



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

Documents de travail

« Coordination failures in network formation »

Auteurs

**Nicolas CARAYOL
Pascale ROUX
Murat YILDIZOGLU**

Document de travail n° 2006-03

Janvier 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

Coordination failures in network formation¹

Nicolas Carayol ^{$\phi\tau^2$} , Pascale Roux ^{$\phi\tau$} , Murat Yıldızoğlu ^{μ} ,

^{ϕ} BETA, CNRS - Université Louis Pasteur (Strasbourg 1)

^{τ} ADIS, Faculté Jean Monnet, Université Paris Sud (Orsay-Paris 11)

^{μ} E3i, IFRéDE-GRES, Université Montesquieu Bordeaux IV

October 29, 2005

¹The authors thank Herbert Dawid for helpful discussion. The participants of the ACEPOL05 Workshop, ZiF, Bielefeld and of the I-neck meeting, BETA, Strasbourg should also be acknowledged. Murat Yıldızoğlu gratefully acknowledges the support of the CCRDT program of Aquitaine Region.

²Corresponding author. BETA, Université Louis Pasteur, 61 avenue de la Forêt Noire, F-67085 Strasbourg Cedex; tel: +33-490244069; fax: +33-390242071, <carayol@cournot.u-strasbg.fr>.

Abstract

In this paper, we make an exploratory use of numerical techniques (genetic algorithms and Monte Carlo simulations) to compute efficient and emergent networks in a spatialized version of the connections model of Jackson and Wolinski (1996). This approach allows us to observe and discuss the coordination failures that arise in a strategic network formation context with link-mediated positive externalities to connections and geographically based connection costs. Our results highlight that, depending on the strength of the externalities, emergent and efficient networks may share several structural properties. Nevertheless, emergent networks have too few local and distant connections and are also too less “coordinated” around some central agents than they should.

Keywords: Strategic Network Formation; Efficiency; Stability; Coordination; Small Worlds; Genetic Algorithms; Monte Carlo Simulations

JEL codes: D85, C63, D62, Z13

1 Introduction

In the last decade, networks structure characterization has been the object of a high and growing interest in various fields. Much effort have been dedicated to understand the topology of real networks. More recently, some studies concentrate on the dynamic processes that determine such topologies. In economics, some authors, recognizing that both individual and collective behaviors and performances are grounded in networks, focus on the micro-behaviors that drive network formation. In the ACE literature, the emergent properties of networks have been studied by Kirman and Vriend (2001) and Tesfatsion (1997) who model the formation of trade networks among strategically interacting buyers and sellers. These agents choose their partners adaptively, on the base of their past experiences with these partners^{1,2}. The aim of such computational approaches is to study complex dynamic systems of interacting agents.

A more theoretical and analytical economic literature on network formation builds upon the seminal contributions of Aumann and Myerson (1988) and of Jackson and Wolinski (1996). Two main questions are central in this literature (Jackson, 2004): Which networks are likely to form when agents choose their connections in order to maximize given individual payoffs structures? And how efficient are networks that emerge from self-interested agents' choices? The first stylized economic model that tackles those two questions is the so-called *connections model* introduced by Jackson and Wolinski (1996). The very simple and realistic specification of the individual payoffs allows the authors to obtain systematic analytical results on graphs efficiency and partial on networks stability. This allows them to show that often efficient networks are unstable. More recently two articles have extended this model in order to study the dynamics of network emergence (*see* Watts (2001) for deterministic dynamics and Jackson and Watts (2002) for stochastic dynamics). In those contributions, the analytical computation of (possibly numerous) emerging networks becomes difficult to handle. More generally, this literature faces important difficulties to generate and discuss non trivial network configurations. Indeed, network structures they analyze are very simple (complete network, empty network, complete star) and have little in common with real social or economic networks. Some frequent features of real social networks are short average distance between agents, high clustering (i.e. there is a high probability that two agents to be neighbors if they have neighbors in common) and heterogenous neighborhood sizes among agents (Watts and Strogatz, 1998; Albert and Barabási, 1999). These properties altogether characterize the so-called *Small World* phenomenon³.

Very recently, Carayol and Roux (2003) and Jackson and Rogers (2005) propose variations of the connections model by giving different forms of geographic locations to individuals and introducing complexities in individual payoff functions through spatial costs of direct link formation⁴. Their aim is to find the minimal specifications that lead to emerging networks that are much richer and that tend to correspond to those empirically observed social networks. In their spatialized connections model, Carayol and Roux (2003) obtain, in a dynamic setting and for a wide set of parameters, endogenous networks that exhibit the Small World properties (i.e. highly clustered connection structures and

¹For a survey on ACE models studying network formation, one can refer to Tesfatsion (2003).

²A similar approach is also used by Dupoët and Yıldızoğlu (2005) for studying the emergence of a particular type of network: communities of practice.

³Evidences concern for instance networks of firm board members (Davis and Greve, 1996), or networks of scientific papers co-authorship (Barabási et al., 2001; Newman, 2001).

⁴Johnson and Gilles (2000) first introduced such spatialized connections model with linear geographical distance.

short average path length). We call *emergent networks* these connection structures since they endogenously result from the connection establishment mechanism assumed in this setup⁵. Carayol and Roux also observe that, for certain values of the parameters, emergent networks are globally and locally too weakly connected since the agents don't want to individually support the costs of social welfare improvement. Nevertheless, their results are limited and partial since, even for a relatively small set of agents, it becomes very difficult to compute network efficiency both analytically and numerically⁶. Indeed, one cannot appreciate the extent to which emerging networks are efficient and whether they are structurally different from the optimal networks. Carayol, Roux and Yıldızoğlu (2005) propose to use Genetic Algorithm (GA) techniques to compute *efficient networks*. They test and calibrate this technique on two simple models among which the connections model⁷ (for which the efficient networks are fully known) and show that the proposed method is quite robust for computing the efficient networks in these models.

In the present paper, we make an exploratory use of such technique to compute the efficient networks in the spatialized connections model. This allows us to compare for the first time the (GA) efficient networks to the emergent networks in this model. Thus we can fully discuss the coordination failures that may arise in the network formation context. Our aim is to provide a general method for exploring the efficient and equilibrium networks. The results of this exploration should be very useful to any ACE model that contains a network formation component. The modeler can get some precious insights on the possible results of his/her model and on the consequences of the different assumptions on network formation. In our context, we show for example that the emergent networks are less dense than the ones that maximizes social wealth. This clearly corresponds to the economic intuition: because agents benefit from indirect connections, there are positive indirect externalities to bond formation. Therefore agents naturally build less links than they should. Moreover the emergent networks are found to be less "coordinated" than they should. Indeed the supplement of connections observed in the efficient networks is preferentially attributed to one or several agents who have central positions in the networks. These connections allow a more efficient distribution of wealth among all agents. Emergent networks do not share such structural property because agents do not want to play a central and costly position. Agents would benefit from mutualizing the costs for increasing the connectivity of one (or several) of them so as to enhance the quality of the indirect connections in the network. Our results underline the difficulties of coordination in strategic network formation and the necessity of public policies (when they are possible) in order to sustain the emergence of central agents in networks.

The paper is structured as follows. The next section begins with some basic definitions on network formation literature and presents the variant of the connections model developed in Carayol and Roux (2003) and their approach for computing emergent networks in a dynamic setting. Section 3 synthetically presents the Genetic Algorithms approach developed in Carayol, Roux and Yıldızoğlu (2005) and its performance in determining efficient networks. Section 4 compares the emergent and

⁵This concept should not be confused with the concept of *emergence* used in the Complex Adaptive System literature, as well as in the ACE approach.

⁶Even for a relatively small numbers of players, the number of possible networks becomes very large. Johnson and Gilles (2000) observe that the number of possible networks for n agents is $\sum_{k=1}^{c(n,2)} c(c(n,2), k) + 1$ where, for every $k \leq n$, $c(n, k) := n! / (k!(n - k)!)$. For example, when $n = 8$, the number of possible networks exceeds 250 million.

⁷The other one is the co-author model of Jackson and Wolinsky (1996).

the efficient networks. Finally, Section 5 concludes.

2 Network formation: the model

In this section, we begin by introducing the notation and the basic notions for studying networks and their efficiency. We limit our attention to the case of non-directed graphs, where bonds are symmetric and built on mutual consent, as it occurs in many real social networks. Then, we introduce the dynamic pertubated process that leads to network formation. Finally, we present the model.

2.1 Basic notions on graphs

We consider a fixed and finite set of n agents, $N = \{1, 2, \dots, n\}$ with $n \geq 3$. Let i and j be two members of this set. Agents are represented by the nodes of a non-directed graph, which's edges 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 unordered pairs of connected and distinct agents. Formally, $\{ij\} \in g$ means that the link 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 all players are connected to 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.

Then 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 neighbors agent i has, 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 $\eta(g) = \#g = \frac{1}{2} \sum_{i \in N} \eta_i(g)$.

A *path* connecting i to j in a non empty graph $g \in G$, is a sequence of edges between distinct agents such that $\{i_1i_2, i_2i_3, \dots, i_{k-1}i_k\} \subset g$ where $i_1 = i$, $i_k = j$. 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 the graph g . The set of *shortest paths* between i and j on g noted $i \overset{\sim}{\longleftrightarrow}_g j$ is such that $\forall k \in i \overset{\sim}{\longleftrightarrow}_g j$, we have $k \in i \longleftrightarrow_g j$ 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 the shortest path between them: $d(i, j) = d_g(i, j) = \#k$, with $k \in i \overset{\sim}{\longleftrightarrow}_g j$. When there is no path between i and j , 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 g .

An external metrics can also be introduced, representing for example the geographic position of agents (Johnson and Gilles, 2000). Such external metrics defines a new distance operator denoted $d'(i, j)$. Following Carayol and Roux (2003), we consider further that agents are located on a circle (or a ring). 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 $d'(i, j) = \min \{|i - j|; n - |i - j|\}$.

Several typical graphs can be described. Let $i \neq j \in N$. First of all, the *empty graph*, denoted g^\emptyset , is such that it does not contain any links. The *ring* g° is a regular network of order $k = 1$, in which all agents are connected and only connected with their two closest geographic 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 every node can be the center.

2.2 Value and efficiency of networks

Traditionally in economics, efficiency refers to a state from which no agent's payoffs can be improved without deteriorating the payoff of at least one other agent. In the context of network efficiency, this property, which corresponds to the Pareto efficiency, means that a network is efficient when it does not exist another network that leads to a higher payoff for at least one individual, without deteriorating the payoff of other agents. Since the pioneering work of Jackson and Wolinski (1996), a 'strongest' notion of efficiency is preferred in the economic literature on networks formation.

Let $\pi_i(g)$ be the net individual payoff that the agent i receives from maintaining his position in the network g , with $\pi_i : \{g \mid g \subseteq g^N\} \rightarrow \mathfrak{R}$. The network social value, denoted $\pi(\cdot)$, can be computed by simply summing individual payoffs. The total value of a graph g , with $\pi(\emptyset) = 0$ is thus given by:

$$\pi(g) = \sum_{i \in N} \pi_i(g). \quad (1)$$

A network is then said to be efficient if it maximizes this sum. The formal definition follows.

Definition 1 *A network $g \subseteq g^N$ is said to be efficient if it maximizes the value function $\pi(g)$ on the set of all possible graphs $\{g \mid g \subseteq g^N\}$ i.e. $\pi(g) \geq \pi(g')$ for all $g' \subseteq g^N$.*

It should be noticed that several networks can lead to the same maximal total value. For example, if we consider strictly homogenous agents, any isomorphic graph of an efficient network is also efficient.

2.3 Network formation: pairwise stability and dynamics

We turn now toward the stability of graphs. 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. In such a process, the formation of a link between two agents requires the consent of both of them, but not its deletion, which can unilaterally emanate from one of them. The formal definition of this notion is the following.

Definition 2 *A network $g \subseteq g^N$ is said to be 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 $\pi_j(g + ij) < \pi_j(g)$.*

In the present paper we are interested in the dynamic network formation presented in Jackson and Watts (2002) which is consistent and encompasses the stability notion presented above. It roughly corresponds to the following scheme. At each period, two agents $i, j \in N$ are randomly chosen with the same probability $p_{ij}^t = p^t > 0$. They can decide to form, maintain or break links. Let's assume that agents are myopic: they take their decisions on the basis of the immediate impact of links on their current payoffs. If these agents are already connected, they consider whether they may unilaterally sever the link or bilaterally keep it. If they are not directly connected, they consider whether they should add this connection or stay disconnected. Formally, the dynamic process can be described as follows:

- i) if $ij \in g_t$, the link is saved if and only if $\pi_i(g_t) \geq \pi_i(g_t - ij)$ and $\pi_j(g_t) \geq \pi_j(g_t - ij)$,
- ii) if $ij \notin g_t$, a link is created if and only 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 at least one of the two agents.

Following Jackson and Watts (2002), we then introduce small random perturbations ε 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 ε , it can be shown that the discrete-time Markov chain becomes irreducible and aperiodic and has thus a unique corresponding stationary distribution (μ^ε). Such perturbed stochastic processes are said to be ergodic. Intuitively ergodicity implies that 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 state of the system to become independent of its initial conditions.

Usually, the modeler let $\varepsilon \rightarrow 0$ (once the long run is reached) in order to restrict the number of states selected in the long run. State z is said to be a stochastically stable state (Young, 1993) if it has a non null probability of occurrence in the stationary distribution: $\lim_{\varepsilon \rightarrow 0} \mu^\varepsilon(z) > 0$. In the network formation context, Jackson and Watts (2002) show that stochastically stable networks are either pairwise stable or part of a closed cycle⁸ (of the unperturbed process). 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. Therefore we propose a slightly different regime for the perturbation process. We let the error term decrease in time according to the following simple rule:

$$\varepsilon^t = \begin{cases} \varepsilon & \text{if } t < T \\ 1/t & \text{otherwise} \end{cases}, \quad (2)$$

with $\varepsilon > 0$ the initial noise and T some finite time. This rule ensures that the noise does affect the dynamics while it decreases down to zero when time increases with $\lim_{t \rightarrow \infty} \varepsilon^t = 0$. It also preserves the ergodicity property of the system. Notice that 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, 2005). Therefore we use Monte Carlo experiments to approximate the unique limiting stationary distribution (of networks) of the perturbed dynamic process presented above. The experiments are stopped at $t = 10,000$, date after which the process is proven to have almost surely stabilized on a given pairwise stable state⁹. If not, the process still goes until it reaches one. Our *emergent networks*¹⁰ hence correspond to the support of this limiting stationary distribution.

2.4 The spatialized “Connections Model”

In the *Connections model* introduced by Jackson and Wolinski (1996), links represent individuals' relationships (for example, friendships). In such a context, agents benefit from their direct and indirect connections, through the relational network of their partners. But, the communication is not perfect: the positive externality deteriorates with the relational distance of the connection. Formally,

⁸A closed cycle is a set of networks that may be reached from any one of them without errors and that cannot lead to any other network.

⁹See Carayol and Roux (2003) for more precise time series analysis of the network formation process.

¹⁰This concept should not be confused with the concept of *emergence* used in the Complex Adaptive System literature, as well as in the ACE approach.

there is a decay parameter which represents the quality of links used for information flows. Moreover, agents bear costs for maintaining direct connections. As a consequence, agents try to maximize the value generated from direct and indirect connections, avoiding superfluous connections. This simple specification of the individual payoffs allows the authors to obtain systematic analytical results on graphs' efficiency and partial results on networks' stability. Nevertheless, the efficient and stable network structures they discuss are very simple (complete network, empty network, complete star) and not very realistic.

In order to obtain emerging networks which tend to correspond to the empirically observed social or economic networks, Carayol and Roux (2003) propose a variation of the connections model of Jackson and Wolinski (1996), by giving geographic locations to individuals and introducing complexities in individual payoff functions through spatial costs for direct links formation. Let's assume that agents are located on a circle (or a ring). The net profit received by any agent i is now given by the following expression¹¹:

$$\pi_i(g_t) = \sum_{j \in N \setminus i} \delta^{d(i,j)} - c \sum_{j:ij \in g_t} d'(i,j), \quad (3)$$

where $d(i,j)$ is the geodesic distance between i and j . $\delta \in]0;1[$ is the decay parameter and $\delta^{d(i,j)}$ 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 unity. Notice that if there is no path between i and j , then $d(i,j) = \infty$ and thus $\delta^{d(i,j)} = 0$. The second part of the right-hand side of the equation describes the costs of direct links. $c \in]0;1[$ is a parameter which gives the costs that agents have to bear for each of their direct connection. $d'(i,j)$ gives the geographic distance between any two agents on the external metrics (a circle) we consider. Thus costs increase linearly with the geographic distance separating neighboring agents.

3 Efficient networks: the Genetic Algorithms approach

Searching for efficient network structures is in general a difficult analytical task. But, once the pay-off structure is well defined in relation with the connection structure, one is tempted to explore this question using more heuristic strategies. As a matter of fact, the connection structure of the network can be expressed as a matrix of bits (1 for connection or 0 for absence of connection) and the pay-off structure can assign a value to each of such matrices. The search for efficient networks can hence be seen as an optimization problem in the connection-matrix space, *i.e.* the space of all possible networks. This optimization problem yields analytical solutions only for simple pay-off structures. We examine here a numerical tool for optimization: genetic algorithms (**GA**) that have proved their efficacy in optimization problems where the potential solutions can be represented as binary strings. Our networks can effectively be quite easily represented as binary strings.

¹¹Johnson and Gilles (2000), relying on Debreu's (1969) hypothesis according to which closely located players incur lower costs to sustain communications, first introduced spatial costs for direct link formation with a linear geographical distance. We thus obtain the same payoffs specification as theirs with a circle as an external metrics instead of a line.

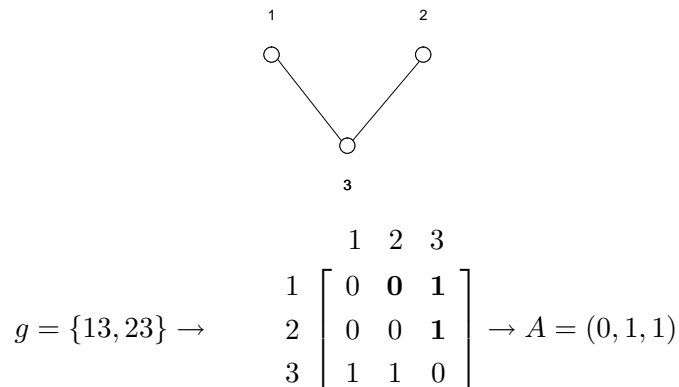
3.1 Representing networks as binary strings

Our problem is to find the network g which maximizes social value π as given by the equation 1 over the set of all possible networks G . In order to use the GA for this optimization problem, we need to represent our networks as binary strings (sequences of bits – 1 or 0).

Consider first that any network with n agents (whether directed or not, eventually with self-connections) can, without loss of generality, be represented by a connection matrix of size $n \times n$ of binary elements. Given that all networks we consider are undirected (i is connected to j iff j is also connected to i) and that self-connections are excluded, the upper triangular part of this connection matrix, excluding the diagonal, provides complete information on the network structure. As a consequence, the vector composed by all the connection *bits* of this upper triangular part in some conventionally chosen order sums up the network structure. Thus for a network of n agents, this vector is a binary string of length $l = (n^2 - n) / 2$.

From the point of view of a genetic algorithm, undirected networks can hence be formally represented as *chromosomes* defined as sequences of binary elements: $A = (a_1, a_2, \dots, a_l)$ with $a_i \in \{0, 1\}, \forall i \in \{1, 2, \dots, l\}$.

In the example below with $n = 3$ agents, the undirected network $g = \{13, 23\}$ is fully characterized by the chromosome $A = (0, 1, 1)$, which's length is $l = (3^2 - 3) / 2 = 3$.



Once we represent it, we can compute the value of a connection matrix (its *fitness*) using the equation 1 and utilize the Genetic Algorithms to search for matrices with the highest value.

3.2 Genetic Algorithms: How do they work?

Genetic algorithms (**GA**) are numerical optimization techniques developed by John Holland (*see for example Holland (2001), which has initially been published in 1975*). GA transpose to other problems the strategies that the biological evolution has successfully used for *exploring* complex fitness landscapes. The search for an optimum by a GA corresponds to the evolution of a population of candidate solutions through *selection*, *crossover* (combination) and *mutation* (random experiments). The GA have been used for solving a very large set of problems directly, or indirectly as a component of a

classifier system. Goldberg (1991) gives quite an exhaustive account of the characteristics of the GA and of their applications. For applications of the GA as a learning algorithm, see Yıldızoğlu (2002).

<i>Population at date t</i>	<i>Fitness: f(x)=x²</i>	<i>Expected number: f(x)/Mean_f</i>	<i>Effective number proportionally drawn</i>	<i>New population</i>
1) 00011 (=3) 2) 01100 (=12) 3) 10111 (=23)	1) 9 2) 144 3) 529 Mean _f =227	1) 0.0 2) 0.6 3) 2.4 Sum=3	1) 0 2) 1 3) 2 Sum=3	1) 10111 (=23) 2) <u>01100</u> (=12) 3) 10111 (=23)
<i>Crossover (3)-(2) at bit 3</i>				
1) 10111 (=23) 2) <u>01111</u> (=15) 3) 10100 (=20)	<i>Mutation (1)</i>			
	1) 11111 (=31 *) 2) 01111 (=15) 3) 10100 (=20)	<i>Population at date t+1</i>		
		1) 11111 (=31) 2) 01111 (=15) 3) 10100 (=20)		

Simple application of GAs to optimisation of the function $f(x)=x^2$ over the interval 0-31. Integers are coded with five bits binary code: 00001=1, 11111=31. The example uses an initially random population of 3 members and the GA constructs a new population through selective reproduction, combination (crossover) and random experiments (mutation). In this schematic example, the GA attains the optimum (31) in one period. For each string, the crossover, its position and the partner, as well as mutation position are chosen randomly. The mutation bit simply switches its value: 0->1 or 1->0. **This process is controlled by: population size, bit-string size, probability of crossover and probability of mutation.**

Figure 1: A simple example of genetic algorithm

The *canonical genetic algorithm* makes evolve a population of binary strings (chromosomes composed of 1 and 0). The size of the population m is given. It is the source of one of the strengths of the GA: implicit parallelism (the exploration of the solution space using several candidates in parallel). The population of chromosomes at step t (a generation) is denoted $P(t) = \{A_j\}_t$ with $\#P(t) = m$, and $\forall t = 1, 2, \dots, T$ with T the given total number of generations. Notice that T is the other source of the strengths of the GA. The algorithm (randomly) generates an initial population $P(0)$ of candidate chromosomes which are evaluated at each period using the fitness (value) function. They are used for composing a new population at the next period $P(t+1)$. For illustrative purposes, Figure 1 gives a deliberately trivial example of optimization by GA. Each chromosome has a probability of being **selected** that is increasing in its fitness. The members included in the new population are recombined using a **crossover** mechanism (see Figure 2). The crossover operation introduces controlled innovations in the population since it combines the candidates already selected in order to invent new candidates with a potentially better fitness. Moreover, the **mutation** operator randomly modifies the candidates and introduces some random experimenting in order to more extensively explore the state space and escape local optima. Typically, the probability of mutation is rather low in comparison with the probability of crossover because otherwise the disruption introduced by excessive mutations can destruct the hill-climbing capacity of the population. Finally, an **elitism** operator can be used which ensures that the best individual of a population will be carried to the next generation¹².

¹²See Dawid (1999) and Michalewicz (1996) for general properties of genetic algorithms.

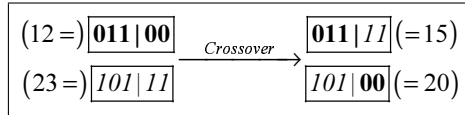


Figure 2: A simple example of crossover operation

4 Emergent and efficient networks: coordination failures in network formation

We present our research protocol in the next subsection. The second subsection exposes the first results on efficient networks and the last one compares the emergent networks to GA efficient networks.

4.1 Simulation protocol : Numerical settings, indicators and controls

The research protocol combines the techniques first used and presented in Carayol and Roux (2003) for generating emergent networks with Monte Carlo experiments and in Carayol, Roux and Yıldızoğlu (2005) for computing efficient networks with GA. Standard indicators are used for characterizing their structures and thus comparing them. In order to allow unbiased comparisons (since the density of the computed networks can introduce a bias), we correct these indicators using measures from comparable (control) random networks.

4.1.1 Numerical settings

Our numerical experiments all correspond to $n = 20$ agents. For simplification purposes, we consider that the unit cost in the payoff function (3) is fixed as $c = \frac{2}{n}$ (0.1 in our case). For the generation of both the GA-efficient networks and the emergent networks, all experiments are performed with randomly drawn values of δ over the value space $]0, 1[$.

The first series of experiments relate to the computation of the GA-efficient networks. The Java *JGAP*¹³ library is used to implement the GA based on binary chromosomes. The GA that we use is *elitist* and its probabilities of crossover and mutation are both computed by *JGAP*¹⁴. Carayol, Roux and Yıldızoğlu (2005) test the relevance of the GA as a search algorithm for efficient networks in the two stylized models for which the efficient networks are known. It establishes the robustness of the GA using an extensive set of Monte Carlo simulations. When $n = 20$, we know that the GA performs correctly with a population of chromosomes of size $m = 500$ evolving over $T = 500$ generations. We will use these numerical values in the present paper. A fixed number (500) of random draws of δ are performed in order to reasonably cover the parameter space $]0, 1[$. In this article, we slightly adapt the simulation protocol in order to increase the robustness of our results. For each configuration, we now fully run three times the GA in order to obtain three final candidate networks among which we keep the one that generates the highest social value as the *GA-efficient network*.

The second series of numerical experiments are designed to search the emerging networks (the networks which are on the unique limiting stationary distribution of the perturbed dynamic process

¹³<http://jgap.sourceforge.net/>

¹⁴Probability of crossover is 0.5 and the probability of mutation is 1/15.

of Jackson and Watts, 2002) developed in Section 2. We run 1,000 experiments with randomly drawn values of $\delta \in]0, 1[$. All experiments are stopped at $t = 10,000$, date after which the process is proven to have almost surely stabilized on a given pairwise stable state.

4.1.2 Indicators

Several indicators are used in order to provide a synthetic characterization of the structural properties of networks.

Density of network. We compute the density of the network as follows:

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

This indicator hence corresponds to the average degree in the network.

Average distance. (or Average Path Length) We compute the average distance of (directly or indirectly) connected agents. It is given by

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

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 is verified and zero otherwise.

Average Clustering. (or Average Cliquishness as it is often referred to in the physics of networks literature) The average clustering indicates to what extent the neighborhoods of connected agents overlap. It is given by

$$c(g) = \frac{1}{n} \sum_{i \in N} \sum_{jl} \frac{1\{j, l \in N_i(g); j \neq l; j \in N_l(g)\}}{\#\{lj \mid j \neq l; j, l \in N_i(g)\}}. \quad (6)$$

It is the frequency with which agents' neighbors are also neighbors together.

Next, we intend to know more on the distribution of neighborhoods size in the population.

Global asymmetry. It is computed using the difference between the largest neighborhood size and the lowest one in the network:

$$r(g) = \max_{i \in N} \eta_i(g) - \min_{j \in N} \eta_j(g). \quad (7)$$

This indicator measures the global asymmetry of neighborhood sizes in the network.

Last we are also interested in the extent of neighborhood asymmetry between directly connected agents.

Local asymmetry. It is computed as the sum over all direct connections of the absolute value of the difference between neighborhood sizes. That is:

$$u(g) = \frac{1}{\eta(g)} \sum_{ij \in g} |\eta_i(g) - \eta_j(g)|. \quad (8)$$

This indicator gives the propensity of the highly connected agents to be linked to agents that have few connections. It can be understood as a measure of non assortativity of connections as regard agents' neighborhood sizes.

4.1.3 Controls for density

The last four indicators presented above are affected by the density of the network (the first indicator) that is likely to vary with both δ and the generating process (GA or stochastic stability). Therefore, these indicators are somehow biased and we must find an efficient control for density. We propose to build, for each generated network, *control random graphs* which have exactly the same number of agents and connections (thus the same density). Such networks are simply built by allocating the given number of edges to randomly chosen pairs of agents¹⁵. For each given number of edges and agents, the four above mentioned indicators are then numerically computed for 1,000 of such randomly drawn networks. The average of the observations is used as the control value. For instance, instead of looking at $r(g)$, where g is a given selected stable network, we will look at the *ratio* $r(g)/r(g^{rd})$, where $r(g^{rd})$ is the average global asymmetry of neighborhood sizes over a set of 1,000 random networks that have exactly the same density as g . Each of the indicators is corrected using the corresponding ratio.

This method allows us (i) to analyze the structural properties of efficient networks given their density; (ii) to compare efficient and emergent networks for the different values of δ while controlling for their density. The results are given in the following sections.

4.2 The structural properties of GA-efficient networks

We find that the GA-efficient networks are globally only 6% longer than random graphs, and in an intermediary region of δ values ($\delta \in [0.5, 0.7]$), they are even 3% shorter (see the ratio $d(g)/d(g^{rd})$ in Table 1). This is noticeably low since, in random graphs, the average distance between any two indirectly connected agents is already known to be very “short”. In the meantime, on the global level (see the last column of Table 1), the clustering ratio of the efficient networks is close to 2, which means that efficient networks are nearly twice clustered (94% more) than their controls (the random networks). The conjunction of these two characteristics qualifies the efficient networks as Small Worlds in the sense of Watts and Strogatz (1998), a property that many real networks do share.

Proposition 1 *When δ is neither close to 0 nor close to 1, efficient networks have an average distance (path length) similar to the one of the control random networks while they are significantly more clustered than these random networks. In this sense they correspond to Small Worlds.*

As regard the two other indicators, an interesting contrast arises. The efficient networks have 9% less global asymmetry of neighborhoods than their control random graphs, while they have 45% more local asymmetry of neighborhoods. This indicates that, though the global asymmetry is quite similar (slightly lower) to a case where the connections would be simply allocated at random, agents with fewer links have a significantly higher chance to be directly connected to agents who have many connections.

Proposition 2 *Efficient networks tend to provide central positions to some agents in a similar extent as random networks, but, in the first case, these central agents are more likely to be connected to agents with few links.*

¹⁵In the random graphs literature, such networks are known as random graphs of the kind of Erdős and Rény (1960).

Ratios	$g \setminus \delta$	0.1 – 0.5	0.5 – 0.7	0.7 – 0.85	0.85 – 0.95	0.95 – 1	0.1 – 1
$\frac{d(g)}{d(g^{\text{rd}})}$	GA efficient	1.19 (0.17)	0.97 (0.00)	0.94 (0.00)	0.90 (0.00)	1.21 (0.14)	1.06 (0.09)
	Emergent	1.51 (0.29)	1.01 (0.00)	1.01 (0.00)	1.10 (0.01)	1.86 (0.06)	1.23 (0.20)
$\frac{c(g)}{c(g^{\text{rd}})}$	GA efficient	2.23 (1.64)	1.54 (0.03)	1.70 (0.09)	2.03 (0.54)	0.37 (0.35)	1.94 (0.91)
	Emergent	1.60 (1.86)	1.34 (0.01)	1.13 (0.09)	0.55 (0.28)	0.00 (0.00)	1.25 (1.07)
$\frac{r(g)}{r(g^{\text{rd}})}$	GA efficient	0.64 (0.09)	0.75 (0.04)	1.09 (0.07)	2.10 (0.50)	2.10 (0.32)	0.91 (0.31)
	Emergent	0.24 (0.03)	0.40 (0.00)	0.50 (0.02)	0.54 (0.01)	0.58 (0.04)	0.37 (0.04)
$\frac{u(g)}{u(g^{\text{rd}})}$	GA efficient	1.01 (0.24)	0.98 (0.12)	1.60 (0.36)	4.10 (6.87)	2.81 (1.43)	1.45 (1.80)
	Emergent	0.36 (0.07)	0.39 (0.01)	0.48 (0.02)	0.60 (0.02)	0.71 (0.10)	0.43 (0.06)

Table 1: Average values (and variance in parentheses) of the various indexes computed on GA-efficient networks and emergent networks relative to their control random networks drawn for various regions of δ . The region for which $\delta < 0.1$ is never taken into considerations because both efficient and emergent networks are empty and thus most indexes are not computable.

The most connected agents reduce the distance between all agents and thus contribute to increasing of social wealth. The need for such global coordination is considerably high when $\delta \approx 0.9$, where complete stars tend to become efficient. The superiority of such networks in terms of social surplus decreases sharply when delta becomes closer to 1, because the connections of various lengths tend to provide the same wealth (the decay phenomenon becomes negligible).

Proposition 3 *When $\delta \approx 0.9$, efficient networks provide central positions to a few agents (or even to only one) in social networks, which tend to be a complete star.*

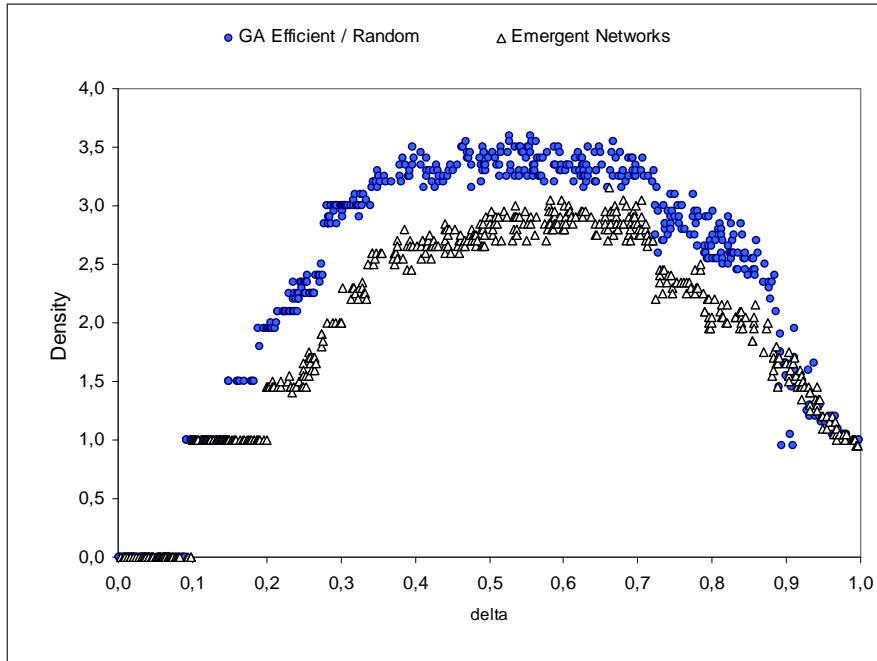


Figure 3: The density of GA-efficient networks and emergent networks for the various values of δ .

4.3 Comparing emergent networks and efficient networks

The first series of results relate to networks density. Apart from the regions where δ is either close to 0 or close to 1, the density of selected stable networks is lower than the density of efficient networks (see Figure 3). Agents generate less connections than they should as regard social surplus. This clearly confirms the economic intuition that arises from the basic payoff function of the connections model: there are positive externalities to link formation since his neighbors, and potentially all agents to whom a given agent is directly and indirectly connected, may benefit from any new connections that this agent would establish. Therefore, it is not surprising that selfish agents, who do not take into account the social returns of link formation, establish too few connections.

Proposition 4 *When δ is neither close to 0 nor close to 1, the density of emergent networks is lower than the density of efficient networks; agents generate less connections than they should as regard social surplus.*

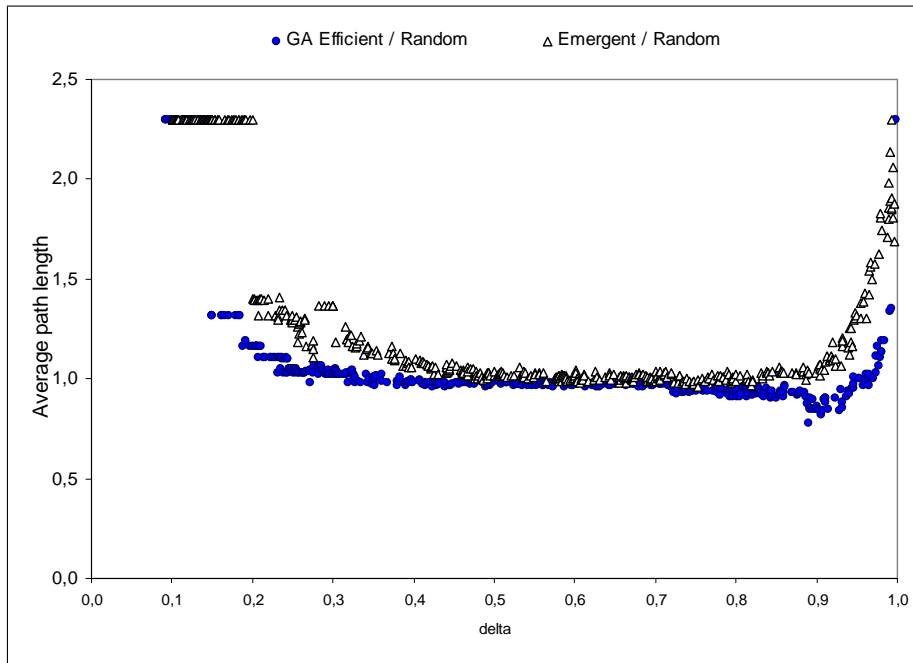


Figure 4: The average distance $d(\cdot)$ of GA-efficient networks and emergent networks for the various values of δ relative to the average distance of their control random networks

The relational distances (see Figure 4) between agents in the network constitute key factors for wealth generation since they directly intervene in the payoff function. We find that such distances are in average significantly longer (controlling for density) in selected stable networks than in efficient networks, when δ is either low¹⁶ or high (when $\delta \lesssim 0.45$ or $\delta \gtrsim 0.80$ see Figure 4 and Table 1). In these extreme regions of δ , agents do not generate (geographically) long distance costly shortcuts. When δ is low, the rewards of such connections are too low as regard the costs. When δ is high, the difference between rewards from (relational) shorter and longer connections become too low as compared to the costs. It is only for intermediary regions of δ that agents are provided with nearly

¹⁶This indicator is not defined when δ is lower than 0.1, that is when the network is empty.

sufficient incentives to bear the costs associated with long distance connections. In particular, for $\delta \in [0.5, 0.7]$ the emergent networks are in average only 1% longer than control random graphs while efficient networks are slightly shorter than their own control random graphs.

Proposition 5 *While the average distance of emergent networks is globally longer than the one of efficient networks, there is an intermediary region of δ ($[0.5, 0.7]$) for which the average distance of emergent and efficient networks becomes very close to the average distance of their control random graphs.*

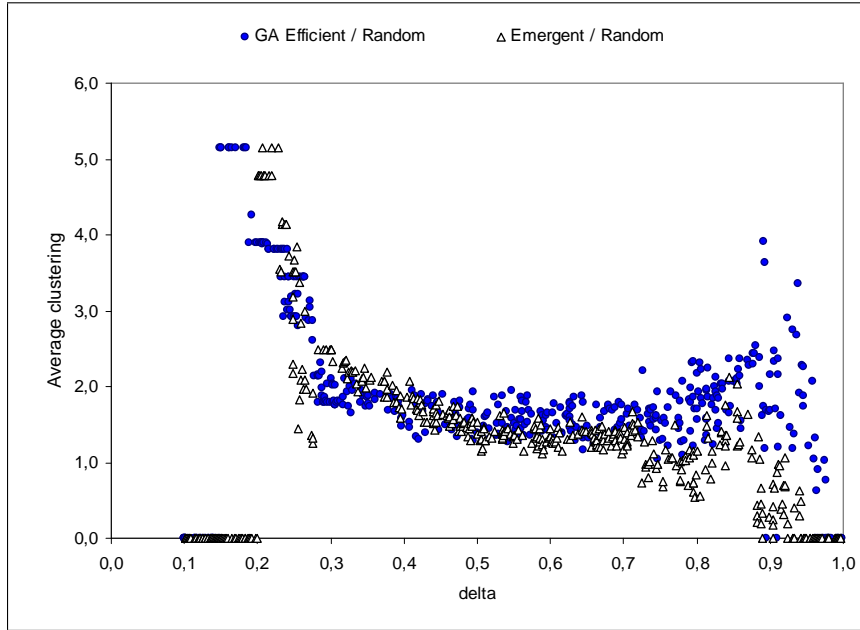


Figure 5: The average cliquishness (clustering) $c(\cdot)$ of GA-efficient networks and emergent networks for the various values of δ relative to the average distance of their control random networks.

What does the payoff function of the spacialized connection model imply in terms of clustering? Except for extreme values of δ , there is no simple answer to this question. On the one hand, the connections model provides positive externalities that are conveyed through the network: agents benefit from other agents with whom they are indirectly connected. Such externalities are the highest at distance 2. Thus agents have low incentives to form triangles (which increase clustering), that is to connect to those agents whom they already benefit from. This is particularly true when δ is far from 0.5 (that is when $\delta - \delta^2$ decreases) holding constant the costs of link formation¹⁷. On the other hand, such positive externalities also provide agents with incentives to form triangles in order to benefit more from these agents whom they already benefit from just by reducing their relational distances to them by one. Efficient networks are thus expected to be more clustered than emerging ones since the social returns to overlapping connections (i.e. forming triangles) are higher than the

¹⁷For extreme values of δ (larger but close to 0.1 or close to 1), agents have very low incentive to form triangles since, for such values, δ becomes very close to δ^2 : the marginal benefits become very low as compared to the costs of forming such links. Thus, in these two extreme configurations, the emergent network should be the ring with a clustering equal to 0.

associated private returns because agents do not consider the positive externalities they generate on other players when forming such triangles.

What do we find? When $\delta \in [0.1, 0.2]$, agents in emergent networks are only connected to their two nearest geographic neighbors and clustering is null (see Figure 5). For a $\delta < 0.3$, clustering of both efficient and emergent networks rise up to 4 or 5 times their control random networks. We also find (see Table 1) that for $\delta \in [0.5, 0.7]$, efficient networks are 54% more clustered than their control random networks while the emergent networks are 34% more clustered than their own controls. When δ increases ($\delta \in [0.7, 0.85]$), efficient networks are still 70% more clustered while emergent ones are 13% more. When $\delta \in [0.85, 0.95]$, efficient networks are even more clustered while emergent ones are now 45% less. Clustering of emergent networks falls down when δ becomes close to unity because as evidenced above the private returns to such overlapping connections tend to zero. It should be noticed that for $\delta \approx 0.3$, emergent networks are too much clustered. This is because agents are not provided with sufficient incentives to form long distant connections for this low value of δ . (we have seen above that agents do not generate costly shortcuts in such a case).

Proposition 6 *For a large intermediary region of δ ($0.2 \lesssim \delta \lesssim 0.7$), efficient and emerging networks are both significantly more clustered than their control random networks. Globally, while controlling for their density, efficient networks are more clustered than emergent networks. Nevertheless, in a narrow region of δ ($0.3 \lesssim \delta \lesssim 0.35$), emergent networks do not generate distant connections which decrease the clustering ratio of efficient networks.*

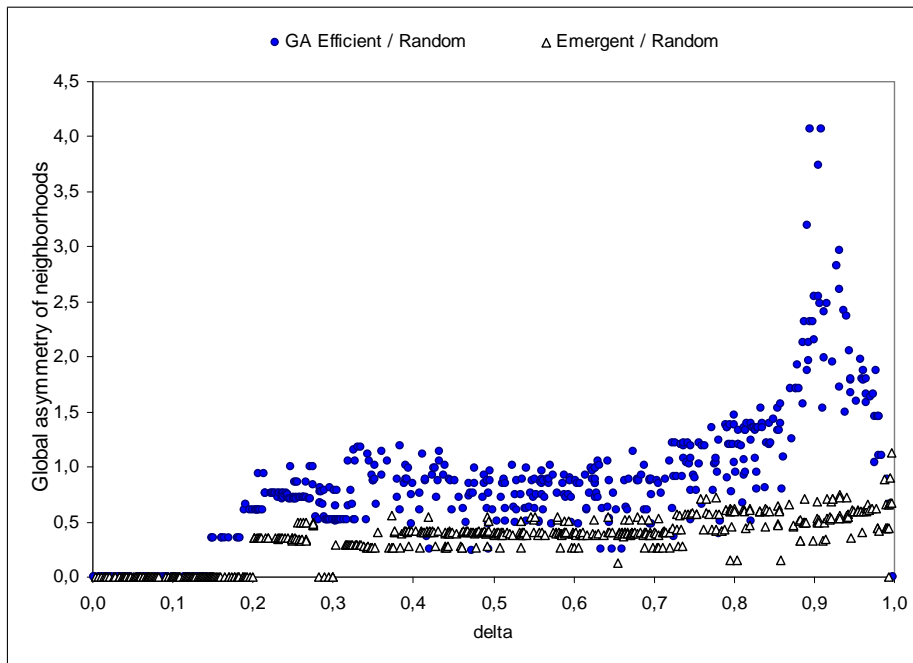


Figure 6: The global asymmetry of neighborhoods sizes (range $r(\cdot)$) of GA-efficient networks and emergent networks for the various values of δ relative to the average distance of their control random networks.

We now consider the distribution of connections over agents in the efficient and the emergent networks. The global asymmetry ratio of neighborhood sizes of efficient networks is 91% while

the global asymmetry ratio of emergent networks is only 37% (Figure 6 and Table 1). Therefore, efficient networks have a much more uneven distribution of connections than emergent networks. Computations of local asymmetry (see Figure 7) of efficient and emergent networks supports the idea that the emergent networks have a much too balanced distribution of connections over agents¹⁸. This statement applies for a very large region of δ : for δ higher than 0.2 and at least slightly lower than 1. The simultaneity of both geographically embedded connections and some more central agents in the network also explains why clustering is so high in efficient networks. When $\delta \approx 0.9$, the complete stars tend to become the efficient networks and the asymmetry ratios of efficient networks become very high. The absence of such stars in emergent networks stresses another coordination failure in network formation: the emergent networks are much less coordinated around central agents that contribute to increase the wealth in the population.

Proposition 7 *Emergent networks have a too much balanced distribution of connections over agents. Decentralized strategic interactions in network formation do not favor the emergence of central agents which would improve social wealth.*

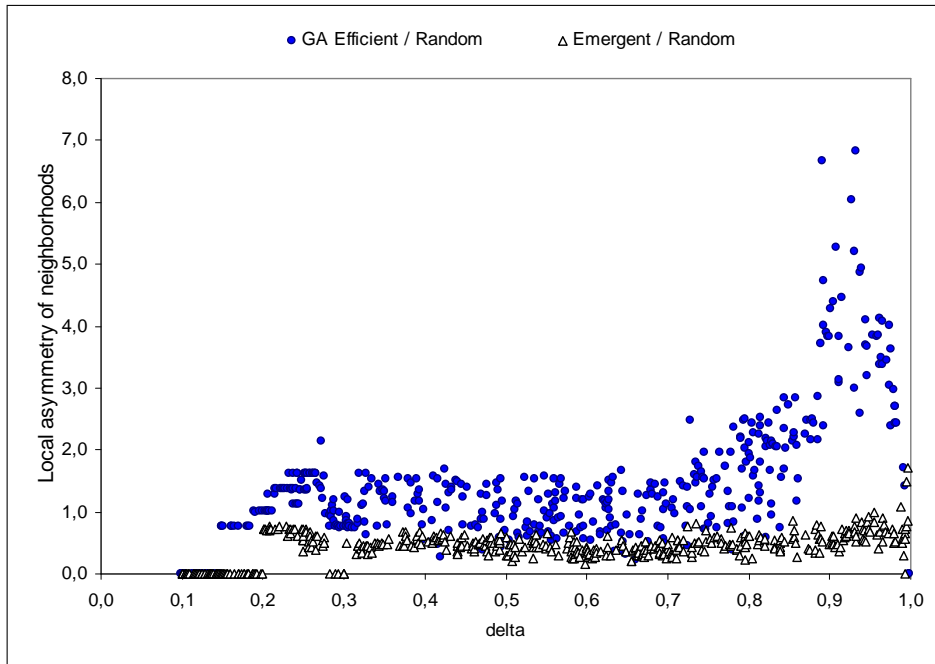


Figure 7: The local asymmetry of neighborhoods sizes $u(\cdot)$ of GA-efficient networks and emergent networks for the various values of δ relative to the average distance of their control random networks.

5 Conclusion

In this paper, we make a simultaneous use of two new approaches for computing emergent networks (Monte Carlo experiments, Carayol and Roux, 2003) and efficient networks (Genetic Algorithms,

¹⁸The global asymmetry is in average 37% of the one of their control random networks. The local asymmetry is 43% of the one of control random networks.

Carayol, Roux and Yıldızoğlu, 2003) in strategic network formation models. It allows us to systematically compare the structural properties of emergent and efficient networks in a spatialized version of the connections model which has proven to lead to emergent networks that resemble more to real social networks (Small Words, see Carayol and Roux, 2003). Our results highlight that welfare allocation within social networks introduces coordination failures that lead to inefficiency in networks formation.

Our first result is that emergent networks are significantly less dense than the efficient networks: agents generate fewer connections than they should as regard social surplus because the model exhibits indirect positive externalities to bond formation. Selfish agents do not naturally build those links for which the social returns overbalance the establishment costs, while the private returns do not. Secondly, efficient networks are more clustered than emergent networks: the social returns to triangular connections are again higher than the private returns. Thirdly, agents are in average (socially) too distant from each other and, emergent networks lack some costly distant connections, even if this problem is reduced when the decay parameter (that conditions positive networks externalities) takes intermediary values. Lastly, emergent networks are globally and locally not enough “coordinated”: emergent networks do not exhibit enough asymmetries between agents (the distribution of links among agents is too balanced). This should be contrasted with the efficient networks, where supplementary connections are preferentially attributed to a few agents who thus gain central positions in the network. Emergent networks do not share this structural property because no agent wants to bear the costs associated with such a central position, even if increasing the connectivity of some (initially identical) agents enhances the quality of many indirect social connections.

Agents would socially benefit from the compensation of some of them for internalizing networks externalities. This would increase the local density of the network and ensure the establishment of more costly (geographically) long distance connections that are particularly important in reducing relational distance between agents. Agents would also be better off if some agents (or only one of them, depending on the decay parameter) were selected and subsidized to play central roles. Global policies could aim at subsidizing these central agents, while local policies could aim at reducing the costs of some long distance connections. Decentralized bargaining that would lead to side payments among agents could also contribute to achieve such a goal if agents can subsidize connections between other agents to whom they are not directly connected (Bloch and Jackson, 2005).

References

- Albert, R. and Barabási, A.L. (1999). The diameter of the World Wide Web, *Nature*, 401, 130-131.
- Aumann, R. and Myerson, R. (1988) Endogenous Formation of Links Between Players and Coalitions: An Application of the Shapley Value in A. Roth (ed), *The Shapley Value*, Cambridge University Press, 175-191.
- Barabási, A.L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A. and Viscek, T. (2001) Evolution of the social network of scientific collaborations, *Physica A* 311, 590-614.
- Bloch, F., and Jackson, M.O. (2005) The formation of networks with transfers among players, mimeo Caltech and Greqam, january.
- Carayol, N. and Roux, P. (2003) Network formation in a model of distributed innovation, Mimeo BETA, Université Louis Pasteur and Université Toulouse 1.

- <http://cournot2.u-strasbg.fr/EEAAS/carayol/proxy_networks.pdf>
- Carayol, N., Roux, P. and Yildizoglu, M. (2005) Efficiency of Network Structures: The Needle in the Haystack, *Cahier du GRES* No. 2005-06, Université Montesquieu Bordeaux IV, Bordeaux, France.
- Davis, G. F. and Greve, H.R. (1997) Corporate elite networks and governance changes in the 1980s, *American Journal of Sociology*, 103, 1-37.
- Dawid, H. (1999) *Adaptive Learning by Genetic Algorithms*, Springer, Berlin.
- Dupouët, O. and Yildizoglu, M. (2005) Organizational performance in hierarchies and communities of practice, forthcoming in *Journal of Economic Behaviour and Organization*.
- Erdős, 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. and Wentzell, A. (1984) *Random Perturbations of Dynamical Systems*, Springer Verlag, New York.
- Goldberg, D. (1991) *Genetic Algorithms*, Addison Wesley, Reading:MA.
- Holland, J.H. (2001) *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge:MA, sixth printing.
- Jackson, M.O. (2004) A Survey of Models of Network Formation: Stability and Efficiency, Chapter 1 in Demange G. and M. Wooders (eds), *Group Formation in Economics: Networks, Clubs and Coalitions*, Cambridge University Press, Cambridge U.K.
- Jackson, M.O. and Rogers, B. (2005) The Economics of Small Worlds, forthcoming in the *Journal of the European Economic Association - Papers and proceedings*, <http://www.hss.caltech.edu/jacksonm/netsmall.pdf>.
- Jackson, M.O. and Wolinski, A. (1996) A strategic model of social and economic networks, *Journal of Economic Theory*, 71, 44-74.
- Jackson, M.O. and Watts, A. (2002) The evolution of social and economic networks, *Journal of Economic Theory*, 106, 265-295.
- Johnson, C. and Gilles, R.P. (2000) Spatial social networks, *Review of Economic Design*, 5, 273-299.
- Kirman, A.P. and Vriend, N.J. (2001) Evolving Market Structure: An ACE Model of Price Dispersion and Loyalty. *Journal of Economic Dynamics and Control*, 25(3-4), 459-502.
- Michalewicz, Z. (1996) *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag, Berlin.
- Tesfatsion, L. (1997) A trade network game with endogenous partner selection. In Amman, H., Rustem, B. and Whinston, A. B. (eds), *Computational approaches to economic problems*, Kluwer Academic Publishers, Dordrecht, 249-269.
- Tesfatsion, L. (2003) Agent-Based Computational Economics, *ISU Economics Working Paper*, No. 1.
- Vega-Redondo, F. (2005) Building up social capital in a changing world, Forthcoming in *Journal of Economic Dynamics and Control*.
- Watts, A. (2001) A dynamic model of network formation, *Games and Economic Behavior*, 34, 331-341.
- Watts, D.J. and Strogatz, S.H. (1998) Collective dynamics of 'small worlds' networks, *Nature*, 393, 440-442.
- Yıldızoğlu, M. (2002) Competing R&D Strategies in an Evolutionary Industry Model, *Computational Economics*, 19, 52-65.
- Young, H.P. (1993) The evolution of conventions, *Econometrica*, 61(1), 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 KØESSLER, 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 KØESSLER, 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 KØESSLER, 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/manager dans 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 ?*

- 2001–01 Sylvie DUCHASSAING, Laurent GAGNOL, décembre 2000.
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 SPÆETER, 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 régions de l'Union Européenne.*
Jalal EL OUARDIGHI, juin 2001.
- 2001–14 *Persuasion Games with Higher Order Uncertainty.*
Frédéric KÆSSLER, 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.*

- 2001–19 Raphaël NKAKLEU, octobre 2001.
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 KÆSSLER, janvier 2002.
- 2002–02 *Strategic Knowledge Sharing in Bayesian Games : Applications.*
Frédéric KÆSSLER, janvier 2002.
- 2002–03 *Partial Certifiability and Information Precision in a Cournot Game.*
Frédéric KÆSSLER, 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 SPÆETER, 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 KÖESSLER, 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 SPÆETER, 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 SPÆETER, Alban VERCHÈRE, juin 2002.
- 2002–17 *Parimutuel Betting under Asymmetric Information.*
Frédéric KÖESSLER, 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 KÖENIG, 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 CO₂ 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 OUARTIGHI, 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 ITERSHEIM, 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.

La présente liste ne comprend que les Documents de Travail publiés à partir du 1^{er} janvier 2000. La liste complète peut être donnée sur demande.

This list contains the Working Paper written after January 2000, 1st. The complet list is available upon request.
