Knowledge flows and the geography of networks: A strategic model of small world formation

Nicolas Carayol*, Pascale Roux

GREThA, Université Bordeaux IV, CNRS, France

ABSTRACT

This paper aims to demonstrate that the strategic approach to link formation can generate networks that share some of the main structural properties of most real social networks. For this purpose, we introduce a spatialized variation of the Connections model [Jackson, M.O., Wolinsky, A., 1996. A strategic model of social and economic networks. Journal of Economic Theory 71, 44–74] to describe the strategic formation of links by agents who balance the benefits of forming links resulting from imperfect knowledge flows against their costs, which increase with geographic distance. We show, for intermediate levels of knowledge transferability, clustering occurs in geographical space and a few agents sustain distant connections. Such networks exhibit the small world property (high clustering and short average relational distances). When the costs of link formation are normally distributed across agents, asymmetric degree distributions are also obtained.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

There is an increasing consensus in the economic literature to recognize that network structures significantly influence the outcomes of many social and economic activities. Such networks are often strategically shaped by the participating agents, as recently highlighted by the theoretical economic literature (Jackson and Wolinsky, 1996; Bala and Goyal, 2000). Models of this type encompass various contexts such as job-contact networks (Calvó-Armengol, 2004), oligopolies and R&D collaborations (Goyal and Moraga, 2001; Goyal and Joshi, 2003), buyer–seller networks (Kranton and Minehart, 2000), and so on. Nevertheless, since the networks obtained exhibit very simple structures (such as empty, star or complete networks), the study of the conditions that may lead to the emergence of more realistic networks remains on the agenda.

What do we know about the structure of real social networks? A series of contributions showed that most of them, though they are complex and heterogeneous in many respects, share some common structural features. Real networks are proven to be very short in the sense that, when counting the minimal number of inter-individual social connections (that is the social distance) between them, the agents are, on average, very close to one another (Milgram, 1967). At the same time, they are highly clustered: the probability that an agent's neighbors are also neighbors to one another is high (Newman, 2001; Kogut

* Corresponding author at: GREThA, Université Bordeaux IV Montesquieu, Avenue Leon Duguit, F-33608 PESSAC, France. Tel.: +33 556848640;
fax: +33 556848647.
  E-mail address: nicolas.carayol@u-bordeaux4.fr (N. Carayol).

and Walker, 2001). In addition to those structural properties, some authors have recently studied the spatial distribution of networks and show that some of them tend to correlate with geography (e.g. Gastner and Newman, 2006). Indeed, earlier studies had already documented the strong inverse relationship between geographical distance and social ties formation (e.g. Bossard, 1932; Zipf, 1946).

How do such complex networks form? Watts and Strogatz (1998) proposed a simple process that allows them to construct networks that share these properties. They begin with a regular one-dimensional lattice network that can be interpreted as a locally connected network (all agents are connected to their four closer neighbors). Then they depart from this network by deleting each link with a small probability and reallocate it to one of the two agents it was attached to and one other agent chosen at random in the population. They show that randomly allocating a few "short cuts" between local clusters is sufficient to minimize the average social distance. The resulting networks, which are both short and highly cliquish, are said to be *small worlds à la* Watts and Strogatz. Nevertheless, this model does not view network formation as the outcome of the explicit decentralized behavior of agents.[1] As a matter of fact, the authors do not explain why most connections remain local while some are formed at a distance.

This paper first aims to demonstrate that the strategic approach to network formation can lead to the emergence of complex networks that share these properties.[2] Therefore, we propose an explanation of why networks with this type of structure come to be formed. For these purposes, we introduce a strategic model of network formation built on a simple variation of the Connections model of Jackson and Wolinsky. Myopic self-interested agents form costly links so as to benefit from agents with whom they are directly or indirectly connected in a network. The longer the distance in the relational network, the weaker the positive externality. Moreover, the costs of direct connections increase linearly with the geographic distance separating agents in a close way to that chosen by Johnson and Gilles (2000).

This model typically applies to knowledge flows in interpersonal connections between inventors for instance. Indeed, it has been repeatedly asserted that invention activity is far from being the outcome of isolated agents efforts but that of interactive and collective processes (e.g. Allen, 1983; Von Hippel, 1989) in which networks of interpersonal relations play a key role by improving and facilitating information and knowledge transfers. Very recently, empirical evidence has shown that inter-individual connections are the support of knowledge spillovers and that knowledge flow decreases sharply with social distance between individuals (e.g. Singh, 2005; Breschi and Lissoni, 2007). In the model, the decay parameter can be interpreted as the quality of knowledge transmission through interpersonal relations. In this context, a straightforward issue relates to how the relative degree of tacitness vs. codification generates qualitatively different network architectures. Thus, whereas the theoretical literature has mainly focused on how some fixed network structures affect information, knowledge or technology diffusion (e.g. David and Foray, 1994; Valente, 1996; Young, 2002; Cowan and Jonard, 2003), our model highlights that the structure of networks depends on the nature of knowledge.

We study the structural attributes of efficient and pairwise stable networks according to the decay parameter (knowledge transferability). A unique upper bound condition is provided, which shows that the more intermediate the value of this parameter, the denser the local connectivity of efficient and pairwise stable networks since the returns to direct connections are higher than the returns to indirect connections. When the population is not too small, this result applies for a large range of the value of the decay. Though agents have disincentives to connect directly with agents they already (indirectly) benefit from, clustering still occurs at the local level because local relations are much less costly. Nevertheless, at this stage, we still do not know if distant connections are formed in the long term equilibrium and to what extent.

Therefore, we propose a dynamic time-inhomogeneous process of network formation, which allows us to examine numerically the structure of emergent networks. We show that indeed, for a large region of intermediate levels of knowledge transferability, the stochastic process leads to emergent networks that share the Small World property: they are highly clustered and exhibit a low average distance thanks to the existence of some distant connections. Such distant connections are also "weak ties" (Granovetter, 1973) or "bridging links" (Burt, 1992) since they are formed between separated and distant communities. This also explains why such costly distant connections are formed: they allow the agents who establish them to gain access and then benefit from relationships with a set of socially distant agents. Moreover, when a distant connection is formed, it dissipates the incentives to form more of these relations. This is the reason why most connections are local (implying high clustering) and some of the short cut relations are endogenously formed, which makes the network short. Thus, the model offers a strategic foundation for the formation of social networks that share the small world property often observed in empirical studies. Lastly, as an extension, we introduce a normal (Gaussian) heterogeneity among agents in the costs of direct connections. Emergent networks are then also characterized by asymmetric distributions of neighborhood sizes: a few agents have many links while most others are weakly connected, an additional characteristic they have in common with real social networks (Albert and Barabási, 1999).

The paper is organized as follows. Section 2 presents basic formal definitions. Section 3 is devoted to the static features of the model and to the presentation of some analytical results on pairwise stability and efficiency. In Section 4 we introduce

---

[1] This statement applies also to many other models that also incorporate, in various extents, randomness and ad-hoc settings (Erdös and Rényi, 1960; Watts and Strogatz, 1998; deSolla Price, 1965; Barabási and Albert, 1999). Survey papers on this literature are Strogatz (2001) and Albert and Barabási (2002).

[2] This aim is shared by Carayol and Roux (2004), Jackson and Rogers (2005) and Galeotti et al. (2006).

the dynamic stochastic process and the methodology we use. The results obtained for the dynamics are presented in Section 5. The last section concludes.

2. Basics on graphs and network formation

We begin with some basic notions on graphs and next provide formal definitions of network stability and efficiency.

2.1. Graphs

Consider a finite set of n agents, $N = \{1, 2, \ldots, n\}$ with $n \geq 3$, and let i and j be two members of this set. Agents are represented by the nodes of a non-directed graph the edges of which represent the links between them. The graph constitutes the relational network between the agents. A link between two distinct agents i and $j \in N$ is denoted ij. A graph g is a list of non-ordered pairs of connected and distinct agents. Formally, $\{ij\} \in g$ means that ij exists in g. The complete graph $g^N = \{ij|i, j \in N\}$ is the set of all subsets of N of size 2, where each agent is connected with all others. Let $g \subseteq g^N$ be an arbitrary collection of links on N. We define $G = \{g \subseteq g^N\}$ as the finite set of all possible graphs between the n agents.

Let $g' = g + ij = g \cup \{ij\}$ and $g'' = g - ij = g \setminus \{ij\}$ be the graph obtained by adding ij and the one obtained by deleting ij from the existing graph g respectively. Graphs g and g', as well as the graphs g and g'', are said to be *a*djacent. For any g, we define $N(g) = \{i|\exists j : ij \in g\}$, the set of agents who have at least one link in the network g. We also define $N_i(g)$ as the set of i's neighbors, that is, $N_i(g) = \{j|ij \in g\}$. The cardinal of that set $\eta_i(g) = \#N_i(g)$ is called the *degree* of node i. The total number of links in the graph g is $\eta(g) = \#g$.

A *path* in a non-empty graph $g \in G$ connecting i to j is a sequence of edges between distinct agents such that $\{i_1 i_2, i_2 i_3, \ldots, i_{k-1} i_k\} \subset g$ where $i_1 = i$, $i_k = j$. A cycle is a path such that $i_1 = i_k$. A network is acyclic if it does not contain any cycles. Otherwise it is cyclic. The length of a path is the number of edges it contains. Let $i \leftrightarrow_g j$ be the set of paths connecting i and j on graph g. The set of *shortest paths* between i and j on g noted $i \widetilde{\leftrightarrow}_g j$ is such that $\forall k \in i \widetilde{\leftrightarrow}_g j$, then $k \in i \leftrightarrow_g j$ and $\#k = \min_{h \in i \leftrightarrow_g j} \#h$. The *geodesic distance* between two agents i and j is the number of links on the shortest path separating them: $d(i, j) = d_g(i, j) = \#k$, with $k \in i \widetilde{\leftrightarrow}_g j$. When there is no path between i and j, then their geodesic distance is conventionally infinite: $d(i, j) = \infty$. A graph $g \subseteq g^N$ is said to be connected if there exists a path between any two vertices of $N(g)$. It is fully connected if there exists a path between any two vertices of N (it is connected and there is no isolated agents). The subgraph $g' \subset g$ is a connected component of g, if for all $i \in N(g')$ and $j \in N(g')$ with $i \neq j$, there exists a path in g' connecting i and j and if $i \in N(g')$ and $j \notin N(g')$, with $i \neq j$, there is no path in g connecting i and j.

Agents have fixed positions in a given space, representing for example their geographic location. We consider that agents are located on a circle (or a ring) with fixed intervals. They are ordered according to their index such that i is the immediate geographic neighbor of agent $i + 1$ and agent $i - 1$ except in the case of agent 1 and agent n who are also direct geographic neighbors. Let's define the operator $l(i, j)$, which simply counts the minimal number of inter-individual intervals along the ring separating i and j. As a consequence, $l(i, j) = \min\{|i - j|; n - |i - j|\}$. Without loss of generality, we assume that the maximum geographical distance on the circle (the distance between any agent and the most distant other agent, the one located on the other side of the circle) is equal to unity. Therefore, the geographic distance between i and j is simply given by $s_{ij} = l(i, j)\lceil n/2 \rceil^{-1}$, with $\lceil n/2 \rceil$ the smallest integer higher than or equal to $n/2$.[3] Notice that the distance between i and j varies negatively with the number of agents n simply because agents are located on a ring of a fixed unitary dimension with uniform intervals. Thus clearly, the larger the number of agents, the higher the density of the population, and thus the less the size of these intervals.

Several typical graphs can be described. First of all, the *empty graph*, denoted $g^\emptyset$, is such that it does not contain any links. The *ring* g° is a network in which all agents are connected and only connected with their two closest geographic neighbors. A *chain* g^c is a connected subset of the ring ($g^c \subset g^\circ$) so that $\forall i, j \in N(g^c)$, $\#i \widetilde{\leftrightarrow}_{g^c} j = 1$. This definition implies that the ring cannot be a chain since the ring is cyclic: in the ring, there are always two paths between any two agents. Let g^{mc} be a maximally connected chain such that $\#g^{mc} = \#g^\circ - 1$. If g^c is such that $\#g^c \leq \#g^{mc}$, there is always one and only one path between two connected agents i and j (the set $i \leftrightarrow_{g^c} j$ is a singleton). The *double ring* denoted $g^{2\circ}$ is a network such that all agents are only connected with their four closest geographic neighbors. In the *triple ring*, denoted by $g^{3\circ}$, all agents are only connected with their six closest neighbors. Finally, a *star*, denoted $g^\star$, is such that $\#g^\star = n - 1$ and there exists an agent $i \in N$ such that if $jk \in g^\star$, then either $j = i$ or $k = i$. Agent i is called the center of the star. It should be noted that there are n possible stars, since each node can be the center.

Let the (geographically) *covering graph* g^{cg} of a network g be built as follows: for all $i, j \in N$: if $ij \in g$ then $k \subset g^{cg}$, with k a path in $i \widetilde{\leftrightarrow}_{g^\circ} j$. If g is connected, then g^{cg} is a chain or the ring. Otherwise, g^{cg} can either be a chain, a set of disconnected chains, or the ring. Note that $i \widetilde{\leftrightarrow}_{g^\circ} j$ is always a singleton except when $l(i, j) = n/2$. Then, the set $i \widetilde{\leftrightarrow}_{g^\circ} j$ is a couple. If one finds such a link in g, it is assumed that only one of the two paths is retained (i.e. the one that has the most edges in common with the covering graph of $g \setminus \{ij\}$). If this criterion does not make it possible to distinguish the two paths, then one of the two is chosen without any consequence since the network is fully symmetric on both sides of edge ij. Therefore also, all non-empty

---

[3] In the circle metric, the maximum number of interindividual connections is given by $\max_{i,j \in N} l(i, j) = \lceil n/2 \rceil$.

networks have one and only one covering graph. Finally, a network g is characterized by regional overlap if $\exists i \in N(g)$ and $\exists jh \in g$, with $j, h \neq i$, such that $i \in N(g_{-i}^{cg})$, with g_{-i}^{cg} the covering graph of $g_{-i} = g \setminus \{ij|j \in N(g)\}$.

2.2. Network stability and efficiency

We consider a network formation game in which pairs of agents meet and decide to form, maintain, or break, links. The formation of a link requires the consent of both agents but the deletion of the link can result from a unilateral decision. Moreover, agents are myopic: they make their decisions on the basis of the immediate impacts on their current payoffs. Formally, let $\pi_i : \{g|g \subseteq g^N\} \to \mathbb{R}$, the payoffs received by i from his position in the network g, with $\pi_i(\emptyset) = 0$.

Jackson and Wolinsky introduce the notion of *pairwise stability*, which departs from the notion of Nash equilibrium since the process of network formation is both cooperative and non-cooperative. A network is said to be pairwise stable if no incentive exists for any two agents to form a new link or for any agent to break one of his existing links. The formal definition of the pairwise stability notion follows. A network $g \subseteq g^N$ is pairwise stable if (i) for all $ij \in g$, $\pi_i(g) \geq \pi_i(g - ij)$ and $\pi_j(g) \geq \pi_j(g - ij)$, and (ii) for all $ij \notin g$, if $\pi_i(g + ij) > \pi_i(g)$ then $\pi_j(g + ij) < \pi_j(g)$.

As far as network efficiency is concerned, we use the 'strong' notion introduced by Jackson and Wolinsky. It is based on the computation of the *total value* of a graph g given by $\pi(g) = \sum_{i \in N} \pi_i(g)$. A network g is then said to be efficient if it maximizes this sum on the set of all possible graphs $\{g|g \subseteq g^N\}$, that is, $\pi(g) \geq \pi(g')$ for all $g' \subseteq g^N$. It should also be noted that several networks can lead to the same maximal total value. For example, if we consider strictly homogeneous agents, any isomorphic graph of an efficient network is also efficient.

3. The spatialized connections model

In this section, we present our model, which is a simple variation of the *Connections model* introduced by Jackson and Wolinsky, and provide some analytical results obtained on the structure of pairwise stable and efficient networks.

3.1. The model

In this model, agents benefit from knowledge that flows through bilateral relationships. Nevertheless, the communication is not perfect: the positive externality deteriorates with relational distance. Formally, there is a decay parameter representing the quality of knowledge transfer through each bilateral connection. In a given network, agents cannot strategically control the circulation of knowledge. Moreover, agents bear the costs of maintaining direct connections. We let the costs of links increase linearly with the geographic distance between the agents on the circle.[4] Our underlying assumption is that geographically distant research collaborations are as effective as geographically close connections *per se* but require more monitoring and costly interactions.[5]

The net profit received by any agent i is given by the following expression:

$$\pi_i(g) = \sum_{j \in N \setminus i} \delta^{d(i,j)} \omega_{ij} - \sum_{j:ij \in g} c_{ij}, \tag{1}$$

where $d(i, j)$ is the geodesic distance between i and j and c_{ij} the costs borne by i for a direct connection with j. ω_{ij} denotes the "intrinsic value" of individual j's knowledge to individual i. For simplicity, it is assumed that such value is fixed across agents and is normalized to unity: $\forall i \neq j : \omega_{ij} = \omega = 1$. The parameter $\delta \in]0, 1[$ is the decay parameter representing the share of knowledge effectively transmitted through each edge. It may be associated with the characteristics of knowledge: communication quality is likely to decrease with the degree of tacitness of knowledge, while it should increase with the codification of knowledge. The effective positive impact of agent j on i's payoffs ($\delta^{d(i,j)}$) is geometrically decreasing with geodesic distance since δ is less than the unity. Note that if there is no path between i and j, $d(i, j) = \infty$ and then $\delta^{d(i,j)} = 0$. Thus, the first part of the right side of (1) expresses the gross payoffs obtained by i thanks to the knowledge flows he receives through all his direct and indirect connections (assuming no time lag for simplicity). We further assume a one-to-one relation between distance and link costs (in a similar fashion as in Johnson and Gilles) so that the costs of maintaining a direct connection between i and j is simply equal to the geographic distance separating them, that is, $c_{ij} = s_{ij} = l(i,j)\lceil n/2 \rceil^{-1}$.

---

[4] Johnson and Gilles first introduced such costs but consider a linear world. Jackson and Rogers have considered separate islands. Our circular specification allows us to avoid *ex ante* asymmetry between agents. For instance, on a line, the two agents at the far ends are, by assumption, in the periphery while agents in the middle of the line are offered a central position.

[5] This feature stresses a specific impact of geography on network formation that traces back to Stouffer (1940) and Zipf (1946), who, thought they did not explicitly model the (individual) benefits and costs of connections formation, assume and interpret their results as geographical distance affecting positively connection costs.

3.2. Efficiency and stability

The analytical results obtained on network efficiency and stability in the model described above are summed up in the following two propositions. The main results in this section concern the manner in which efficient and pairwise stable networks relate to geography, so we focus our attention on this question.

Proposition 1. *Efficiency*

(i) *If $\delta + \frac{(n-2)}{2}\delta^2 < \lceil n/2 \rceil^{-1}$, the empty network $g^\emptyset$ is the only efficient network.*
(ii) *The value of any non-empty acyclic graph g that exhibits no regional overlap is less than its associated covering graph g^{cg}.*
(iii) *If $\delta - \delta^2 > k \lceil n/2 \rceil^{-1}$, $\forall k \in \mathbb{N}^*$, then the efficient network(s) $\hat{g}$ contain(s) all the links it is possible to form between agents located at a geographic distance less or equal to $k \lceil n/2 \rceil^{-1}$.*

Proofs.. See Appendix A. □

Proposition 2. *Stability*

(i) *If $\delta < \lceil n/2 \rceil^{-1}$, the empty graph is the unique acyclic pairwise stable graph and no network containing a peripheral agent (has only one connection) is pairwise stable. If $\delta = \lceil n/2 \rceil^{-1}$, the empty graph is pairwise stable. If $\delta > \lceil n/2 \rceil^{-1}$, the empty graph $g^\emptyset$ is never pairwise stable.*
(ii) *If $\delta - \delta^2 > k \lceil n/2 \rceil^{-1}$, $\forall k \in \mathbb{N}^*$, then any pairwise stable network g contains all the links that can be formed between agents located at geographic distance less or equal to $k \lceil n/2 \rceil^{-1}$.*
(iii) *The star $g^\star$ and the complete network g^N are never pairwise stable.*

Proofs.. See Appendix A. □

The statements in (iii)-Proposition 1 and in (ii)-Proposition 2 have straightforward implications regarding the existence of cycles in efficient and pairwise stable networks. Indeed, if $\delta - \delta^2 > \lceil n/2 \rceil^{-1}$, then any efficient network and any pairwise stable network contain the ring that itself contains a cycle. They also indicate that the more intermediate δ is (close to 1/2), the greater the need for direct connections from both the social and individual points of view, and thus the more beneficial it is for geographical neighbors to be directly connected. To appreciate the extent to which the cost structure imposes the formation of local connections for a significantly large range of intermediate values of δ, let us consider the following numerical example.

Example 1. Let $n = 20$. Then (iii)-Proposition 1 and (ii)-Proposition 2 imply that any efficient or pairwise stable equilibrium must contain all links that can be formed between agents geographically distant the ones from the others by one interindividual interval when $\delta \in]0.112, 0.887[$. The collection of these links is the ring g°, that is included in any pairwise stable and efficient network g when the condition on δ is verified ($g^\circ \subseteq g$). When $\delta \in]0.276, 0.723[$, any pairwise stable and efficient network g must contain all the links that can be formed between agents geographically distant the ones from the others by one or two interindividual intervals. Thus the double ring, which is the collection of such local links, is included in these networks as a subset ($g^{2\circ} \subseteq g$).

When n increases, the links formation costs decrease, and thus the local connectivity increases. Nevertheless, the upper bound expression in (iii)-Proposition 1 and (ii)-Proposition 2 indicates that the number of locally connected agents is a fixed proportion of the total population. To see this, let us assume, for simplicity, that n is even. The right term of the inequality then equals $2k/n$. It can be interpreted as the fixed fraction of agents to whom any agent is connected and who are found in his closest geographic neighborhood.

Our results are incomplete in that we do not know much about the formation of distant connections, in addition to the local connections when δ is intermediary. The existence of distant connections, their proportion among all links, in relation with their impact on individual payoffs are of special interest to us in the context of the equilibrium. Indeed, the literature on social networks highlights that a few of these connections are formed (Granovetter, 1973; Burt, 1992). It also highlighted that such bridging connections are critical for all agents in that a few of these links enable mostly locally connected networks to reduce significantly the average social distance (Watts and Strogatz, 1998). We next introduce a dynamic approach to the formation of networks, which, applied to our model, will allow us to explore the structural attributes of the networks that are formed for the different values of the decay and in particular to investigate fully the issue of the distant links formation.

4. The dynamic process of network formation

This section is a self-contained presentation of our dynamic model. We first introduce the standard perturbed stochastic process of network formation and then focus on the terms that differentiate our process from the standard one.

4.1. The standard perturbed stochastic process of network formation

Let $g_t \in G$ denote the state of the social network at period t (with $t = 1, 2, \ldots$). At each time period, two agents i and $j \in N$ are randomly selected. If they are directly connected, they can jointly decide to maintain their relation or unilaterally decide to sever the link between them. If they are not connected, they can jointly decide to form a link, or each agent can unilaterally decide against it. Formally, these two situations are the following:

(i) if $ij \in g_t$, the link is maintained if $\pi_i(g_t) \geq \pi_i(g_t - ij)$ and $\pi_j(g_t) \geq \pi_j(g_t - ij)$. Otherwise, the link is deleted.
(ii) if $ij \notin g_t$, a new link is created if $\pi_i(g_t + ij) \geq \pi_i(g_t)$ and $\pi_j(g_t + ij) \geq \pi_j(g_t)$, with a strict inequality for one of them.

The evolution of the system at any time t depends only on the present state of the system given by the graph structure g_t. Thus the stochastic process is Markovian. The evolution of the system $\{g_t, t > 0\}$ can be described by the probability matrix (P) describing the one-step transition probabilities between all possible states of the finite state space G.

Jackson and Watts (2002) introduce small random perturbations $\bar{\varepsilon} > 0$ that invert agents' correct decisions to create, maintain or delete links. These perturbations may be understood as mistakes or as mutations. For low but non-null values of $\bar{\varepsilon}$, it can be shown that the discrete-time Markov chain associated to the transformed transition matrix $P(\bar{\varepsilon})$ is irreducible and aperiodic and thus has a unique corresponding stationary distribution $\mu(\bar{\varepsilon})$ that is the solution of $\mu(\bar{\varepsilon}) \times P(\bar{\varepsilon}) = \mu(\bar{\varepsilon})$. Such perturbed stochastic process is ergodic. Intuitively ergodicity occurs when it is possible to shift directly or indirectly between any two states in a potentially very long period of time. It allows the long run equilibrium ($\mu(\bar{\varepsilon})$) of the system to be unique and independent of the initial conditions.

Usually, the modeler computes $\mu^* = \lim_{\bar{\varepsilon} \to 0} \mu(\bar{\varepsilon})$ and a state g (a network here) is said to be a stochastically stable state (Young, 1993) if it has a non-null probability of occurrence in the limit stationary distribution. Thus the set of stochastically stable states is $G^* = \{g \in G | \mu_g^* > 0\}$. In the network formation context, Jackson and Watts show that stochastically stable networks are either pairwise stable or part of a closed cycle (of the unperturbed process).

4.2. A time-inhomogeneous process of network formation

In practice the precise computation of the stochastically stable networks requires the identification of all the recurrent classes of the unperturbed process (Young, 1998) that, in the network context, are likely to be extremely numerous. To make that point clear, we shall remind the reader that there is a recurrent class for each pairwise stable network and that in models such as the Connections model or the spatialized connections model presented in Section 3, thousands of networks are possibly pairwise stable. Thus, the standard process described above is not well designed for our purpose.

We propose to let the error term decrease in time according to the following simple rule:

$$\varepsilon_t = 1/(t+1) + \bar{\varepsilon}, \tag{2}$$

with $\bar{\varepsilon} > 0$. This rule ensures that a significant noise affects the dynamics at the beginning while it decreases monotonically with time down to a small strictly positive limit: $\lim_{t \to \infty} \varepsilon_t = \bar{\varepsilon}$. According to Robles (1998), the long run equilibrium $\psi(\bar{\varepsilon})$ of such a time-inhomogeneous Markov chain exists, is unique, and is equal to the equilibrium of the Markov chain perturbed by the constant error $\bar{\varepsilon}$: $\psi(\bar{\varepsilon}) = \mu(\bar{\varepsilon})$.[6] It is then obviously ergodic, thus enabling us to examine with good confidence the long run behavior of the system through simulations (Vega-Redondo, 2006). Note that the assumption of the time-inhomogeneity of noise is more satisfactory than that of time-homogeneity: agents are likely to make fewer and fewer errors over time, but a very small error probability persists in the long run. Moreover, the time-inhomogeneity of noise is interesting since it makes numerical experiments more tractable.

We call the networks on which the process stabilizes in the long run *emergent networks*. Formally the set of emergent networks is $\hat{G} = \{g \in G | \psi_g(\bar{\varepsilon}) > 0\}$. This set is broader than the set of stochastically stable networks defined above that is included in $\hat{G}$. That is, for all g such that, if $\lim_{\bar{\varepsilon} \to 0} \mu_g(\bar{\varepsilon}) > 0$, then $\psi_g(\bar{\varepsilon}) > 0$.[7]

5. Results: the structural properties of emergent networks

In this section we study the emergent networks selected by the stochastic process described in the previous section. We use Monte Carlo experiments to generate networks that are on the support of the unique limiting stationary distribution $\psi(\bar{\varepsilon})$ (of networks) of the perturbed dynamic process. The limit error term is $\bar{\varepsilon} = 10^{-4}$. The size of the network is set to $n = 20$ agents because this value of n is sufficiently high to obtain local connectivity for a large space of δ values (cf. Example 1). All experiments are stopped at $T = 20,000$, a period after which the process is proven to have almost certainly

---

[6] See Proposition 3.1 of Robles (p. 211).
[7] This can be easily proved by recalling the Freidlin and Wentzell (1984) theorem that states $\forall g, \psi_g(\varepsilon) = \mu_g(\varepsilon)$ is of the form $\psi_g(\varepsilon) = \upsilon_g(\varepsilon)/\sum_{g'} \upsilon_{g'}(\varepsilon)$ with $\upsilon_g(\varepsilon)$ a polynomial in ε.
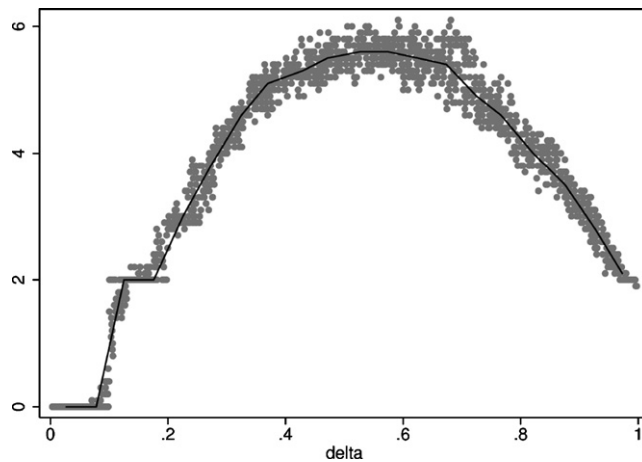
Fig. 1. Average degree $\hat{\eta}(g)$ of emergent networks for 1500 simulations with randomly drawn values of $\delta \in]0; 1[$.

stabilized on a given pairwise stable state.[8] We shall focus on the characterization of the structural properties of emergent networks depending on the decay parameter δ, which is interpreted as the quality of knowledge transfer in the context of our application. For this purpose we ran 1500 experiments performed with values of δ randomly drawn over its value space $]0, 1[$.

5.1. How is density affected by knowledge transferability?

In order to provide a first synthetic characterization of the structural properties of networks, we compute the average neighborhood size (or *average degree* as it is usually referred to in the literature on networks) as follows:

$$\hat{\eta}(g) = 2\eta(g)/n. \tag{3}$$

Let us first examine how the average degree $\hat{\eta}(g)$ is affected by δ. As shown in Fig. 1, the average degree is null when $\delta \leq \lceil n/2 \rceil^{-1} = 0.1$; the process converges to the empty graph that is pairwise stable for such values of the parameters (see Proposition 2). However when δ becomes close to 0.1, some non-empty networks begin to be selected. These networks are somehow in a phase transition between the empty graph and networks for which agents have in average two connections (average degree equal to unity). This structure is mostly selected for $0.1 \leq \delta \leq 0.2$. The average degree of the network then increases up to nearly $\hat{\eta}(g) = 3$ for $\delta \approx 0.55$. From $\delta \approx 0.6$, the average degree decreases down to a value slightly below unity when δ reaches its maximum.

Let us remember that the transferability of knowledge δ can be associated with the characteristics of knowledge: the quality of communication is likely to decrease with the degree of tacitness of knowledge while it should increase with the codification of knowledge. Thus, our model provides a prediction of how the nature of knowledge could affect the density of emergent networks. We find that in the case when knowledge is highly tacit (δ close to 0) and in the case when it is highly codified (δ close to 1), the emergent networks are weakly connected. In both cases, the marginal returns from direct distant and/or local overlapping connections do not make up for their associated costs. When the quality of knowledge transfers is very high, agents do not need to build direct connections since indirect connections are almost as effective. Conversely, when the quality is weak, only cheap local ties are formed due to the low amount of knowledge sourced through direct connections. Thus, in both cases, only local minimally connected networks emerge (when knowledge is extremely tacit, no connection is formed). It is when knowledge is neither highly codified nor highly tacit that such private returns are sufficiently high for the network average degree to increase. Then, the difference between the private returns to direct connections and the returns to indirect ties is the greatest.

5.2. When do small worlds emerge?

Behind the dots of Fig. 1, one can find networks with many different structures. This diversity calls for a statistical analysis of the structural properties of emergent networks. For this purpose, we compute two dedicated indexes. The former is the

---

[8] In average, non-pairwise stable networks are rarely found after 20,000 periods (less than 4%). Let us be precise that emergent networks are not necessarily pairwise stable. They can also be part of a closed cycle in the associated time-homogeneous unperturbed process ($\varepsilon_t = 0$). We also cannot exclude that some non-pairwise stable networks are just emergent networks since $\bar{\varepsilon} > 0$. Nevertheless, in practice, when such a non pairwise stable network was found, it was because an error recently affected the system. Then, we dropped these experiments to ensure that the results are not contingent to recent arbitrary changes. Thus, our approximations of the emergent networks distributions are restricted to the emergent networks that are pairwise stable.

average distance (or average path length) of (directly or indirectly) connected agents. It is given by

$$d(g) = \frac{\sum_i \sum_{j \neq i} d(i,j) \times 1\{i \leftrightarrow_g j \neq \emptyset\}}{\#\{i,j | i \neq j \in N, i \leftrightarrow_g j \neq \emptyset\}}, \tag{4}$$

if $\eta(g) > 0$, with $\#\{\cdot\}$ denoting the cardinal of the set defined into brackets and $1\{\cdot\}$, the indicator function that is equal to unity if the condition in brackets is verified and zero otherwise. This index allows us to appreciate the extent to which directly or indirectly connected agents are "relationally" distant.

The second index is the *average clustering* (or average cliquishness as it is often referred to in Physics). It indicates the extent to which neighborhoods of connected agents overlap. It is given by

$$c(g) = \frac{1}{n} \sum_{i \in N; \eta_i(g) > 1} \frac{\#\{jl \in g | j \neq l \in N_i(g)\}}{\#\{j, l | l \neq j \in N_i(g)\}}. \tag{5}$$

In words, this index measures the propensity with which an agent's neighbors are also neighbors to one another.

The two indicators presented above are affected by the average degree of the network ($\hat{\eta}(g)$), which is likely to vary with δ. Therefore, these indicators are somehow biased, and we must find for an efficient way of controlling for the average degree. In the spirit of Watts and Strogatz, we associate *control random graphs* with emergent networks characterized by the same number of agents and links (thus the same average degree). Such random networks are simply built by allocating a given number of edges to randomly chosen pairs of agents (Erdös and Rényi). For each given number of edges of emergent networks, the average distance and the average clustering are numerically computed and averaged over 1000 such random graphs. Thus, instead of looking at $c(g)$, where g is an emergent network, we compute the ratio $c(g)/c(g^{rd})$, where $c(g^{rd})$ denotes the mean average clustering of the 1000 random networks that have exactly the same average degree as g. Similarly we compute $d(g)/d(g^{rd})$. These ratios are plotted in Fig. 2.

Such ratios can be used to identify a specific but frequently observed network structure called a *small world* (by Watts and Strogatz). It is characterized by the two following properties:

$$c(g)/c(g^{rd}) \gg 1 \quad \text{and} \quad d(g)/d(g^{rd}) \approx 1. \tag{6}$$

Small world networks are highly clustered as compared to random graphs, and simultaneously their average distance is close to that of random graphs, which are known to have very short average path length. We observe that the average distance of emergent networks becomes close to unity when δ reaches 0.35 and then retains this value until $\delta \approx 0.9$. The average clustering ratio also decreases quite sharply with $\delta > 0.2$. Nevertheless, the clustering of emergent networks remains significantly higher than their corresponding random networks, at least until $\delta \leq 0.7$. Therefore, we conclude that small world configurations are selected for the whole region characterized by $\delta \in [0.35, 0.7]$.

5.3. The spatial structure of emergent networks

We shall now address the question of the typical structures of emergent networks. Fig. 3 provides some intuitions about typical networks shapes obtained for several values of the parameter δ. When $0.09 \lesssim \delta \leq 0.2$, it is the ring g° that emerges most often (characterized by an average degree equal to one in Fig. 1 and a null average clustering in Fig. 2). Thus agents not only have two neighbors in average, but they are all connected to their two closest geographic neighbors. When δ is equal
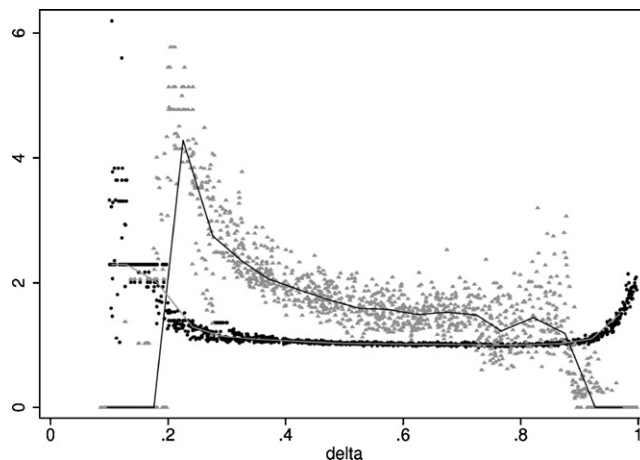

Fig. 2. Average distance (black circles: $d(g)/d(g^{rd})$), average clustering (grey triangles: $c(g)/c(g^{rd})$) and their fitted values of emergent networks for 1500 simulations with randomly drawn values of $\delta \in]0, 1[$.
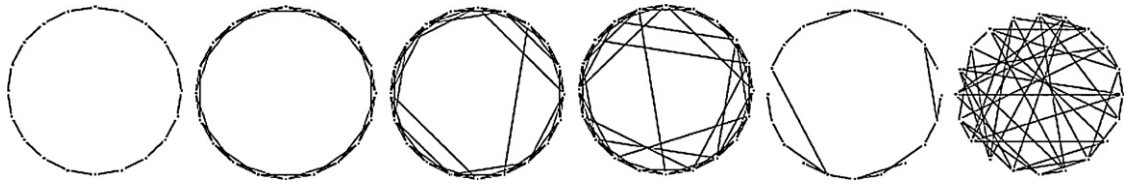
Fig. 3. Typical emergent networks obtained with $\delta = 0.15$, $\delta = 0.3$, $\delta = 0.35$, $\delta = 0.7$, $\delta = 0.98$. The last network has been generated with the standard connections model of Jackson and Wolinsky with $\delta = 0.7$ and a connection cost equal to 0.5 (A (fixed) link cost in the standard connections model set to 0.5 is approximately equal to the average (potential) link cost borne in the spatialized connections model with 20 agents).

to 0.3, networks in which all agents are connected to their four closest geographic neighbors are likely to emerge. Such a situation corresponds to the double geographic ring $g^{2\circ}$. These results are fully consistent with the upper bound condition in (ii)-Proposition 2 for $k = 1, 2$ on pairwise stability, numerically computed for $n = 20$ in Example 1. At the other far end of the spectrum, emergent networks tend to become maximal chains (g^{mc}): when δ approaches unity, direct and indirect connections are likely to provide the same wealth, and thus overlapping connections become redundant. Between both these extremes, when $0.35 \leq \delta \leq 0.7$, we find structurally distinguishable configurations characterized by the conjunction of (i) a prevalence of local connections, and (ii) the existence of some short cut connections. The specificity appears clearly when one compares these structures with the typical networks obtained with the standard connections model.

In order to provide a systematic analysis of the correlation of social connections with the ring-like geography, we propose to study $p(\cdot)$, the density distribution of direct connections according to the number of inter-individual intervals between two linked agents. It is formally defined as

$$p(h) = \frac{1}{\eta(g)} \sum_{ij \in g} 1\{l(i,j) = h\}, \tag{7}$$

for all $h = 1, \ldots, \lfloor n/2 \rfloor$, with $\lfloor \cdot \rfloor$ meaning "the highest integer smaller than or equal to".

In order to explore the distribution in cases when the emergent networks have the small world structure, we perform 100 additional numerical experiments for each of the following values of $\delta : \delta = 0.35, 0.5, 0.7$. Averaged distributions for each value of δ are presented in Fig. 4. Unsurprisingly, networks do correlate with the geographic metric. Indeed, between 70% and 80% of links connect two agents distant of at most two inter-individual intervals. Clustering is thus achieved in local space. Nevertheless, the networks also exhibit distant connections that make the distribution $p(\cdot)$ long tailed on the upper side of the spectrum. The originality of these networks clearly relies on the fact that they are both correlated with space and have some uncorrelated connections. The result provides the missing information mentioned in Section 3: when δ is intermediate, not only are the networks fully locally connected (and thus clustered in space), but some distant connections are also formed.

Overall, the propensity to connect decreases with geographical distance when it is greater or equal to three. Notice that there is a specific reduced probability that connections are made at four inter-individual intervals, particularly when δ equals 0.5 or 0.7. Agents do not find it profitable to establish such medium distance connections because they already benefit from agents at distance four with one intermediary (since the double ring is there). They prefer to connect with agents located further away to reach beyond the local cluster from which they already benefit.

This result supports the idea that sustainable short cuts are not made at a random distance (and cost), as in Watts and Strogatz, but are driven by some strategic behavior. Since distant connections are necessary to obtain the small world property
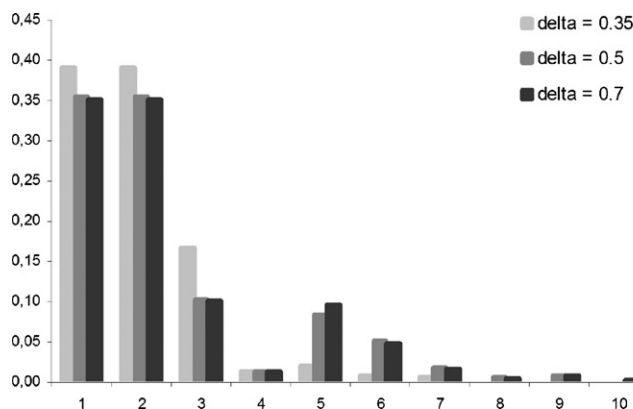

Fig. 4. Distribution of links $p(\cdot)$, according to the number of inter-individual intervals $l(i,j)$ between the two directly connected agents (averaged over 100 experiments for each value of $\delta = 0.35, 0.5, 0.7$).
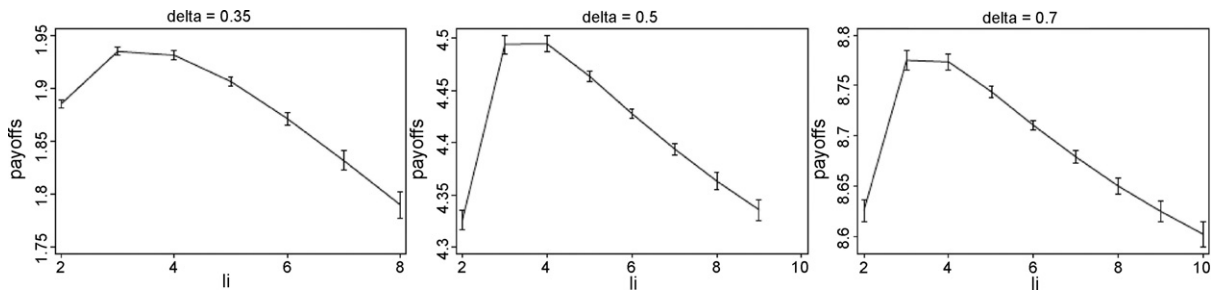
Fig. 5. Fractional polynomial estimates (with 90% confidence interval) of individual payoffs π_i (in ordinates) in relation to l_i (in abscises). 2,000 agents are considered for each of the three values of δ : 0.35 (left), 0.5 (middle), 0.7 (right).

and since these distant connections are also very costly, a naturally arising question is why some agents bear such costs? In other words: does it pay to sustain shortcuts?

In order to answer this question, we record for each of the experiments computed for $\delta \in \{0.35, 0.5, 0.7\}$ the individual payoffs (π_i) and l_i, defined as the longest connection of i measured by counting the number of inter-individual intervals: $l_i \equiv \max_{j \in N_i(g)} l(i, j)$. Two-way plotted results are presented in Fig. 5. The latter shows that individual payoffs decrease on average when agents sustain a connection that is longer or equal to four intervals. Therefore we can conclude that it does not pay to sustain short cuts in the long run,[9] especially when such connections are very long distance ones. Nevertheless, for any shortcut to exist, it has to remain advantageous to sustain it: payoffs must increase with rather than without it. As a matter of fact, the agents who sustain a distant connection are locked in, in that they are better off with this connection although they cannot appropriate the full social value it generates, which is to a large extent captured by their local neighbors. If the latter (sustaining no distant connection) have higher payoffs, it is clearly because they indirectly benefit from these connections while they do not bear their high associated costs: they free ride.

5.4. The heterogeneity of agents and degree distribution

Small worlds *à la* Watts and Strogatz are characterized by a low average path length and a high average clustering. Another generic attribute of real social networks is that their degree distribution is usually right asymmetric. In short, this means that many agents have only a few connections and that very few agents have many connections. In the model presented in Eq. (1) agents have no incentive to form many connections. Rather they have incentives to be connected to a "star". This is why we naturally do not find some densely connected agents among emergent networks as observed in most real networks. Therefore, we now explore whether slight modifications of the payoff function might increase the disparity in neighborhood sizes. A natural question rises: is a normal heterogeneity of agents sufficient to generate a right asymmetry in the degree distribution, or is it necessary to assume the distribution of agents abilities to be also asymmetric?

Let us consider that the costs agents bear for link formation are heterogeneous. The payoff function (1) thus becomes

$$\pi_i(g_t) = \sum_{j \in N \setminus i} \delta^{d(i,j)} \omega_{ij} - a_i \sum_{j: ij \in g_t} c_{ij}, \tag{8}$$

where still $\omega_{ij} = \omega = 1$ and with $a_i > 0$, its mean $\langle a_i \rangle$ equalling 1 to keep consistency with (1). This implies that the costs borne by two agents to be linked together ($a_i c_{ij}, a_j c_{ij}$) may now differ. Moreover, we assume that the a_i are distributed in a Gaussian fashion, numerically given in Fig. 6. The a_i are randomly allocated to agents independently of geography.

Let $\rho(k)$ denote the degree distribution. It is defined as

$$\rho(k) = \frac{1}{n} \sum_{i \in N(g)} 1\{\eta_i(g) = k\}, \tag{9}$$

for all $k = 0, \ldots, n - 1$.

As in the previous subsection, we perform, with the heterogeneous agents model (the distribution of the a_i is presented in Fig. 6), 100 additional numerical experiments for each of the following values of δ :$\delta = 0.35, 0.5, 0.7$. For each of these values, the averaged degree distributions are computed and plotted in Fig. 7, which also presents the averaged degree distributions for the homogeneous agents model ($a_i = 1$).

We find that while most agents have 4, 5 or 6 connections in the homogeneous model,[10] the degree distribution is indeed much more dispersed when agents are heterogeneous. Unsurprisingly, the agents who bear lower costs are much more

---

[9] A simple econometric estimation confirms the inverse-U relation between l_i and payoffs when we control for network density, which explains much of the variance observed in Fig. 5. Due to space constraints, we did not report it here.

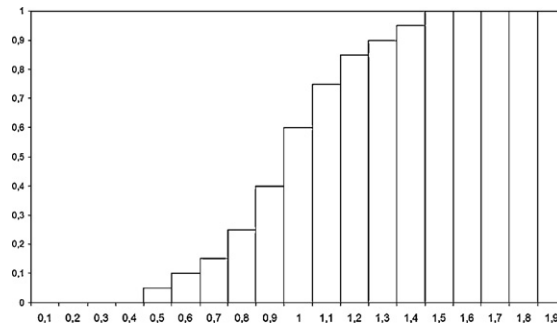[10] Bimodality on 4 and 6 neighbors is found when $\delta = 0.35$.

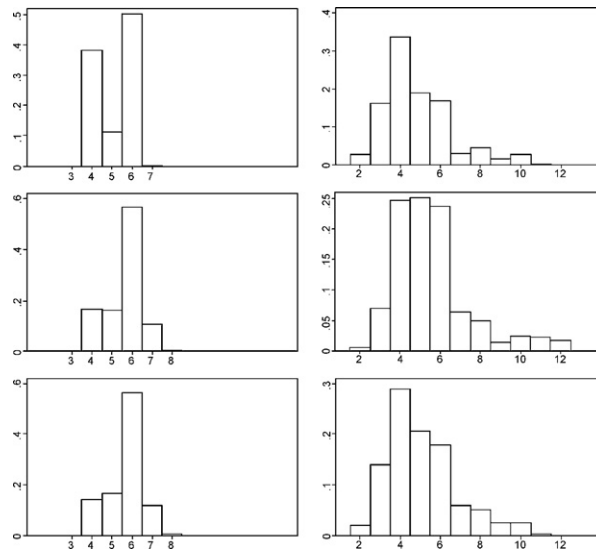**Fig. 6.** Distribution of a_i among agents.

Fig. 7. Degree distribution $\rho(k)$ averaged over 100 experiments for each value of δ : 0.35 (top), 0.5 (middle), 0.7 (bottom), for the two models described in (1)(homogeneous agents, left graphs) and (8)(heterogeneous agents, right graphs).

inclined to form numerous connections. Moreover, the Gaussian distribution of costs across agents generates an asymmetric and long tailed degree distribution. Though we consider a limited number of agents, this result is consistent with the skewed degree distributions often observed in most social networks. Moreover, it should be noted that unreported experiments (due to space constraints) performed with the heterogeneous agents model and random values of $\delta \in \,]0;1[$ show that the small world property is fully preserved for $0.35 \le \delta \le 0.7$.[11] Therefore, all three characteristics (short average path length, high average clustering and skewed degree) are then simultaneously observed on emergent networks.

6. Conclusion

In this paper, we introduced a model of network formation in which agents, arranged on a circle, myopically decide to establish, maintain or sever links. For this purpose, they balance the benefits of obtaining knowledge that flows imperfectly through bilateral relationships (positive externalities deteriorate with relational distance) against the costs of maintaining direct connections, which linearly increase with geographic distance. We also proposed an original time-inhomogeneous stochastic process to study the structural properties of networks that emerge in the long run.

We showed that this simple model of strategic network formation, which is a spatialized variation of the Connections model developed by Jackson and Wolinsky, leads to emergent networks that share the main structural properties of most real social networks. Indeed, for a large region of intermediate values of the parameter that accounts for the decay of externalities along connections, we found emergent networks that have the small world configuration, characterized by both a high average clustering and a short average social distance.

---

[11] On the contrary, it tends to increase slightly the range of relevant values of delta for which a small world à la Watts and Strogatz is found. Numerical results are available from the authors on request.

Clustering occurs through a high local connectivity thanks to the lower costs. This result is consistent with the empirical literature on the suggested application to knowledge networks, which shows that most links are established in the local space (Saxenian, 1994; Almeida and Kogut, 1999; Singh, 2005; Breschi and Lissoni, 2007), generating, in turn, a higher level of local knowledge diffusion (Jaffé et al., 1993). It is also when the decay has intermediate values, which can be interpreted as the knowledge being neither highly tacit nor highly codified, that networks are likely to be more dense, a hypothesis that needs testing empirically.

The short average social distance is obtained thanks to the formation of some distant connections. As stressed in the empirical literature, costly bridging connections are formed to gain significant returns from accessing some other local communities: bridging links are associated with a higher probability of finding a job (Granovetter, 1974), a higher rate of good ideas generation for Burt (2004), and obtaining faster promotions for Burt (1997). Nevertheless, the agents who sustain such connections are locked in: their net returns are lower than their neighbors who free ride on their distant connections. This also explains why such links are strategically formed with parsimony: the neighbors of agents sustaining shortcuts have no incentive to duplicate distant connections with communities of agents they already benefit from indirectly.

We further studied, for this region, the effects of introducing agents' heterogeneity in terms of their connection costs (distributed in a Gaussian manner). We found that emergent networks are, in addition, characterized by an asymmetric distribution of neighborhood sizes. Therefore, we conclude that the three characteristics of a short average distance, a high average clustering and an asymmetric degree distribution are simultaneously gathered for a large and intermediate region of the decay parameter. Such properties have been repeatedly found as structural properties of real networks, in particular for networks of collaboration for scientific knowledge production (Newman, 2001; Goyal et al., 2006).

Acknowledgements

Appendix A

A.1. Proof of Proposition 1

Part i) The proof proceeds as follows. We give an upper bound expression for connected networks value of $k \geq n - 1$ links. We show that this expression is at its maximum when $k = n - 1$ links. We then derive the condition under which this expression (with $k = n - 1$) is negative. Lastly, we demonstrate that the value of any network having $k < n - 1$ links is negative if no more connected adjacent network has a positive value. Since the value of the empty graph is zero, it is certainly the efficient network under the given condition (for which all networks have negative value).

An upper bound value of a network of $k \geq n - 1$ links may be given by

$$\pi^{\max}(g|\#g = k) = 2k\delta + [n(n-1) - 2k]\delta^2 - 2k\lceil n/2 \rceil^{-1}.$$

This expression considers that all the agents who are not directly connected benefit from each other just as they were at relational distance 2. It also assumes that bonds costs are minimal (as if all links were connecting immediate geographic neighbors).

Let us now see how the upper bound expression of the network value behaves when k varies. Adding a link, from k to $k + 1$ links, the upper bound value varies as follows:

$$\pi^{\max}(g|\#g = k+1) - \pi^{\max}(g|\#g = k) = 2\delta - 2\delta^2 - 2\lceil n/2 \rceil^{-1},$$

which is independent of k and strictly negative when $\delta - \delta^2 < \lceil n/2 \rceil^{-1}$. Thus, given that $k \geq n - 1$, the upper bound value is maximal when $k = n - 1$. Now assuming that $k = n - 1$, the upper bound network value becomes

$$\pi^{\max}(g|\#g = n-1) = 2(n-1)\left(\delta + \left(\frac{n}{2} - 1\right)\delta^2 - \lceil n/2 \rceil^{-1}\right). \tag{A.1}$$

It is negative when

$$\delta + \frac{(n-2)}{2}\delta^2 < \lceil n/2 \rceil^{-1}. \tag{A.2}$$

Now let us assume that there is no network with $n - 1$ links that has a positive value. This occurs for instance when condition (A.2) applies. Then, no network with $n - 2$ links can have a positive value. To see this, let us simply consider that any one of these networks is necessarily adjacent to at least one network with $n - 1$ links that can be built by adding an

edge of (geographical) distance equal to one interval and would connect to a component an agent who would otherwise remain alone. The marginal social value of this bond is greater than any of the other links of the component while it can be less or equal but not more costly. Thus all graphs with $n-2$ links necessarily have a negative value if all graphs composed of $n-1$ links also have a negative value. We can apply this reasoning recurrently to all networks of connectivity $k = n-2$, $n-3, \ldots, n-(n-1)$. Thus, under condition (A.2), all networks with k links such that $n-1 > k > 0$ have a negative value. Thus, the empty network is the efficient network. This completes the proof. $\square$

Part ii) Any non-empty acyclic graph is a tree or a set of disjoint trees (with potentially some isolated agents). A tree of m nodes always has $k = m-1$ links. A tree of $m+u-1$ nodes generates 17 more utility than a graph composed of two distinct trees of m and u nodes; that is, because with the same number of links, it generates several more indirect connections. Thus we can restrict our analysis to connected acyclic graphs that are necessarily trees. Consider now the trees that do not exhibit any regional overlap. In this situation, any link ij in g generates a cost equal to its covering graph while it generates less utility since fewer agents are thus directly or indirectly connected. This applies for all bonds with no regional overlap. Thus the value of such network is always below the one of its associated covering graph. This completes the proof. $\square$

Part iii) The gross social value of any link is at least equal to $2\delta - 2\delta^2$, while their costs are equal to $2k\lceil n/2\rceil^{-1}$ when they are formed at geographic distance $k\lceil n/2\rceil^{-1}$ with $k \in \mathbb{N}^*$. It is straightforward to infer that the formation of any link at geographic distance $k\lceil n/2\rceil^{-1}$ or less is increasing social surplus when $\delta - \delta^2 > k\lceil n/2\rceil^{-1}$. Therefore, these links must be included in the efficient network. $\square$

A.2. Proof of Proposition 2

Part i) When $\delta < \lceil n/2\rceil^{-1}$, it is easy to show that the empty network is always stable. Being on the empty net, no agent has any interest in forming a link even with his direct (geographic) neighbors since this connection will always cost him more than the (direct) gross payoff it may bring to him. Moreover, as showed by Jackson and Wolinsky, in such a situation, stability implies no *loose end* (i.e. no agent i is connected to only one other agent j). That is because j will always find interest in severing this connection. Since all acyclic networks but the empty graph always have loose ends, the empty network is the only acyclic pairwise stable network. When $\delta = \lceil n/2\rceil^{-1}$, no agent will may find strictly beneficial to form a link. Finally, when $\delta > \lceil n/2\rceil^{-1}$, the empty graph is trivially unstable since two geographic neighbors always have interest in forming a connection. $\square$

Part ii) An under bound expression for the gross individual returns of any direct connection is $\delta - \delta^2$. Thus, if $\delta - \delta^2 > k\lceil n/2\rceil^{-1}$, $\forall k \in \mathbb{N}^*$, then any agent has an incentive to be connected to her neighbors located at geographic distance equal or below $k\lceil n/2\rceil^{-1}$, because the costs of doing so are at max equal to $k\lceil n/2\rceil^{-1}$. Then, any pairwise stable network must contain all links that can be formed at a geographic distance less or equal to $k\lceil n/2\rceil^{-1}$. Otherwise, they would be unstable, since at least two unconnected agents located at geographic distance $k\lceil n/2\rceil^{-1}$ or less from one another would be mutually willing to form such a link.

Part iii) In the star network, the center of the star is never interested in maintaining a link with his most distant neighbor. This link costs him 1, which is strictly more than his gross utility, which is simply $\delta < 1$. In the complete network, no agent has an incentive to maintain his most distant connection since its deletion would reduce costs by 1, while gross utility would decrease by only $\delta - \delta^2 < 1$. Then, neither the star nor the complete graphs are pairwise stable. $\square$

References

Albert, R., Barabási, A.L., 2002. Statistical mechanics of complex networks. Review of Modern Physics 74, 47–97.
Albert, R., Barabási, A.L., 1999. The diameter of the World Wide Web. Nature 401, 130–131.
Allen, R.C., 1983. Collective invention. Journal of Economic Behavior and Organization 4, 1–24.
Almeida, P., Kogut, B., 1999. The localization of knowledge and the mobility of engineers in regional networks. Management Science 45, 905–917.
Bala, V., Goyal, S., 2000. A non-cooperative model of network formation. Econometrica 68, 1181–1229.
Barabási, A.L., Albert, R., 1999. Emergence of scaling in random networks. Science 286, 509–512.
Bossard, J., 1932. Residential propinquity as a factor in marriage selection. American Journal of Sociology 38, 219–222.
Breschi, S., Lissoni, F., 2007. Cross-firm inventors and social networks: localised knowledge spillovers revisited. Annales d'Economie et de Statistiques, forthcoming.
Burt, R., 2004. Structural holes and good ideas. American Journal of Sociology 110, 349–399.
Burt, R., 1997. The contingent value of social capital. Administrative Science Quarterly 42 (2), 339–365.
Burt, R., 1992. Structural Holes: The Social Structure of Competition. Harvard University Press, Cambridge.
Calvó-Armengol, A., 2004. Job contact networks. Journal of Economic Theory 115, 191–206.
Carayol, N., Roux, P., 2004. Behavioral foundations and equilibrium notions for social network formation processes. Advances in Complex Systems 7 (1), 77–92.
Cowan, R., Jonard, N., 2003. The dynamics of collective invention. Journal of Economic Behavior and Organization 52, 513–532.
David, P.A., Foray, D., 1994. Percolation structures, Markov random fields and the economics of EDI standard diffusion. In: Pogorel, G. (Ed.), Global Telecommunication Strategies and Technological Changes. North-Holland, Amsterdam, pp. 135–170.
deSolla Price, D.J., 1965. Networks of scientific papers. Science 149, 510–515.
Erdös, P., Rényi, A., 1960. On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences 5, 290–297.
Freidlin, M., Wentzell, A., 1984. Random Perturbations of Dynamical Systems. Springer Verlag, New York.
Galeotti, A., Goyal, S., Kamphorst, J., 2006. Network formation with heterogeneous players. Games and Economic Behavior 54, 353–372.
Gastner, M.T., Newman, M.E.J., 2006. The spatial structure of networks. European Physical Journal B 49, 247–252.
Goyal, S., Joshi, S., 2003. Networks of collaboration in oligopoly. Games and Economic Behavior 43, 57–85.
Goyal, S., Moraga, J.L., 2001. R&D networks. RAND Journal of Economics 32, 686–707.

Goyal, S., Moraga, J.L., van der Leij, M., 2006. Economics: an emerging small world. Journal of Political Economy 114, 403–412.
Granovetter, M., 1974. Getting a Job: A Study of Contact and Career. Harvard University Press, Cambridge.
Granovetter, M., 1973. The strength of weak ties. American Journal of Sociology 78, 1360–1380.
Jackson, M.O., Rogers, B.W., 2005. The economics of small worlds. Journal of the European Economic Association 3 (2–3), 617–627.
Jackson, M.O., Watts, A., 2002. The evolution of social and economic networks. Journal of Economic Theory 106, 265–295.
Jackson, M.O., Wolinsky, A., 1996. A strategic model of social and economic networks. Journal of Economic Theory 71, 44–74.
Jaffé, A., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics 63, 577–598.
Johnson, C., Gilles, R.P., 2000. Spatial social networks. Review of Economic Design 5, 273–299.
Kogut, B., Walker, G., 2001. The small world of Germany and the durability of national networks. American Sociological Review 66, 317–335.
Kranton, R., Minehart, D., 2000. Competition for goods in buyer–seller networks. Review of Economic Design 5, 301–332.
Milgram, S., 1967. The small world problem. Psychology Today 2, 60–67.
Newman, M.E.J., 2001. The structure of scientific collaborations. Proceedings of the National Acadademy of Science USA 98, 404–409.
Robles, J., 1998. Evolution with changing mutation rates. Journal of Economic Theory 79, 207–223.
Saxenian, A., 1994. Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Harvard University Press, Cambridge.
Singh, J., 2005. Collaborative networks as determinants of knowledge diffusion patterns. Management Science 51 (5), 756–770.
Stouffer, S., 1940. Intervening opportunities: a theory relating mobility and distance. American Sociological Review 5, 845–867.
Strogatz, S.H., 2001. Exploring Complex Networks. Nature 410 (6825), 268–276.
Valente, T., 1996. Social network thresholds in the diffusion of innovations. Social Networks 18, 69–89.
Vega-Redondo, F., 2006. Building up social capital in a changing world. Journal of Economic Dynamics and Control 30, 2305–2338.
Von Hippel, E., 1989. Cooperation between rivals: informal knowhow trading. In: Carlsson, B. (Ed.), Industrial Dynamics, Technological, Organizational and Structural Changes in Industries and Firms. Kluwer, Boston, pp. 157–176.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small worlds' networks. Nature 393, 440–442.
Young, H.P., 2002. The diffusion of innovations in social networks. Mimeo SantaFe Institute.
Young, H.P., 1998. Individual Strategy and Social Structure. Princeton University Press, Princeton.
Young, H.P., 1993. The evolution of conventions. Econometrica 61, 57–84.
Zipf, G., 1946. The P1P2/D hypothesis: On the intercity movement of persons. American Sociological Review 11, 677–686.