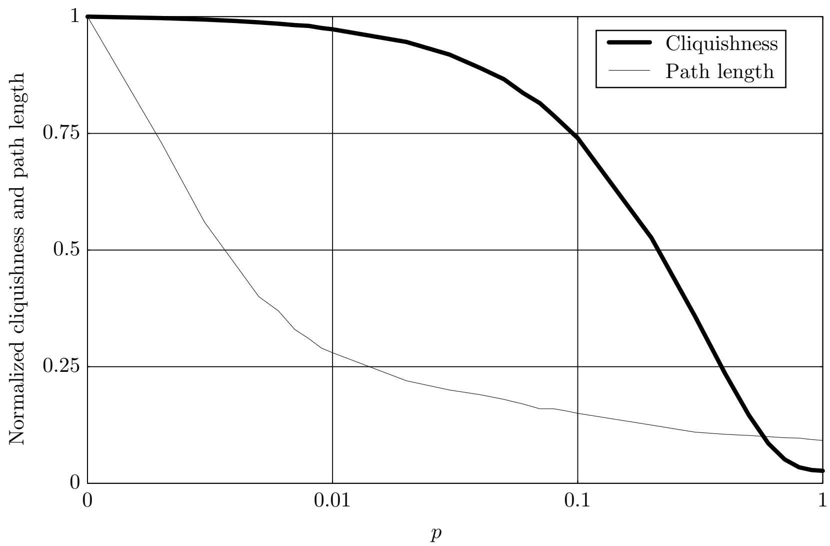

Fig. 1. Average cliquishness and average path length as functions of the degree of randomness in network structure, p

and the variance in knowledge allocation is

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

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 (McGrew and Monroe 1993). Dropping time indexes, the coefficient is written

$$S=\frac{1}{\sigma^2}\sum_{i\in I}\sum_{j\neq i}w_{i,j}(\mu_i-\mu)(\mu_j-\mu), \tag{7}$$

where

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

This statistic is written in terms of the type of cross-products found in the standard correlation, but instead of two variables, the pairs of adjoining vertices are used. Values of the Moran coefficient S tend to be in $[-1,+1]$, though are not restricted to this range. Values near $+1$ indicate similar values tend to cluster; values near -1 indicate dissimilar values tend to cluster; and values near $-1/(\#I - 1)$ (going to 0 as $\#I$ gets large) indicate values tend to be randomly scattered.

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 $G(I, 10, p)$ with $I = \{1, \ldots, 500\}$.[6] Each agent is endowed with a 5-category knowledge vector, with each element randomly initialized from a uniform distribution $U[0, 1]$. In simulation time, in each period one randomly selected agent broadcasts his knowledge. The knowledge is received by all those who are directly connected to him and are sufficiently similar. Formally, the broadcast of agent i is received by the set of agents $\{j \in \Gamma(i) \mid \Delta(i, j) < 0.2\}$. We consider a large number τ of periods, to ensure that each agent takes part in a reasonably large number of knowledge transmissions, both as sender and receiver. With $\tau = 30,000$, the average agent broadcasts 60 times. We run 50 replications for each value of α and p.

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 0.96 to 1.2 with an increment of 0.01. Values of $\alpha < 1$ indicate a regime in which tacit knowledge is important, and knowledge received is only absorbed with some 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 (1983) 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 and equity 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

[6] The issue of the endogeneity of the network structure is not addressed here. This is typical in this literature. Notable exceptions, in which endogenous network strutures are the focus of the analysis, are Bala and Goyal (1998); Jackson and Wolinsky (1996); Plouraboue et al. (1998).

in the economy μ^τ. The parameter space can be partitioned into two distinct regions corresponding to $\alpha < 1$ and $\alpha > 1$. In the first, new 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.

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 (1999), 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. 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 (short) paths along which a piece of knowledge can travel.

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 2 depicts the relationship between the overall knowledge level and the re-wiring probability p and absorptive/innovative capacity α. We plot normalized rather than absolute values, to preserve legibility.[7] For α lower than 1.05, no clear distinction exists between ordered and random graphs. With finer graining, we 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, though it remains confined to the small-world region, i.e. the area between $p = 0.01$

[7] What we plot is actually μ^τ divided, for each α-value, by the maximal value it takes for $p \in [0,1]$, i.e. $\mu^\tau(\alpha, p) / \max_{p'} \mu^\tau(\alpha, p')$.

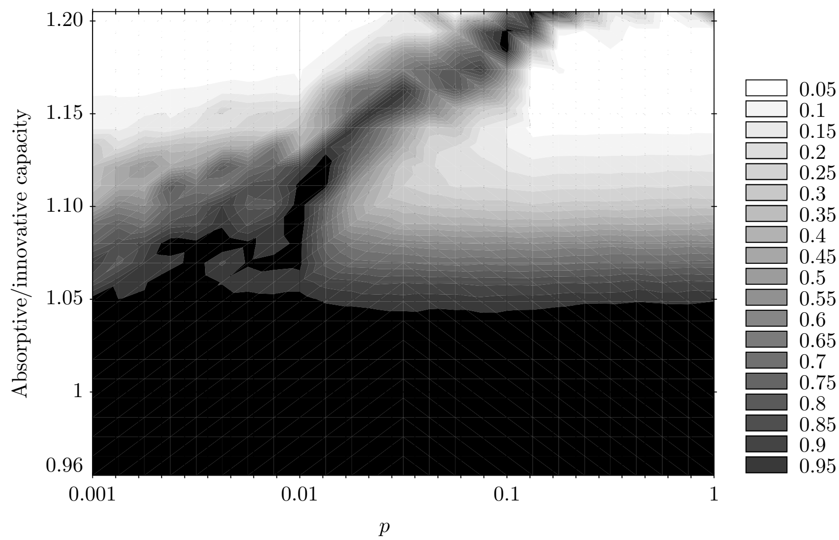

Fig. 2. Knowledge levels in the (p, α)-space

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 in this region. Outside the small-world 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 time a significant amount of local order and a few random connections, achieves the best overall performance in our broadcast economy.

A remark is in order here. In our experiments we ignored values of α greater than $1 + \theta$. Recall that 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 increase so much 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 receive from you. This seems unreasonable, so we have excluded this possibility by restricting α to be less than $1 + \theta$.

The interplay of path length and cliquishness In order to explain the mechanism underlying this result, we focus on 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 analyzed 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. In the long run, this means that 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 is, the higher the limit they jointly achieve also is. So large 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. Figure 3 summarizes three typical outcomes of the process of knowledge broadcast which we comment in turn.

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 a set of neighbouring agents. 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. Strong inequalities arise between groups that advance quickly and groups which lag behind. The first (upper) graph in Figure 3 represents long-run knowledge states in the regular world.

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 because homogeneity is bad for recombination. 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 3, 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 random graph, but relatively homogeneous even at the global level. Homogenization is stronger as soon as shortcuts are in-

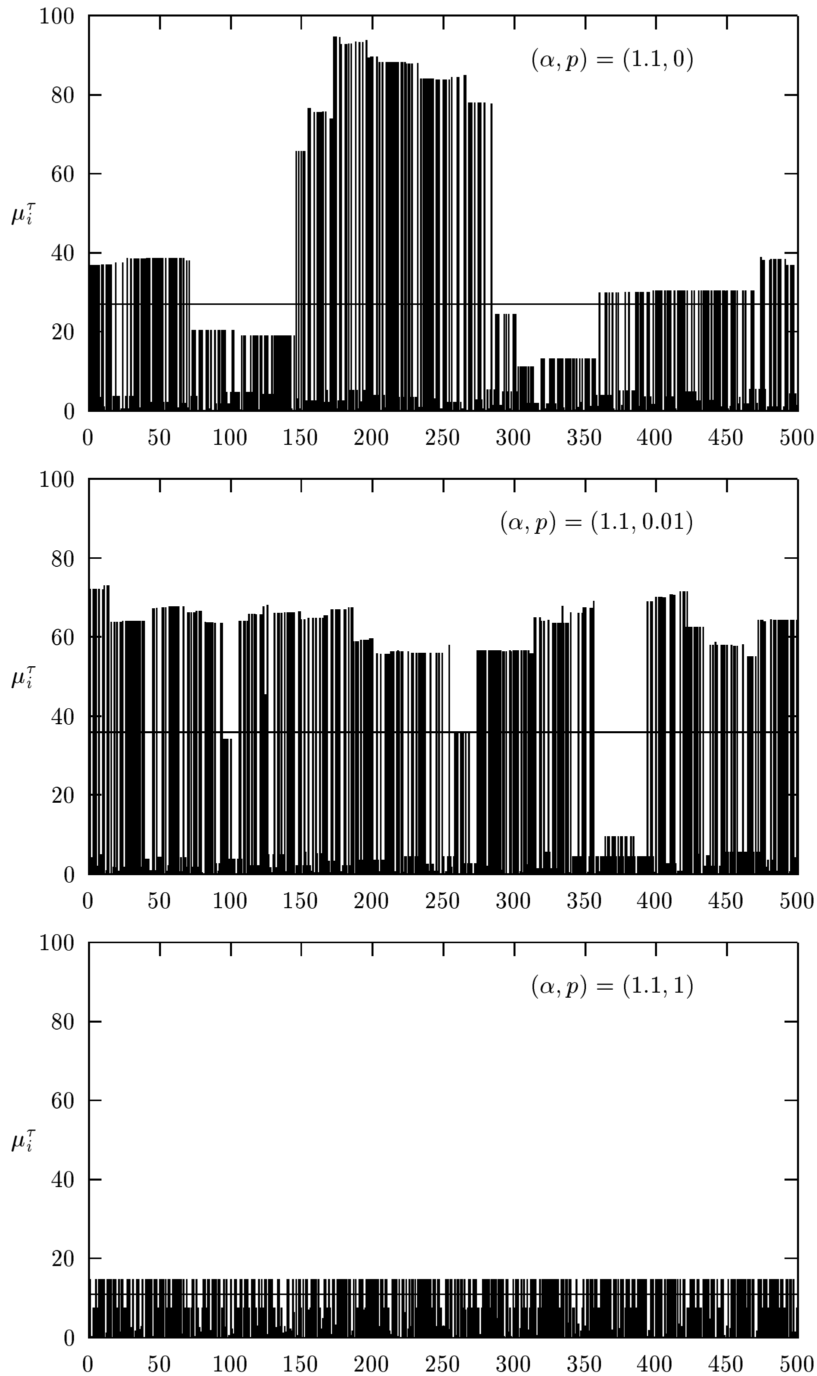

Fig. 3. The average individual long-run knowledge levels in three typical simulations runs for $\alpha = 1.1$ and values of p corresponding to a regular world, a small world and a random graph

troduced, though as long as there are not too many the local reinforcement stemming from large degrees of cliquishness is preserved. This is shown in the middle part of Figure 3. Corresponding average values are represented in Figure 3 by horizontal lines, and we have $\mu^{\tau}(1.1, 0) = 26$, $\mu^{\tau}(1.1, 0.01) = 36$ and $\mu^{\tau}(1.1, 1) = 11$.

A second comment 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. 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.

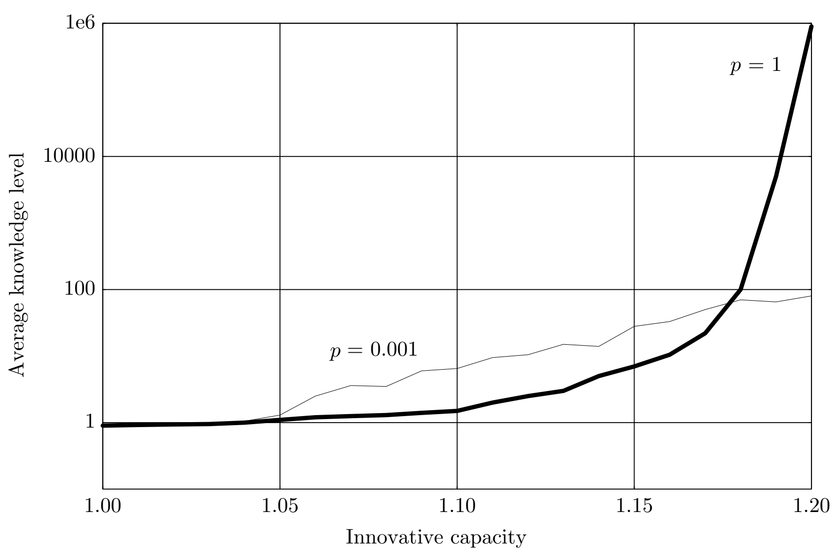

Fig. 4. The population average long-run knowledge level in two polar cases ($p = 0.001$ and $p = 1$), as a function of α

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 **??**.

From this Figure, it is clear that the benefits of cliquishness increase slowly but steadily with α, as evidenced by the thin black curve; the benefits of short paths rises rapidly when α becomes relatively large (thick black curve growing at a more than exponential rate). This suggests that effects of decreasing path length does indeed increase faster than the costs of decreasing cliquishness. Figure **??** shows 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 5 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.

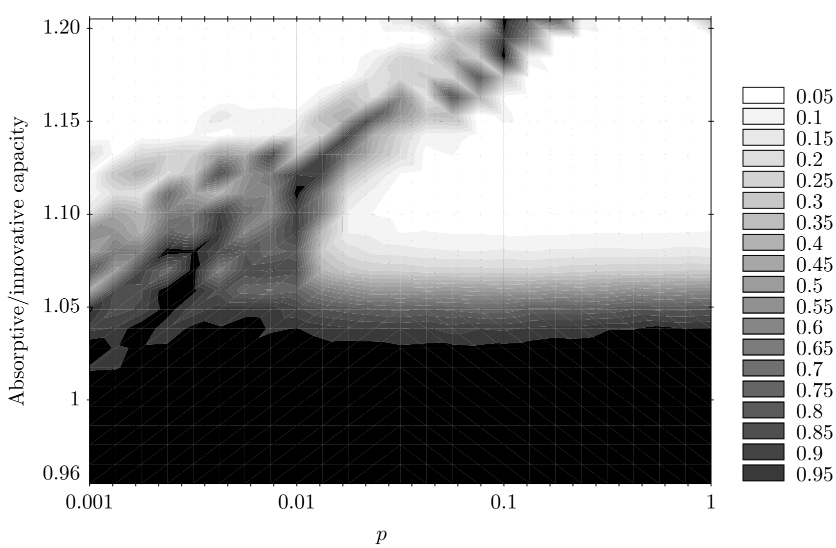

Fig. 5. Long run heterogeneity in knowledge allocation in the (p, α)-space

Surprisingly, Figure 2 and Figure 5 are very similar (though dispersion is lower in Figure 5), 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 $p = 0.1$.[8] We therefore come to the somewhat uncomfortable conclusion that the small world region at the same time generates the best overall performance in terms of how much knowledge is produced by the system, and the worst overall performance when homogeneity of allocation is considered desirable. Figure 3 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 or 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 neighbours of laggards and no clustering of knowledge exists. 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 (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 6 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.[9] On the time scale we consider, though, this relationship between p, α and spatial correlation contains considerable randomness.

Figure 7 gives the Moran coefficient 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

[8] The coefficient of variation of knowledge levels follows the same pattern, indicating that the pattern in the variance is not driven simply by re-scaling.

[9] This was checked using a standard two-tailed t-test.

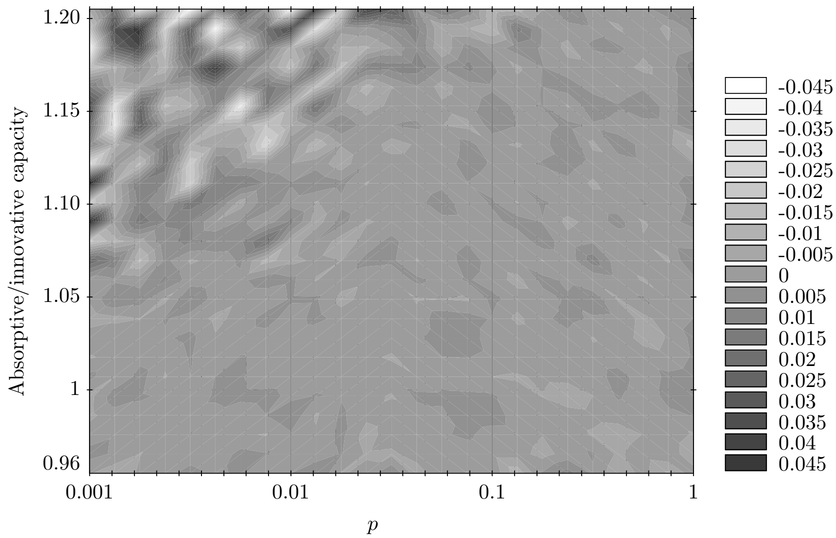

Fig. 6. Spatial correlation in the geographical space

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.

It is clear from the results presented in Figures 6 and 7 and the discussion, that there is no small world effect on spatial correlation or knowledge clustering. Cliquishness is a manifest source of knowledge clustering and introducing shortcuts to possibly distant parts of the graph only tends to impede clustering.

5 Conclusion

The use and creation of knowledge is central to economic growth and development. The fact that knowledge diffusion takes time forces analysts to

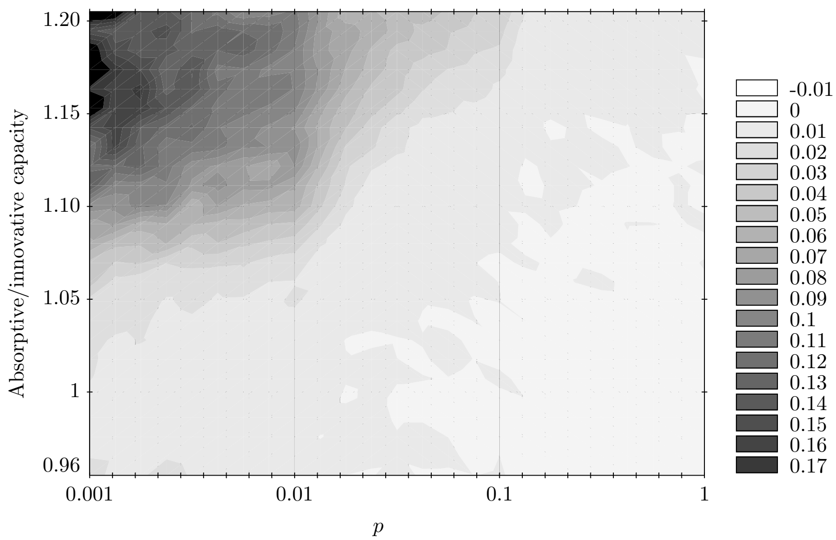

Fig. 7. Spatial correlation in the network space

acknowledge that agents are heterogeneous in economically important ways. Modelling this heterogeneity, and understanding its consequences for economic performance is thus a central task in the attempt to understand any dynamic economy. The model in this paper has explicitly included in the diffusion process the idea that knowledge is not merely information, and thus is diffused through specific contacts between individual agents. Repeated interactions are valuable in facilitating diffusion, and a group of agents can, by collectively sharing knowledge advance together. Groups of agents sharing knowledge more or less freely is something observed historically, and Allen (1983) has coined the phrase "collective invention" to describe it. Common to these 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.

References

1. Allen, R. (1983) Collective Invention, Journal of Economic Behavior and Organization, 4, 1–24
2. Arrow, K. (1962) The Economic Implications of Learning-by-Doing, Review of Economic Studies, 29, 155–173
3. Bala, V., and S. Goyal (1998) Learning from Neighbours, Review of Economic Studies, 65, 595–621
4. Cohen, W., and D. Levinthal (1989) Innovation and Learning: The Two Faces of Research and Development, The Economic Journal, 99, 569–596
5. Cowan, R., P. David, and D. Foray (2000) The Explicit Economics of Knowledge Codification, paper prepared for the European Contract "TIPIK: Technology and Infrastructure Policy in the Knowledge-Based Economy", programme TSER, DG XII (1998/2000)
6. Cowan, R., and D. Foray (1997) Quandaries in the Economics of Dual Technologies and Spillovers from Military to Civilian Research and Development, Research Policy, 24, 851–868
7. Cowan, R., and N. Jonard (1999) Network Structure and the Diffusion of Knowledge, Discussion Paper 99028, MERIT/Maastricht University, Maastricht, The Netherlands
8. Jackson, M. O., and A. Wolinsky (1996) A Strategic Model of Social and Economic Networks, Journal of Economic Theory, 71, 44–74
9. Jaffe, A., M. Tratjenberg, and R. Henderson (1993) Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, Quarterly Journal of Economics, 63, 577–598
10. Kirman, A. P. (1992) Whom and What Does the Representative Individual Represent?, Journal of Economic Perspectives, 6, 117–136
11. McGrew, C., and J. Monroe (1993) An Introduction to Statistical Problem-Solving in Geography. W. C. Brown Publishers.
12. Nelson, R. R. (1959) The Simple Economics of Basic Scientific Research, Journal of Political Economy, 67, 297–306
13. Plouraboue, F., A. Steyer, and J.-B. Zimmermann (1998) Learning Induced Criticality in Consumers Adoption Pattern: A Neural Network Approach, Economics of Innovation and New Technology, 6, 73–90
14. Von Hippel, E. (1998) The Sources of Innovation. Oxford University Press.
15. Watts, D., and S. Strogatz (1998) Collective Dynamics of Small-World Networks, Nature, 393, 400–403