Bureau d'économie théorique et appliquée (BETA)
UMR 7522

Documents de travail

Documents de travail

« Knowledge flows and the geography of neworks. A strategic model of small worlds formation »

Auteurs

Nicolas CARAYOL
Pascale ROUX

Document de travail n° 2006–**16**

Mai 2006

Faculté des sciences économiques et de gestion
Pôle européen de gestion et d'économie (PEGE)
61 avenue de la Forêt Noire
F-67085 Strasbourg Cedex

Secrétariat du BETA
Christine Demange
Tél. : (33) 03 90 24 20 69
Fax : (33) 03 90 24 20 70
demange@cournot.u-strasbg.fr
http://cournot.u-strasbg.fr/beta

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITÉ LOUIS PASTEUR STRASBOURG

Knowledge flows and the geography of networks. A strategic model of small worlds formation

Nicolas Carayol and Pascale Roux

BETA, CNRS - Université Louis Pasteur (Strasbourg 1)

ADIS, Faculté Jean Monnet, Université Paris Sud (Orsay-Paris 11)

First version : October 2003 / Revised version : May 2006

Abstract

This paper aims to demonstrate that the strategic approach of network formation can generate networks that share the main structural properties of most real social networks. We introduce a spatialized variation of the Connections model (Jackson and Wolinski, 1996) in which agents balance the benefits of forming links resulting from imperfect knowledge flows through bonds against their costs which increase with geographic distance. We show that, for intermediary levels of knowledge transferability, our time-inhomogeneous process selects networks which exhibit high clustering, short average distances and, when the costs of link formation are normally distributed across agents, skewed degree distributions.

JEL classification : D85 ; C63 ; Z13

Keywords : Strategic network formation ; Time-inhomogeneous process ; Knowledge flows ; Small worlds ; Monte Carlo simulations

Corresponding author. Nicolas Carayol, BETA (UMR CNRS 7522), Université Louis Pasteur, 61 avenue de la Forêt Noire, F-67085 Strasbourg Cedex ; tel : +33-388352204 ; fax : +33-390242071 ; email : carayol@cournot.u-strasbg.fr.

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 Wolinski, 1996 ; Bala and Goyal, 2000). Models of such kind 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), etc. Nevertheless, this literature has not yet dedicated much attention to the full characterization of the endogenous networks and most importantly to the strategic conditions that may lead to the emergence of more complex networks that might resemble real networks.

At the same time, the structural properties of real networks have been extensively explored and characterized by some recent contributions in the field of Statistical Physics. In short, this (still highly active) literature has shown that in most social networks, the average distance between any two nodes (computed as the minimal number of inter-individual connections) is remarkably short[1] while the agents remain highly clustered. Networks that simultaneously share those two structural properties are said to be *small worlds* à la Watts and Strogatz (1998)[2]. Moreover, most real networks also exhibit a skewed (right-asymmetric) distribution of connections among agents (Albert and Barabási, 1999) which means that few agents have many links while most agents are weakly connected[3].

In this paper, we demonstrate that the strategic approach to network formation can lead to the emergence of networks that share such structural properties[4]. For that purpose, we introduce a strategic model of network formation built on a simple variation of the Connections model of Jackson and Wolinski (1996). Myopic self-interested agents form costly links 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 manner as Johnson and Gilles, 2000). The model typically applies to knowledge flows in interpersonal connections between inventors for instance. Indeed, it has been repeatedly argued that invention activity is far from being the outcome of isolated agents efforts but 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[5].

[1] This idea can be traced back to the "six degree of separation" experiment of Milgram (1967).

[2] Such properties have for instance been evidenced for web sites links or coauthorship of scientific papers (see Barabási and Albert, 1999 ; Newman, 2001).

[3] Survey papers on this literature are Strogatz (2001) and Albert and Barabási (2002).

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

[5] Several previous theoretical works focus on how some fixed network structures affect information, knowledge or technology diffusion (for example, David and Foray, 1994 ; Valente, 1996 ; Cowan and Jonard, 2003 ; Young, 2002).

We study, in a dynamic time-inhomogeneous process of networks formation, how the degree of knowledge transferability in networks, which may be related to its relative degree of tacitness vs. codification, generates qualitatively different network architectures. In particular, the model provides a theoretical ground for the relative density of invention networks depending on the nature of knowledge : emergent networks are less dense when knowledge is either mostly tacit or mostly codified. The model also provides an economic rationale for why connections between inventors are mainly established in the local space and thus why knowledge mostly diffuses locally (Almeida and Kogut, 1999 ; Breschi and Lissoni, 2006). Our most striking result is that, for a large region of intermediary levels of knowledge transferability, the stochastic process leads to emergent networks that share the *small world* properties : they are highly clustered and exhibit a low average distance thanks to some distant connections (Granovetter, 1973). The latter are strategically formed while such shortcuts are usually thought as randomly attributed. Lastly, when we introduce a normal (Gaussian) heterogeneity among agents on the costs they bear for their direct connections, then emergent networks also exhibit skewed distributions of neighborhoods, that is some few agents tend to play more central roles in the population.

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 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 turn to formal definitions on networks stability and efficiency.

2.1. Graphs

Consider a finite set of n agents, $N = \{1, 2, ..., n\}$ 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. We define the complete graph $g^N = \{ij \mid i, j \in N\}$ as the set of all subsets of N of size 2, where each player is connected with all others. Let $g \subseteq g^N$ be an arbitrary collection of links on N. We define $G = \left\{g \subseteq g^N\right\}$ as the finite set of all possible graphs between the n agents.

Let $g' = g + ij = g \cup \{ij\}$ and $g'' = g - ij = g \backslash \{ij\}$ be respectively the graph obtained by adding ij and the one obtained by deleting ij from the existing graph g. The graphs g and g' are said to be *adjacent* as well as the graphs g and g''. For any g, we define $N(g) = \{i \mid \exists j : ij \in g\}$, 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 : $N_i(g) = \{j \mid ij \in g\}$. 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 $\{i_1i_2, i_2i_3, ..., i_{k-1}i_k\} \subset g$ 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. The length of a path is the number of edges it contains. Let $i \longleftrightarrow_g j$ be the set of paths connecting i and j on graph g. The set of *shortest paths* between i and j on g noted $i\widetilde{\longleftrightarrow}_g j$ is such that $\forall k \in i\widetilde{\longleftrightarrow}_g j$, then $k \in i \longleftrightarrow_g j$ and $\#k = \min_{h \in i \longleftrightarrow_g j} \#h$. We define the *geodesic distance* between two agents i and j as the number of links of a shortest path between them : $d(i,j) = d_g(i,j) = \#k$, with $k \in i\widetilde{\longleftrightarrow}_g j$. When there is no path between i and j then their geodesic distance is conventionally infinite : $d(i,j) = \infty$.

If agents have fixed positions in a given space, representing for example their geographic location, one can define a new distance operator denoted $l(i,j)$. In our model, we consider that agents are equidistantly located on a circle (or a ring) with unitary intervals. Without loss of generality, agents are ordered according to their index, such that i is the immediate geographic neighbor of agent $i+1$ and agent $i-1$ but agent 1 and agent n who are neighbors. As a consequence, the geographic distance between any two agents is given by $l(i,j) = \min\{|i-j|\,; n - |i-j|\}$.

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. The *chain* g^c, is defined as a connected subset of the ring, that is $g^c \subset g^{\circ}$ and $\forall i, j \in N(g^c), i \longleftrightarrow_{g^c} j \neq \emptyset$. If $\#g^c = \#g^{\circ}$ then $g^c = g^{\circ}$. 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 $i \longleftrightarrow_{g^c} j$ is a singleton). The covering chain of the graph g is a chain g^{cc} such that for all $i, j \in N : i\widetilde{\longleftrightarrow}_{g^{\circ}} j \subset g^{cc}$ iff $ij \in g$. 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 $g^{3\circ}$, all agents are only connected with their six closest neighbors. Finally, a (complete) *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.

2.2. Networks 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 not its deletion which can emanate from one of them unilaterally. Moreover, agents are myopic : they take their decisions on the basis of the immediate impacts on their current payoffs. Formally, let $\pi_i : \{g \mid g \subseteq g^N\} \to \mathbb{R}$, the payoffs received by i from his position in the network g, with $\pi_i(\emptyset) = 0$.

Jackson and Wolinski (1996) introduce the notion of *pairwise stability* which departs from the 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 regard network efficiency, we use the 'strong' notion introduced by Jackson and Wolinski (1996). It relies 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 $\{g \mid g \subseteq g^N\}$, that is : $\pi(g) \geq \pi(g')$ for all $g' \subseteq g^N$. It should also be noticed 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 Wolinski (1996) and present some new analytical results we obtain 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 which represents the quality of knowledge transfer through each bilateral connection. In a given network, agents can not strategically control the circulation of knowledge. Moreover, agents bear costs for maintaining direct connections. Relying on Debreu's (1969) hypothesis according to which closely located players incur lower costs to establish communications, we let links costs linearly increase with the geographic distance between agents on the circle[6].

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

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

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. $\delta \in]0;1[$ is the decay parameter which gives 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 would increase with the codification of knowledge. Thus, $\delta^{d(i,j)}$ gives the payoffs resulting from the (direct or indirect) connection between i and j. It is a decreasing function of the geodesic distance since δ is less than the unity. Notice 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

[6] Johnson and Gilles (2000) first introduced such costs but consider a linear world.

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). The second part describes the costs of direct links. We assume that the cost of a direct connection between i and j is proportional to the geographic distance separating them $c_{ij} \propto l(i,j)$, with $l(i,j)$ computed on the external metrics (the circle) considered. As a normalization device and in order to account for an inverse relation between the costs and the size of the population, we assume that the costliest connection costs unity : $\max_{i,j \in N} c_{ij} \equiv 1$, which is also the upper bound of a positive externality. In the ring metric we have $\max_{i,j \in N} l(i,j) = \lceil n/2 \rceil$, the smallest integer higher than or equal to $n/2$. Thus, it comes :

$$c_{ij} = l(i,j) / \lceil n/2 \rceil . \tag{2}$$

3.2. Some analytical results

The analytical results obtained on networks efficiency and stability in the model described above are summed up in the following two propositions.

Proposition 1. *Efficiency.*

i) The empty network $g^{\emptyset}$ *is the only efficient network when* $\delta + \frac{(n-2)}{4}\delta^2 < \lceil n/2 \rceil^{-1}$.

ii) If $\delta^2 - \delta^{n-1} < \lceil n/2 \rceil^{-1}$, *the value of any acyclic graph* g *is less than its associated covering chain* g^c.

iii) Consider three chains $g^c, g^{c\prime}$ and $g^{c\prime\prime}$, *if* $g^{c\prime} \cap g^{c\prime\prime} = \emptyset$ *and* $\#g^c = \#g^{c\prime} + \#g^{c\prime\prime}$, *then* $\pi(g^c) > \pi(g^{c\prime}) + \pi(g^{c\prime\prime})$. *Moreover a maximal chain* g^{mc} *(a chain such that* $\#g^{mc} = n-1$*) is the most efficient positive value chain.*

Proofs. *See Appendix.*

Proposition 2. *Stability.*

i) When $\delta > \lceil n/2 \rceil^{-1}$, *the empty graph is never pairwise stable. When* $\delta < \lceil n/2 \rceil^{-1}$, *the empty graph* $g^{\emptyset}$ *is the unique acyclic pairwise stable graph and no network containing a peripheral agent (has only one connection) is pairwise stable. The empty graph is pairwise stable when* $\delta = \lceil n/2 \rceil^{-1}$.

ii) The star $g^{\star}$ *and the complete network* g^N *are never pairwise stable.*

Proofs. *See Appendix.*

4. The dynamic process of networks formation

This section is a self-contained presentation of our dynamic model. We first introduce the standard perturbed stochastic process of network formation and next focus on the terms by which our process differentiates.

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 $t = 1, 2, ...$). 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 renounce unilaterally. Formally, those 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 only depends on the present state of the system given by the graph structure g_t. The stochastic process is thus 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 $\overline{\varepsilon} > 0$ which invert agents' right decisions in creating, maintaining or deleting links. These perturbations may be understood as mistakes or as mutations. For small but non null values of $\overline{\varepsilon}$, it can be shown that the discrete-time Markov chain associated to the transformed transition matrix $\mathrm{P}(\overline{\varepsilon})$ is irreductible and aperiodic and has thus a unique corresponding stationary distribution $\mu(\overline{\varepsilon})$ which is the solution of $\mu(\overline{\varepsilon}) \times \mathrm{P}(\overline{\varepsilon}) = \mu(\overline{\varepsilon})$. Such perturbed stochastic process is ergodic. Intuitively ergodicity occurs when it is possible to transit directly or indirectly between any chosen pair of states in a potentially very long period of time. It allows the long run equilibrium ($\mu(\overline{\varepsilon})$) of the system to be unique and independent of the initial conditions.

Usually, the modeler computes $\mu^* = \lim_{\overline{\varepsilon} \to 0} \mu(\overline{\varepsilon})$ 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 (2002) 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) which, in the network context, are likely to be extremely numerous. To make that point clear, we shall recall 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, possibly thousands of networks are 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 :

$$\varepsilon_t = 1/(t+1) + \overline{\varepsilon} , \tag{3}$$

with $\overline{\varepsilon} > 0$. This rule ensures that a significant noise affects the dynamics in the beginning while it decreases monotonically with time down to a small strictly positive limit : $\lim_{t \to \infty} \varepsilon_t = \overline{\varepsilon}$.

According to Robles (1998), the long run equilibrium $\psi(\overline{\varepsilon})$ of such time-inhomogeneous Markov chain exists, is unique and is equal to the equilibrium of the Markov chain perturbed by the constant error $\overline{\varepsilon}$: $\psi(\overline{\varepsilon}) = \mu(\overline{\varepsilon})$ [7]. It is then obviously ergodic. This property is interesting since it renders numerical experiments more tractable in order to examine with good confidence the long run behavior of the system (Vega-Redondo, 2006). Notice that the time-inhomogeneity of noise assumption is more satisfactory than the time-homogeneity of noise one. Agents are likely to make less and less errors through time while still a very small error probability persists in the long run.

We label the networks on which the process stabilizes in the long run as *emergent networks*. Formally the set of emergent networks is $\widehat{G} = \{g \in G \,|\psi_g(\overline{\varepsilon}) > 0\}$. This set is broader than the set of stochastically stable networks defined above which is included in $\widehat{G}$. That is, for all g such that, if $\lim_{\overline{\varepsilon}\to 0} \mu_g(\overline{\varepsilon}) > 0$ then $\psi_g(\overline{\varepsilon}) > 0$[8].

5. Results : The structural properties of emergent networks

In this section we study the emergent networks selected by the stochastic process described in the preceding section. We use Monte Carlo experiments to generate networks that are on the support of the unique limiting stationary distribution $\psi(\overline{\varepsilon})$ (of networks) of the perturbed dynamic process presented above. The limit error term is $\overline{\varepsilon} = 10^{-4}$. The size of the network is arbitrarily set to $n = 20$ agents. All experiments are stopped at $T = 20,000$, period after which the process is proven to have almost surely stabilized on a given pairwise stable state[9]. We shall focus on the characterization of the structural properties of emergent networks depending on the knowledge transfer parameter δ. For this purpose we ran 1,500 experiments performed with randomly drawn values of δ over its value space $]0, 1[$.

5.1 How density is 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 of the network as follows :

$$\hat{\eta}(g) = \eta(g)/n. \tag{4}$$

[7] See Proposition 3.1 of Robles (1998, p. 211).

[8] 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.

[9] In average, non pairwise stable networks are rarely found after 20,000 periods (less than 4%). Let us 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 can not exclude that some non pairwise stable networks are just emergent networks since $\overline{\varepsilon} > 0$. Nevertheless, in practice, when such a non pairwise stable network was found, it is because an error recently affected the system. Thus, we dropped these experiments to ensure that the results are not contingent to recent arbitrary changes. Thus, our approximations of the emergent networks distributions restrict to the emergent networks that are pairwise stable.

Let first examine how the average degree $\hat{\eta}(g)$ is affected by δ. As shown in Figure 1, the average degree is null when $\delta \leq \lceil n/2 \rceil^{-1} = 0.1$: the process converges to the empty graph which 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 configuration appears to be 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 $\delta \approx 0.6$, average degree decreases down to a value slightly below unity when δ reaches its maximum.

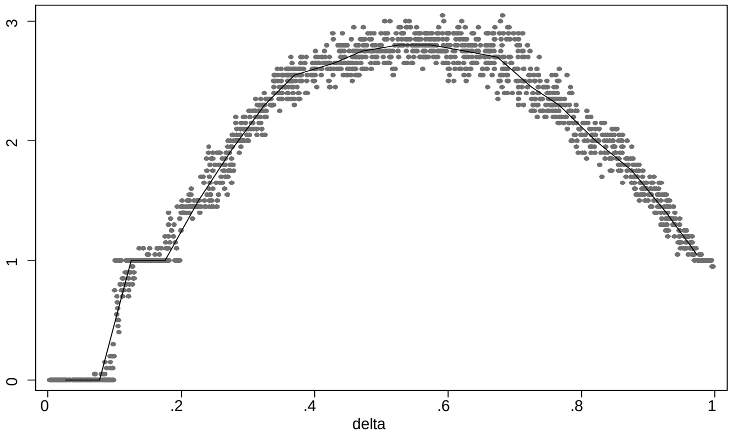

Figure 1. *Average degree $\hat{\eta}(g)$ of emergent networks for 1,500 simulations with randomly drawn values of $\delta \in]0;1[$.*

Let us remind 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 would increase with the codification of knowledge. Our model thus provides an explanation of how the nature of knowledge affects the density of emergent networks. We find that both when knowledge is mainly tacit (δ close to 0) or mainly 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 overbalance their associated costs. Thus, only local minimally connected networks emerge (when knowledge is extremely tacit, no connection is formed). It is when the knowledge has balanced codified and tacit parts that such private returns are sufficiently high for the network average degree to increase.

5.2 When do small worlds emerge ?

Behind the dots of Figure 1, one can find networks with many different structures. This diversity calls for a statistical analysis of the structural properties of emergent networks. For that purpose, we compute two dedicated indexes. The former is the *average distance* (or average path length) of (directly or indirectly) connected agents. It is given by :

$$d(g) = \frac{\sum_i \sum_{j \neq i} d(i,j) \times 1\{i \leftrightarrow_g j \neq \emptyset\}}{\#\{i,j \,|i \neq j \in N, i \leftrightarrow_g j \neq \emptyset\}}, \quad (5)$$

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 into 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 :

$$c(g) = \frac{1}{n} \sum_{i \in N; \eta_i(g) > 1} \frac{\#\{jl \in g \,|\, j \neq l \in N_i(g)\}}{\#\{j, l \,|\, l \neq j \in N_i(g)\}}. \tag{6}$$

In words this index measures the propensity with which an agent's neighbors are also neighbors together.

The two indicators presented above are affected by the average degree of the network ($\hat{\eta}(g)$) that is likely to vary with δ. Therefore, these indicators are somehow biased and we must find an efficient control for average degree. In the spirit of Watts and Strogatz (1998), we associate *control random graphs* to 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 Rény, 1960). For each given number of edges of emergent networks, the average distance and the average clustering are numerically computed and averaged over $1,000$ of 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 $1,000$ random networks that have exactly the same average degree as g. Similarly we compute $d(g)/d(g^{rd})$. These ratios are plotted in Figure 2.

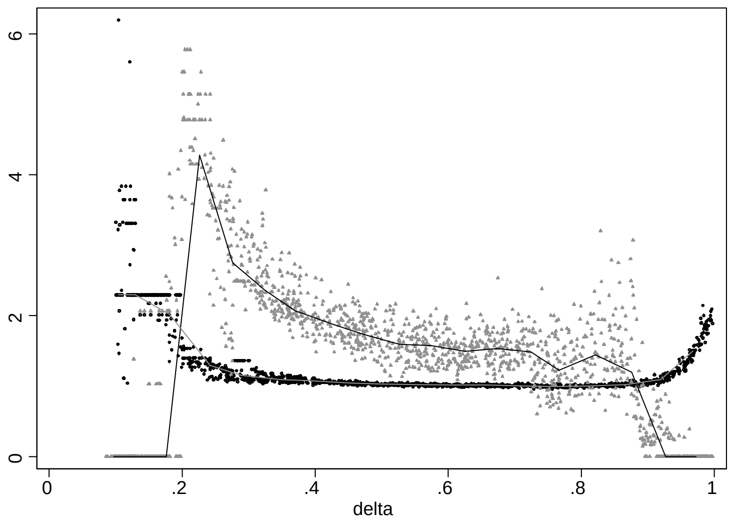

Figure 2. *Average distance (black circles : $d(g)/d(g^{rd})$), average clustering (grey triangles : $c(g)/c(g^{rd})$) and their fitted values of emergent networks for 1,500 simulations with randomly drawn values of* $\delta \in]0;1[$.

Such ratios have been used by Watts and Strogatz (1998) to identify a specific but frequently observed network structure called a *small world*. It is characterized by the two following properties :

$$c(g)/c\left(g^{rd}\right) \gg 1 \quad \text{and} \quad d(g)/d\left(g^{rd}\right) \approx 1. \tag{7}$$

Small world networks are highly clustered as compared to random graphs and simultaneously their average distance is close to the one of random graphs which are known to exhibit 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 stays on 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

Before exploring statistically the extent to which networks correlate and/or discorrelate with the ring-like geography, we provide in Figure 3 some intuitions on the typical networks shapes obtained for several values of the parameter δ. When $0.09 \lesssim \delta \leq 0.2$, it is the ring g° which emerges most often (characterized by an average degree equal to one in Figure 1 and a null average clustering in Figure 2). So agents not only have two neighbors in average, but all are connected to their two closest geographic neighbors. When δ is equal to 0.3, networks in which all agents are connected to their four closest geographic neighbors are likely to emerge. Such situation corresponds to the double geographic ring $g^{2\circ}$. 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 these two 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, *ii*) the existence of some "short cuts". The specificity appears clearly when one compares these structures with the typical networks one obtains with the standard connections model.

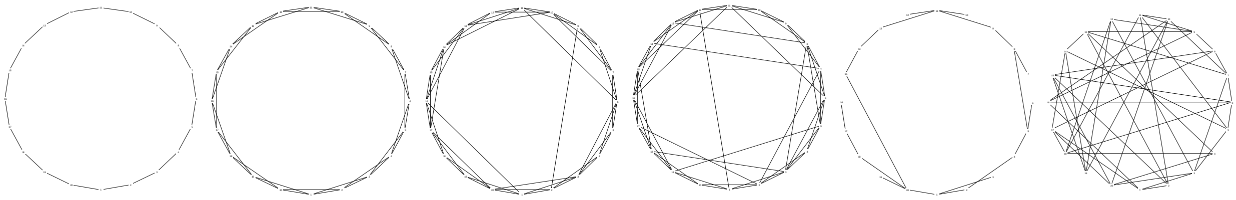

Figure 3. *Typical emergent networks obtained with* $\delta = 0.15$, $\delta = 0.3$, $\delta = 0.35, \delta = 0.7$, $\delta = 0.98$. *The last network has been generated with the standard connections model of Jackson and Wolinski (1996) with* $\delta = 0.7$ *and a connection cost equal to* 0.5[10].

[10] The link cost in the standard connections model is set to 0.5 because this value is close to the average link cost borne by an agent to connect to any other agent in our spatialized 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 geographic distance between two linked agents. It is formally defined as

$$\mathrm{p}\,(h) = \frac{1}{\eta\,(g)} \sum_{ij \in g} 1\,\{l(i,j) = h\}\,, \tag{8}$$

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

In order to explore such a distribution when the emergent networks have the small world properties, 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 Figure 4. Unsurprisingly, networks do correlate with the geographic metrics. Indeed, between 70% and 80% of links connect two agents at a geographical distance equal or less than 2. Clustering is thus achieved in local space. Nevertheless, the networks also exhibit these distant connections that make p$(\cdot)$ be long tailed on the upper side of the spectrum. The originality of these networks clearly relies on the fact that they both correlate with space and have some uncorrelated connections. Lastly, when δ equals 0.5 or 0.7, then there is an increased probability that connections are made at geographical distance 5 and 6. Agents find it profitable to establish such medium distance connections as a benchmark between overcoming the local cluster of which they already benefit from and bearing not too high connection costs. This result supports the idea that sustainable short cuts are not made at a random distance (and cost) like in Watts and Strogatz (1998) but are driven by some strategic behavior.

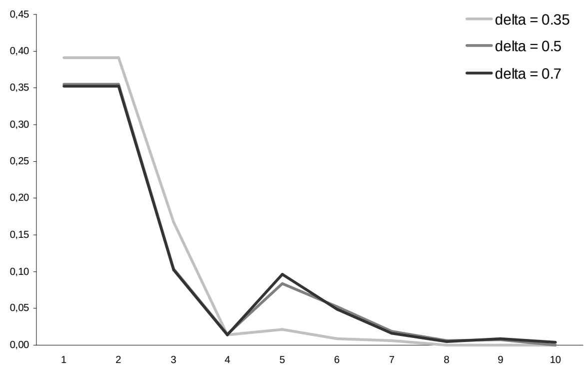

Figure 4. *Distribution of links* p$(\cdot)$*, according to the geographic distance between the two directly connected agents* $l(i,j)$ *(averaged over 100 experiments for each value of* $\delta = 0.35, 0.5, 0.7$*).*

Since distant connections are necessary to meet the small world properties and since these distant connections are also very costly (remember that connection costs increase with geographic distance), a natural issue raises : why do some agents bear such costs ? In other words, does it pay to sustain shortcuts ? For this purpose, 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 (geographic) distance between i and i's neighbors : $l_i \equiv \max_{j \in N_i(g)} l(i,j)$. Two-way plotted results are presented in Figure 5. It shows that individual payoffs decrease in average when agents sustain a connection longer or equal to 4. Therefore we can conclude that it does not pay to sustain distant connections in the long run[11]. Nevertheless, for any shortcut to exist, it must remain advantageous to sustain it : payoffs must be increased with rather than without it. The agents who sustain a distant connection are locked in, in the sense they are better off with this connection though they can not 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.

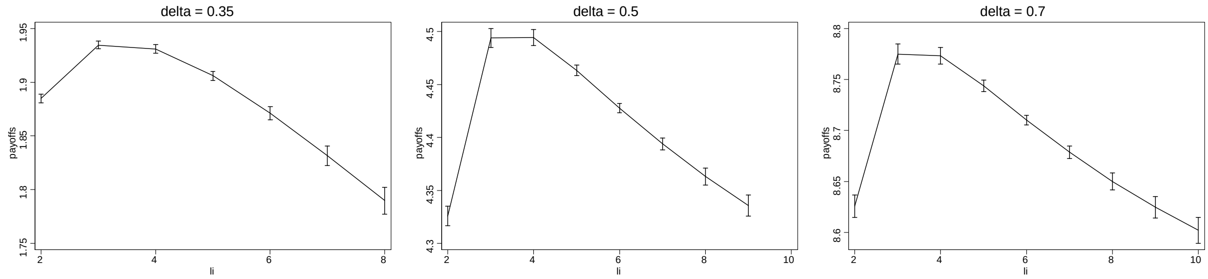

Figure 5. *Fractional polynomial estimates (with 90% confidence interval) of individual payoffs (π_i in ordinates) respective to agent's longest (geographically) connection (l_i in abscises) (*$100 \times 20 = 2,000$ *agents are considered for each of the three values of δ considered :* 0.35 *(left),* 0.5 *(middle),* 0.7 *(right).*

5.4 The heterogeneity of agents and skewed degree distribution

Low average path length and high average clustering characterize small worlds *à la* Watt and Strogatz (1998). Another general result of the physics literature on social networks is that their degree distribution is usually skewed. This last result comes from the *scale free networks* approach (Barabási and Albert, 1999). In short, this means many agents have few connections while few agents have many connections. In the model presented in equation (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.

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

$$\pi_i(g_t) = \sum_{j \in N \setminus i} \delta^{d(i,j)} - a_i \sum_{j: ij \in g_t} c_{ij}, \tag{9}$$

[11] Due to space constraints, we did not report simple econometric estimations that confirm the inverse-U relation between l_i and payoffs when we control for network density that explains much of the variance observed in Figure 5.

where $a_i > 0$ and its mean $\langle a_i \rangle = 1$ to keep consistency with (1). This implies that the costs bore 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 Figure 6. The a_i are randomly allocated to agents independently of geography.

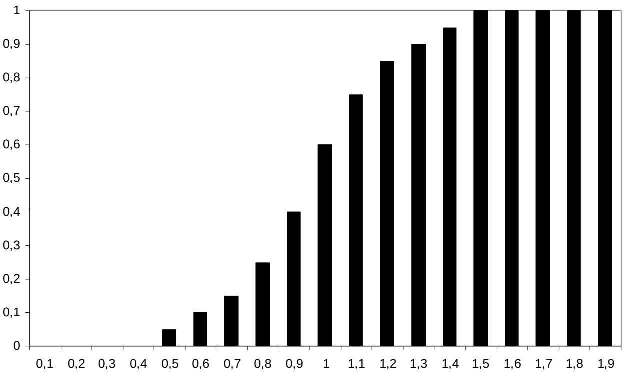

Figure 6. *Distribution of* a_i *among agents.*

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

$$\rho(k) = \frac{1}{n} \sum_{i \in N(g)} 1\{\eta_i(g) = k\}, \tag{10}$$

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

As in the previous subsection, we perform, with the heterogeneous agents model (a_i distributed as presented in Figure 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 Figure 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[12], degree distribution is indeed much more dispersed when agents are heterogeneous. Unsurprisingly, the agents who have lower costs are much more inclined to form numerous connections. What is of much interest is that a Gaussian-like distribution of agents as regard costs 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 noticed 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 properties are fully preserved for $0.35 \leq \delta \leq 0.7$[13]. Therefore, all the three characteristics (short average path length, high average clustering and skewed degree) are then simultaneously observed on emergent networks.

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

[13] 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 (1998) is found. Numerical results are available from the authors on request.

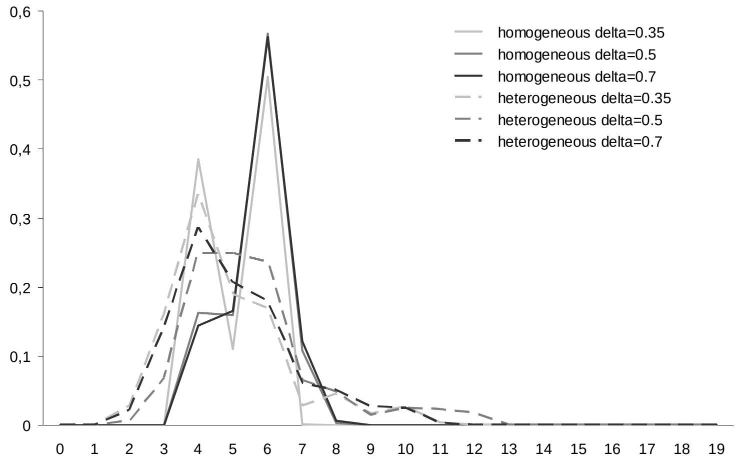

Figure 7. *Degree distribution* $\rho(k)$ *(averaged over 100 experiments for each value of* $\delta : 0.35, 0.5, 0.7$*, for the two models described in* (1) *and* (9)*).*

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 severe links. For that, they balance the benefits obtained from knowledge that imperfectly flows through bilateral relationships (positive externalities deteriorate with relational distance) against the costs they pay to maintain 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 a simple model of strategic network formation which is a spatialized variation of the Connections model developed by Jackson and Wolinski (1996), can lead, in a dynamic setting, to emergent networks that share the main structural properties of most real social networks. We first studied the effects of the parameter which tunes the decay of positive externalities in each bilateral connection, on the structure of emergent networks. We found that density is the highest when the decay is around 55%-60%. It decreases down to the empty network when decay is small and to minimally connected networks when decay is close to the unity. For a large region of intermediary levels of the decay (35% to 70%), we found, while controlling for networks density, emergent networks that exhibit the small world configurations characterized by both a high average clustering and a short average path length. Clustering occurs in local space while few agents sustain distant connections. We further studied, for this region, the effects of introducing agents' heterogeneity as regard their connections costs (distributed in a Gaussian manner). We found that emergent networks exhibit in addition asymmetric distributions of neighborhoods. Therefore, we conclude that the three characteristics of a short average distance, a high average clustering and a skewed degree distribution are simultaneously gathered for a large and intermediary region of the decay parameter.

Let us now briefly discuss some implications of our results as regard the interplay between knowledge, networks and geography. Indeed, this model provides a theoretical ground for the relative density of invention networks depending on the nature of knowledge, assuming that the decay parameter measures the quality of interpersonal knowledge transfers. Emergent networks are less dense both when knowledge is highly tacit or highly codified (knowledge transferability is low or high). The model also provides an economic rationale for why connections between inventors are mainly established in the local space and thus why knowledge mostly diffuses locally. This result is fully consistent with the empirical studies of Saxenian (1994), Almeida and Kogut (1999) or Breschi and Lissoni (2006) which show that local knowledge diffusion (Jaffé et al., 1993) occurs thanks to the higher density of social networks within local areas. Furthermore, if denser and more clustered networks are found in local areas, some long distant connections are likely to be formed when the quality of knowledge transfer is intermediary. Then the private returns of distant connections overcome their high costs and are thus established. Such links are strategically formed with parsimony because, since knowledge flows in local areas thanks to local clustering, the neighbors of agents sustaining shortcuts have no incentive to duplicate distant connections they already benefit from.

This suggests agents free-ride on the distant connections of their neighbors. It can also be shown that, for some high values of knowledge transferability, the social surplus in star networks overbalance the ones of the emergent networks. Agents fail to coordinate in selecting one of them to play the role of knowledge transmitter : no agent intends to bear the costs associated to such a central position. An efficiency analysis of emergent networks is further needed in order to fully appreciate the structural differences between the emergent networks and the efficient ones. This issue constitutes an avenue for further research.

Acknowledgements

We thank Matthew Jackson, Antoni Calvó-Armengol, Fernando Vega-Redondo, Robin Cowan, and Yann Bramoullé for helpful remarks and discussions as well as the participants at the WEHIA 2003 conference, SCE 2003 conference, European Economic Association 2004 conference, I-Neck 2004 meeting, SAET 2005 conference, at regular seminars in University Louis Pasteur (Strasbourg), University of Toulouse 1, University Paris Sud (Orsay-Paris 11) and at Large Networks Seminar at University Paris 6 Jussieu. Remaining errors or omissions are ours.

Appendix

A.1 Proof of Proposition 1

i) The proof uses the following steps. We give an upper bound expression for the connected network value of k links. We show that this expression is at its maximum when $k = n - 1$ links. Next, we show that under the given condition, that expression is negative. Since the value of the empty graph is zero, it is the efficient network under the condition.

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

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

This expression considers that all 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 it is connecting immediate geographic neighbors). Now assume that $k = n - 1$. Thus the upper bound network value becomes :

$$\pi^{\max}(g \,|\#g = n-1\,) = 2(n-1)\left(\delta + 1/2n\delta^2 - 2\,\lceil n/2 \rceil^{-1}\right) \tag{A1}$$

To know whether we should consider cases where $k > n-1$, let us see how the upper expression of the network value behaves when we add a link : The maximal value with k ($k > n - 1$) links minus the max value with $k + 1$ links is equal to :

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

which is independent of k and strictly negative when $\delta < \lceil n/2 \rceil^{-1}$. Thus there is no interest in adding a new link from the beginning that is from $k = n - 1$.

On the other hand if $k < n-1$, the network is not connected, then no node can benefit from all others. For instance if $k = n - 2$, then the agents are associated to at least two connected components. Since the two components are isolated, the total value of the graph is equal to the sum of the value of the two components. Let us assume that it is possible to connect the two components by adding a bond at distance 1 (as it is assumed in the $\pi^{\max}$ expression). Then the value of this bond adds to the total value more than any of the other ones did previously while it costs at most the same. Thus the value of the graph with $n - 2$ links is negative if the one of the connected graph composed of $n - 1$ links is also negative.

Thus expression (A1) gives the maximal value of the network. It is negative when : $\delta + \frac{(n-2)}{4}\delta^2 < \lceil n/2 \rceil^{-1}$. This completes the proof.□

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 has always $k = m - 1$ links. A tree of $m + u - 1$ nodes generates 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 one more direct and several more indirect connections. Thus we can restrict our analysis to connected acyclic graphs which are necessarily trees.

One can get two different types of trees given the following definition. A network g is said to exhibit regional overlap if $\exists i \in N(g)$ such that for an arc $jh \in g, i \in N(g^{cc})$, with g^{cc} the covering chain of g.

a) Let us first consider the trees for which there is no *regional overlap.*

In that situation, any link ij in g generates a cost equal to its covering chain (Definition 3) while it generates less utility since less agents are thus directly or indirectly connected. This applied for all bonds that exhibit no regional overlap. Thus the value of such network is always below the one of its associated covering chain.

b) Consider now connected trees which exhibit some regional overlap. In that situation, each link ij of g such that $l(i,j)$ generates an extra cost of $2\lceil n/2 \rceil^{-1}$ as compared to its covering chain, while it generates at most a gross extra value of $2\delta^2 - 2\delta^{n-1}$. Thus $\pi(g) < \pi(g^c)$ if $\delta^2 - \delta^{n-1} < \lceil n/2 \rceil^{-1}$.□

Part iii) The proof of the first part of the Proposition is trivial since g^c costs as much as $g^{c\prime}$ and $g^{c\prime\prime}$, while it brings more utility due to more indirect connections. As regard, each new node added to the chain costs as much as the preceding ones while it always brings more value due to more indirect connections. Thus, if the maximal chain g^{mc} has a positive value, then it is always the most efficient chain.□

A.2 Proof of Proposition 2

Part i) When $\delta > \lceil n/2 \rceil^{-1}$, the proof is trivial : two geographic neighbors always have interest in forming a connection. 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 cost him always more than the (direct) gross payoff it may bring to him. Moreover, as showed by Jackson and Wolinski (1996), in such a situation, stability implies no *loose end*, that is no agent i is connected to only one other agent j. That is because j will always find interest in severing this connection. Thus, as noticed by Johnson and Gilles (2000), since all acyclic networks but the empty graph always have loose ends (among which the star net), the empty network is the only acyclic pairwise stable network.□

Part ii) 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 of 1 while gross utility would decrease of only $\delta - \delta^2 < 1$. Then, neither the star nor the complete graphs are paiwise stable.□

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.

Breschi, S., Lissoni, F., 2006. Mobility and social networks : Localised knowledge spillovers revisited. Annales d'Economie et de Statistiques, *forthcoming*.

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, 135-170.

Debreu, G., 1969. Neighboring economic agents. La Décision 171, 85–90.

Erdös, P., and Rény, 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.

Goyal, S., Joshi, S., 2003. Networks of collaboration in oligopoly. Games and Economic Behavior 43(1), 57-85.

Goyal, S., Moraga, J.L., 2001. R&D networks. Rand Journal of Economics 32, 686-707.

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.

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.

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. Forthcoming in Journal of Economic Dynamics and Control.

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, 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.

Documents de travail du BETA

2000–**01** *Hétérogénéité de travailleurs, dualisme et salaire d'efficience.*
Francesco DE PALMA, janvier 2000.

2000–**02** *An Algebraic Index Theorem for Non–smooth Economies.*
Gaël GIRAUD, janvier 2000.

2000–**03** *Wage Indexation, Central Bank Independence and the Cost of Disinflation.*
Giuseppe DIANA, janvier 2000.

2000–**04** *Une analyse cognitive du concept de « vision entrepreneuriale ».*
Frédéric CRÉPLET, Babak MEHMANPAZIR, février 2000.

2000–**05** *Common knowledge and consensus with noisy communication.*
Frédéric KŒSSLER, mars 2000.

2000–**06** *Sunspots and Incomplete Markets with Real Assets.*
Nadjette LAGUÉCIR, avril 2000.

2000–**07** *Common Knowledge and Interactive Behaviors : A Survey.*
Frédéric KŒSSLER, mai 2000.

2000–**08** *Knowledge and Expertise : Toward a Cognitive and Organisational Duality of the Firm.*
Frédéric CRÉPLET, Olivier DUPOUËT, Francis KERN, Francis MUNIER, mai 2000.

2000–**09** *Tie–breaking Rules and Informational Cascades : A Note.*
Frédéric KŒSSLER, Anthony ZIEGELMEYER, juin 2000.

2000–**10** *SPQR : the Four Approaches to Origin–Destination Matrix Estimation for Consideration by the MYSTIC Research Consortium.*
Marc GAUDRY, juillet 2000.

2000–**11** *SNUS–2.5, a Multimoment Analysis of Road Demand, Accidents and their Severity in Germany, 1968–1989.*
Ulrich BLUM, Marc GAUDRY, juillet 2000.

2000–**12** *On the Inconsistency of the Ordinary Least Squares Estimator for Spatial Autoregressive Processes.*
Théophile AZOMAHOU, Agénor LAHATTE, septembre 2000.

2000–**13** *Turning Box–Cox including Quadratic Forms in Regression.*
Marc GAUDRY, Ulrich BLUM, Tran LIEM, septembre 2000.

2000–**14** *Pour une approche dialogique du rôle de l'entrepreneur/managerdans l'évolution des PME : l'ISO comme révélateur ...*
Frédéric CRÉPLET, Blandine LANOUX, septembre 2000.

2000–**15** *Diversity of innovative strategy as a source of technological performance.*
Patrick LLERENA, Vanessa OLTRA, octobre 2000.

2000–**16** *Can we consider the policy instruments as cyclical substitutes ?*

Sylvie DUCHASSAING, Laurent GAGNOL, décembre 2000.

2001–**01** *Economic growth and CO2 emissions : a nonparametric approach.*
Théophile AZOMAHOU, Phu NGUYEN VAN, janvier 2001.

2001–**02** *Distributions supporting the first–order approach to principal–agent problems.*
Sandrine SPÆTER, février 2001.

2001–**03** *Développement durable et Rapports Nord–Sud dans un Modèle à Générations Imbriquées : interroger le futur pour éclairer le présent.*
Alban VERCHÈRE, février 2001.

2001–**04** *Modeling Behavioral Heterogeneity in Demand Theory.*
Isabelle MARET, mars 2001.

2001–**05** *Efficient estimation of spatial autoregressive models.*
Théophile AZOMAHOU, mars 2001.

2001–**06** *Un modèle de stratégie individuelle de primo–insertion professionnelle.*
Guy TCHIBOZO, mars 2001.

2001–**07** *Endogenous Fluctuations and Public Services in a Simple OLG Economy.*
Thomas SEEGMULLER, avril 2001.

2001–**08** *Behavioral Heterogeneity in Large Economies.*
Gaël GIRAUD, Isabelle MARET, avril 2001.

2001–**09** *GMM Estimation of Lattice Models Using Panel Data : Application.*
Théophile AZOMAHOU, avril 2001.

2001–**10** *Dépendance spatiale sur données de panel : application à la relation Brevets–R&D au niveau régional.*
Jalal EL OUARDIGHI, avril 2001.

2001–**11** *Impact économique régional d'un pôle universitaire : application au cas strasbourgeois.*
Laurent GAGNOL, Jean–Alain HÉRAUD, mai 2001.

2001–**12** *Diversity of innovative strategy as a source of technological performance.*
Patrick LLERENA, Vanessa OLTRA, mai 2001.

2001–**13** *La capacité d'innovation dans les regions de l'Union Européenne.*
Jalal EL OUARDIGHI, juin 2001.

2001–**14** *Persuasion Games with Higher Order Uncertainty.*
Frédéric KŒSSLER, juin 2001.

2001–**15** *Analyse empirique des fonctions de production de Bosnie–Herzégovine sur la période 1952–1989.*
Rabija SOMUN, juillet 2001.

2001–**16** *The Performance of German Firms in the Business–Related Service Sectors : a Dynamic Analysis.*
Phu NGUYEN VAN, Ulrich KAISER, François LAISNEY, juillet 2001.

2001–**17** *Why Central Bank Independence is high and Wage indexation is low.*
Giuseppe DIANA, septembre 2001.

2001–**18** *Le mélange des ethnies dans les PME camerounaises : l'émergence d'un modèle d'organisation du travail.*

Raphaël NKAKLEU, octobre 2001.

2001–**19** *Les déterminants de la GRH des PME camerounaises.*
Raphaël NK AKLEU, octobre 2001.

2001–**20** *Profils d'identité des dirigeants et stratégies de financement dans les PME camerounaises.*
Raphaël NKAKLEU, octobre 2001.

2001–**21** Concurrence Imparfaite, Variabilité du Taux de Marge et Fluctuations Endogènes.
Thomas SEEGMULLER, novembre 2001.

2001–**22** *Determinants of Environmental and Economic Performance of Firms : An Empirical Analysis of the European Paper Industry.*
Théophile AZOMAHOU, Phu NGUYEN VAN et Marcus WAGNER, novembre 2001.

2001–**23** *The policy mix in a monetary union under alternative policy institutions and asymmetries.*
Laurent GAGNOL et Moïse SIDIROPOULOS, décembre 2001.

2001–**24** *Restrictions on the Autoregressive Parameters of Share Systems with Spatial Dependence.*
Agénor LAHATTE, décembre 2001.

2002–**01** *Strategic Knowledge Sharing in Bayesian Games : A General Model.*
Frédéric KŒSSLER, janvier 2002.

2002–**02** *Strategic Knowledge Sharing in Bayesian Games : Applications.*
Frédéric KŒSSLER, janvier 2002.

2002–**03** *Partial Certifiability and Information Precision in a Cournot Game.*
Frédéric KŒSSLER, janvier 2002.

2002–**04** *Behavioral Heterogeneity in Large Economies.*
Gaël GIRAUD, Isabelle MARET, janvier 2002.
(Version remaniée du Document de Travail n°2001–**08**, avril 2001).

2002–**05** *Modeling Behavioral Heterogeneity in Demand Theory.*
Isabelle MARET, janvier 2002.
(Version remaniée du Document de Travail n°2001–**04**, mars 2001).

2002–**06** *Déforestation, croissance économique et population : une étude sur données de panel.*
Phu NGUYEN VAN, Théophile AZOMAHOU, janvier 2002.

2002–**07** *Theories of behavior in principal–agent relationships with hidden action.*
Claudia KESER, Marc WILLINGER, janvier 2002.

2002–**08** *Principe de précaution et comportements préventifs des firmes face aux risques environnementaux.*
Sandrine SPÆTER, janvier 2002.

2002–**09** *Endogenous Population and Environmental Quality.*
Phu NGUYEN VAN, janvier 2002.

2002–**10** *Dualité cognitive et organisationnelle de la firme au travers du concept de communauté.*
Frédéric CRÉPLET, Olivier DUPOUËT, Francis KERN, Francis MUNIER, février 2002.

2002–**11** *Comment évaluer l'amélioration du bien–être individuel issue d'une modification de la qualité du service d'élimination des déchets ménagers ?*
Valentine HEINTZ, février 2002.

2002–**12** *The Favorite–Longshot Bias in Sequential Parimutuel Betting with Non–Expected Utility Players.*
Frédéric KŒSSLER, Anthony ZIEGELMEYER, Marie–Hélène BROIHANNE, février 2002.

2002–**13** *La sensibilité aux conditions initiales dans les processus individuels de primo–insertion professionnelle : critère et enjeux.*
Guy TCHIBOZO, février 2002.

2002–**14** *Improving the Prevention of Environmental Risks with Convertible Bonds.*
André SCHMITT, Sandrine SPÆTER, mai 2002.

2002–**15** *L'altruisme intergénérationnel comme fondement commun de la courbe environnementale* ***à la*** *Kuznets et du développement durable.*
Alban VERCHÈRE, mai 2002.

2002–**16** *Aléa moral et politiques d'audit optimales dans le cadre de la pollution d'origine agricole de l'eau.*
Sandrine SPÆTER, Alban VERCHÈRE, juin 2002.

2002–**17** *Parimutuel Betting under Asymmetric Information.*
Frédéric KŒSSLER, Anthony ZIEGELMEYER, juin 2002.

2002–**18** *Pollution as a source of endogenous fluctuations and periodic welfare inequality in OLG economies.*
Thomas SEEGMULLER, Alban VERCHÈRE, juin 2002.

2002–**19** *La demande de grosses coupures et l'économie souterraine.*
Gilbert KŒNIG, juillet 2002.

2002–**20** *Efficiency of Nonpoint Source Pollution Instruments with Externality Among Polluters : An Experimental Study.*
François COCHARD, Marc WILLINGER, Anastasios XEPAPADEAS, juillet 2002.

2002–**21** *Taille optimale dans l'industrie du séchage du bois et avantage compétitif du bois–énergie : une modélisation microéconomique.*
Alexandre SOKIC, octobre 2002.

2002–**22** *Modelling Behavioral Heterogeneity.*
Gaël GIRAUD, Isabelle MARET, novembre 2002.

2002–**23** *Le changement organisationnel en PME : quels acteurs pour quels apprentissages ?*
Blandine LANOUX, novembre 2002.

2002–**24** *TECHNOLOGY POLICY AND COOPERATION : An analytical framework for a paradigmatic approach.*
Patrick LLERENA, Mireille MATT, novembre 2002.

2003–**01** *Peut–on parler de délégation dans les PME camerounaises ?*
Raphaël NKAKLEU, mars 2003.

2003–**02** *L'identité organisationnelle et création du capital social : la tontine d'entreprise comme facteur déclenchant dans le contexte africain.*
Raphaël NKAKLEU, avril 2003.

2003–**03** *A semiparametric analysis of determinants of protected area.*
Phu NGUYEN VAN, avril 2003.

2003–**04** *Strategic Market Games with a Finite Horizon and Incomplete Markets.*
Gaël GIRAUD et Sonia WEYERS, avril 2003.

2003–**05** *Exact Homothetic or Cobb–Douglas Behavior Through Aggregation.*
Gaël GIRAUD et John K.–H. QUAH, juin 2003.

2003–**06** *Relativité de la satisfaction dans la vie : une étude sur données de panel.*
Théophile AZOMAHOU, Phu NGUYEN VAN, Thi Kim Cuong PHAM, juin 2003.

2003–**07** *A model of the anchoring effect in dichotomous choice valuation with follow–up.*
Sandra LECHNER, Anne ROZAN, François LAISNEY, juillet 2003.

2003–**08** *Central Bank Independence, Speed of Disinflation and the Sacrifice Ratio.*
Giuseppe DIANA, Moïse SIDIROPOULOS, juillet 2003.

2003–**09** *Patents versus ex–post rewards : a new look.*
Julien PÉNIN, juillet 2003.

2003–**10** *Endogenous Spillovers under Cournot Rivalry and Co–opetitive Behaviors.*
Isabelle MARET, août 2003.

2003–**11** *Les propriétés incitatives de l'effet Saint Matthieu dans la compétition académique.*
Nicolas CARAYOL, septembre 2003.

2003–**12** *The 'probleme of problem choice' : A model of sequential knowledge production within scientific communities.*
Nicolas CARAYOL, Jean–Michel DALLE, septembre 2003.

2003–**13** *Distribution Dynamics of CO_2 Emissions.*
Phu NGUYEN VAN, décembre 2003.

2004–**01** *Utilité relative, politique publique et croissance économique.*
Thi Kim Cuong PHAM, janvier 2004.

2004–**02** *Le management des grands projets de haute technologie vu au travers de la coordination des compétences.*
Christophe BELLEVAL, janvier 2004.

2004–**03** *Pour une approche dialogique du rôle de l'entrepreneur/manager dans l'évolution des PME : l'ISO comme révélateur …*
Frédéric CRÉPLET, Blandine LANOUX, février 2004.

2004–**04** *Consistent Collusion–Proofness and Correlation in Exchange Economies.*
Gaël GIRAUD, Céline ROCHON, février 2004.

2004–**05** *Generic Efficiency and Collusion–Proofness in Exchange Economies.*
Gaël GIRAUD, Céline ROCHON, février 2004.

2004–**06** *Dualité cognitive et organisationnelle de la firme fondée sur les interactions entre les communautés épistémiques et les communautés de pratique..*
Frédéric CRÉPLET, Olivier DUPOUËT, Francis KERN, Francis MUNIER, février 2004.

2004–**07** *Les Portails d'entreprise : une réponse aux dimensions de l'entreprise « processeur de connaissances ».*
Frédéric CRÉPLET, février 2004.

2004–**08** *Cumulative Causation and Evolutionary Micro–Founded Technical Change : A Growth Model with Integrated Economies.*
Patrick LLERENA, André LORENTZ, février 2004.

2004–**09** *Les CIFRE : un outil de médiation entre les laboratoires de recherche universitaire et les entreprises.*
Rachel LÉVY, avril 2004.

2004–**10** *On Taxation Pass–Through for a Monopoly Firm.*
Rabah AMIR, Isabelle MARET, Michael TROGE, mai 2004.

2004–**11** *Wealth distribution, endogenous fiscal policy and growth : status–seeking implications.*
Thi Kim Cuong PHAM, juin 2004.

2004–**12** *Semiparametric Analysis of the Regional Convergence Process.*
Théophile AZOMAHOU, Jalal EL OUARDIGHI, Phu NGUYEN VAN, Thi Kim Cuong PHAM, Juillet 2004.

2004–**13** *Les hypothèses de rationalité de l'économie évolutionniste.*
Morad DIANI, septembre 2004.

2004–**14** *Insurance and Financial Hedging of Oil Pollution Risks.*
André SCHMITT, Sandrine SPAETER, septembre 2004.

2004–**15** *Altruisme intergénérationnel, développement durable et équité intergénérationnelle en présence d'agents hétérogènes.*
Alban VERCHÈRE, octobre 2004.

2004–**16** *Du paradoxe libéral–parétien à un concept de métaclassement des préférences.*
Herrade IGERSHEIM, novembre 2004.

2004–**17** *Why do Academic Scientists Engage in Interdisciplinary Research ?*
Nicolas CARAYOL, Thuc Uyen NGUYEN THI, décembre 2004.

2005–**01** *Les collaborations Université Entreprises dans une perspective organisationnelle et cognitive.*
Frédéric CRÉPLET, Francis KERN, Véronique SCHAEFFER, janvier 2005.

2005–**02** *The Exact Insensitivity of Market Budget Shares and the 'Balancing Effect'.*
Gaël GIRAUD, Isabelle MARET, janvier 2005.

2005–**03** *Les modèles de type Mundell–Fleming revisités.*
Gilbert KOENIG, janvier 2005.

2005–**04** *L'État et la cellule familiale sont–ils substituables dans la prise en charge du chômage en Europe ? Une comparaison basée sur le panel européen.*
Olivia ECKERT–JAFFE, Isabelle TERRAZ, mars 2005.

2005–**05** *Environment in an Overlapping Generations Economy with Endogenous Labor Supply : a Dynamic Analysis.*
Thomas SEEGMULLER, Alban VERCHÈRE, mars 2005.

2005–**06** *Is Monetary Union Necessarily Counterproductive ?*
Giuseppe DIANA, Blandine ZIMMER, mars 2005.

2005–**07** *Factors Affecting University–Industry R&D Collaboration : The importance of screening and signalling.*
Roberto FONTANA, Aldo GEUNA, Mireille MATT, avril 2005.

2005–**08** *Madison–Strasbourg, une analyse comparative de l'enseignement supérieur et de la recherche en France et aux États–Unis à travers l'exemple de deux campus.*
Laurent BUISSON, mai 2005.

2005–**09** *Coordination des négociations salariales en UEM : un rôle majeur pour la BCE.*
Blandine ZIMMER, mai 2005.

2005–**10** *Open knowledge disclosure, incomplete information and collective innovations.*
Julien PÉNIN, mai 2005.

2005–**11** *Science–Technology–Industry Links and the 'European Paradox' : Some Notes on the Dynamics of Scientific and Technological Research in Europe.*
Giovanni DOSI, Patrick LLERENA, Mauro SYLOS LABINI, juillet 2005.

2005–**12** *Hedging Strategies and the Financing of the 1992 International Oil Pollution Compensation Fund.*
André SCHMITT, Sandrine SPAETER, novembre 2005.

2005–**13** *Faire émerger la coopération internationale : une approche expérimentale comparée du bilatéralisme et du multilatéralisme.*
Stéphane BERTRAND, Kene BOUN MY, Alban VERCHÈRE, novembre 2005.

2005–**14** *Segregation in Networks.*
Giorgio FAGIOLO, Marco VALENTE, Nicolaas J. VRIEND, décembre 2005.

2006–**01** *Demand and Technology Determinants of Structural Change and Tertiarisation : An Input–Output Structural Decomposition Analysis for four OECD Countries.*
Maria SAVONA, André LORENTZ, janvier 2006.

2006–**02** *A strategic model of complex networks formation.*
Nicolas CARAYOL, Pascale ROUX, janvier 2006.

2006–**03** *Coordination failures in network formation.*
Nicolas CARAYOL, Pascale ROUX, Murat YILDIZOGLU, janvier 2006.

2006–**04** *Real Options Theory for Law Makers.*
Marie OBIDZINSKI, Bruno DEFFAINS, janvier 2006.

2006–**05** *Ressources, compétences et stratégie de la firme : Une discussion de l'opposition entre la vision Porterienne et la vision fondée sur les compétences.*
Fernand AMESSE, Arman AVADIKYAN, Patrick COHENDET, janvier 2006.

2006–**06** *Knowledge Integration and Network Formation.*
Müge OZMAN, janvier 2006.

2006–**07** *Networks and Innovation : A Survey of Empirical Literature.*
Müge OZMAN, février 2006.

2006–**08** *A.K. Sen et J.E. Roemer : une même approche de la responsabilité ?*
Herrade IGERSHEIM, mars 2006.

2006–**09** *Efficiency and coordination of fiscal policy in open economies.*
Gilbert KOENIG, Irem ZEYNELOGLU, avril 2006.

2006–**10** *Partial Likelihood Estimation of a Cox Model With Random Effects : an EM Algorithm Based on Penalized Likelihood.*
Guillaume HORNY, avril 2006.

2006–**11** *Uncertainty of Law and the Legal Process.*
Giuseppe DARI–MATTIACCI, Bruno DEFFAINS, avril 2006.

2006–**12** *Customary versus Technological Advancement Tests.*
Bruno DEFFAINS, Dominique DEMOUGIN, avril 2006.

2006–**13** *Institutional Competition, Political Process and Holdup.*
Bruno DEFFAINS, Dominique DEMOUGIN, avril 2006.

2006–**14** *How does leadership support the activity of communities of practice ?*
Paul MULLER, avril 2006.

2006–**15** *Do academic laboratories correspond to scientific communities ? Evidence from a large European university.*
Rachel LÉVY, Paul MULLER, mai 2006.

2006–**16** *Knowledge flows and the geography of networks. A strategic model of small worlds formation.*
Nicolas CARAYOL, Pascale ROUX, mai 2006

La présente liste ne comprend que les Documents de Travail publiés à partir du 1er janvier 2000. La liste complète peut être donnée sur demande.
This list contains the Working Paper writen after January 2000, 1rst. The complet list is available upon request.