MERIT-Infonomics Research Memorandum series

Evolving Networks of Inventors

Robin Cowan, Nicolas Jonard & Jean-Benoît Zimmermann

2004-018

MERIT – Maastricht Economic Research Institute on Innovation and Technology

PO Box 616
6200 MD Maastricht
The Netherlands
T: +31 43 3883875
F: +31 43 3884905

http://www.merit.unimaas.nl
e-mail:secr-merit@merit.unimaas.nl

International Institute of Infonomics

c/o Maastricht University
PO Box 616
6200 MD Maastricht
The Netherlands
T: +31 43 388 3875
F: +31 45 388 4905

http://www.infonomics.nl
e-mail: secr@infonomics.nl

Evolving Networks of Inventors*

R. Cowan†, N. Jonard‡, and J-B. Zimmermann♮

†MERIT, University of Maastricht, P.O. Box 616
6200 MD Maastricht, The Netherlands
‡CNRS, CREA Ecole Polytechnique, 1 Rue Descartes
75005 Paris, France
♮CNRS, GREQAM, EHESS, 2 Rue de la Charité
13002 Marseille, France

August 27, 2004

Abstract

In this paper a model for the formation of strategic alliances is studied. Innovation results from the recombination of knowledge held by the partners to the collaboration, and from the history of their collaboration. Innovation brings partners closer together, while at the same time the repetition of partnerships fosters trust and helps improving the outcome of each round of cooperation. A tension exists between innovating with people I know in order to reduce uncertainty at the expense of the net benefit from our joint effort, and innovating with strangers with whom the outcome of joint innovation can be greater but at a larger risk of failure. This "organized proximity", built through the experience of cooperation, can be at the origin of strongly structured networks of innovation, where agents' relations focus on limited cliques of partners.
JEL Classification: L14, Z13, O3
Keywords: Networks, Innovation, Network Formation, Knowledge

*Authors' E-mail addresses are `r.cowan@merit.unimaas.nl`, `nicolas.jonard@shs.polytechnique.fr`, and `jbenoit@ehess.univ-mrs.fr`. Jonard and Zimmermann acknowledge support from the French Ministry of Research, Project "Systèmes complexes en Sciences Humaines et Sociales ", S.Co. 0061.

1 Introduction

One of the long-standing legacies of Schumpeter (1934) is the belief that innovation was a central aspect of a dynamic economy. Schumpeter was also pivotal in his view that innovation consists largely of the recombination of existing knowledge. The practical force of this idea is that if agents have access to more and a wider variety of knowledge or information, this will foster innovation (see Weitzman, 1998; Olsson, 2000). Recently, changes in the technological landscape have made this issue more important than ever before. Several authors have noted that in many industries today the technologies both being used and being produced involve technological expertise that covers a much broader range of 'disciplines' than has hitherto been the case (see Powell et al., 1996, Grandstand and Sjolander, 1990; Grandstand, 1996, Teece and Pisano, 1989). What this implies is that types of knowledge necessary to innovate and compete successfully can lie outside a firm's main area of expertise. A now common way of coping with this problem is to form an alliance with a firm that has the missing expertise. Inter-firm cooperation can be extremely effective in increasing the circulation of tacit knowledge, and in creating possibilities for a firm to acquire knowledge outside its boundaries. Consequently, these cooperative agreements for R&D have grown dramatically in number since the 1970s.[1]

Nevertheless, cooperation is risky, in the sense that it is marked by uncertainty relative both to the actual skills of the partner and to his reliability. (See Powell 1990, p. 318 for a discussion of risks.) Cooperation implies mutual knowledge and sharing of routines, representations, ways of thinking and so on. In other words, it demands a proximity of actors that has to be built through the experience of cooperation (Garcia-Pont and Nohria, 2002). Coriat and Guennif (1998) argue that, in a partnership, trust evolves as a result of a "meaningful" repetition of interactions between the partners. They consider that interaction "provides additional information and meaning about the partners' behaviour", reducing uncertainty and establishing "a minimum of predictability in the behaviour of the partners". In that sense, trust construction and learning are strongly related (Sako, 1991; Dodgson, 1996; Klos and Nooteboom, 2001).

Cooperation also needs the existence of an intermediary common ground between partners, starting point of a common action. Galison (1999) shows that in experimental physics, cooperation between theorists, experimenters and instrument makers is made possible by the emergence of some sort of "creole", an intermediary level knowledge specific to a given pair of actors, that has been built through their repeated interactions.

This discussion suggests that in a variety of ways, after an interaction, assuming it was successful, the two parties will find each other more attractive than they did before it. A successful interaction will increase the partners trust in each others' motives, and increased awareness of each others' skills. It will also have generated more common tacit knowledge which improves communication. Powell et al. (1996) show that firms that have engaged in partnerships in the past are more likely to engage in them in the future. This is a general result. Roijakkers (2003) finds that if two particular firms have allied with each other in the past, these two firms are more likely to have an alliance together in the future.

[1] See Hagedoorn (2001) for a review and discussion of this trend.

But through cooperating, partners' knowledge profiles will tend to move closer together, reducing their mutual attractiveness. Nooteboom (2004) assumes the proposition "that ongoing interaction will yield a reduction of cognitive distance, or, in other words, identification." This is born out empirically: Mowery et al. (1998) find that "technological overlap between joint venture partners after alliance formation is greater than their pre-alliance overlap." (p. 517). (See also Dyer and Nobeoka, 2000.) This makes future cooperation easier. The cost, though, is that the increased similarity means that they can have less to contribute to each other. In the extreme, if they become identical in what they know, (or in their location in technology or knowledge space) there is no reason for them to collaborate, no matter how much they trust each other and how smoothly they can interact.

To model strategic technological alliances in their entirety is far beyond the scope of this paper. Technological alliances can be very rich and varied, in character, motivation and effect. But for our purposes we focus on a single aspect, namely the production of shared knowledge, and how firms' behaviour in this regard leads to the emergence of networks.

We have observed in the last decades a rapid emergence of a relatively new form of organization, in which firm structures are built around networks. Networks differ from markets and hierarchies in a variety of ways, but can be seen as depending on particular types of interactions between pairs agents within the economy. (For a discussion see Powell, 1990.) Sustaining a network connection demands many of the same things as establishing and sustaining a strategic alliance (trust, absence of opportunistic behaviour, familiarity and so on). Thus a natural way to approach network formation is as a result of individuals forming bilateral, non-market relationships.

There is now a growing literature in economics on network formation, but the majority of it treats the problem in game-theoretic terms, looking for stable structures that emerge from agents's one-time decisions about whether to form links. The concern in this literature tends to be whether the stable networks are efficient. (For a recent survey see Dutta and Jackson, 2003.) This work tends not to address the evolutionary nature of network formation and operation. The model developed in this paper continues work using another approach, centred on the idea that agents are continually forming and breaking links with each other, and that this ongoing activity is what underlies the networks that we observe. Agents choose with whom to form a link (or activate an existing link) in order to achieve some immediate goal. The repetition of these actions results in a network that emerges and evolves over time. The aim here is to identify the properties of the structures and behaviours that emerge (see for example Cowan et al. 2003; Kirman and Vriend, 2001; or Plourabouè et al. 1998).

In this paper we incorporate effects of collaboration in a simple model of network formation. The goal is to understand network formation as the consequence of individual firms creating bilateral alliances in which they innovate. Firms attempt to pool their knowledge to create new knowledge, and the amount they can create is determined by the complementarities in their knowledge stocks. In addition, repeated interaction between a pair of firms increases the probability of success, but at the same time can change partners' abilities to complement each others' knowledge. By repeated alliance formation and dissolution a

network structure emerges and continuously evolves. How this structure changes and how the nature of the economy's knowledge stock changes, in response to parameters governing different parts of the process, are the issues that we explore.

2 The model

The dynamics of the model are simple: each period firms form alliances based on their expectations of the amount of output attributable to any collaboration. An alliance having formed, the firms pool their knowledge to create a joint knowledge endowment. Innovation happens as recombination — the joint knowledge endowment serves as an input to knowledge creation. When new knowledge is created, it is added to the respective knowledge stocks of the two partners, and then the alliance is dissolved. This repeated pair formation and dissolution generates a process of network formation and evolution.[2]

The population of agents is finite and denoted $S = \{1, \ldots, n\}$. Each agent $i \in S$ is characterized by a knowledge endowment of several types of knowledge. This is represented as a vector $v(i) = (v^1(i), \ldots, v^\ell(i))$, where each element $v^m(i)$ represents the amount of knowledge of type $m = 1, \ldots, \ell$ held by agent $i \in S$. This representation permits us to treat agents as located in an ℓ-dimensional knowledge space. With this notation we can turn to knowledge creation and allocation.

2.1 Knowledge production

There are many ways to characterize knowledge, none of them without its pitfalls. But a representation of the innovation process should satisfy several minimal requirements. Consider two individuals i and j who conduct innovation jointly. After innovation has taken place, one would expect the following to be true: the knowledge amounts held by i and j have increased; the knowledge "profiles" of i and j have changed; and the similarity of the knowledge profiles of i and j (that is the relative distance between them in the knowledge space) has fallen.

Operationally, each pair of agents (i, j) creates an amount of new knowledge determined by a production function, and this amount is simply added to both partners' existing knowledge endowments. This process satisfies the three requirements just described.

Agents in a pair pool their knowledge to create a joint knowledge profile. This is done through

$$v^m(i,j) = (1-\theta)\min\{v^m(i), v^m(j)\} + \theta \max\{v^m(i), v^m(j)\},\ m = 1, \ldots, \ell. \tag{1}$$

Such a pooling of course remains virtual, as each of the partners remains the owner of his own skills. Nevertheless, it permits a useful formalization: after having aggregated

[2] We have assumed here that agents are pursuing knowledge for its own sake. This is unrealistic in general for firms, who pursue knowledge more generally for the sake of profits. To incorporate that explicitly in the model adds significant complication, demanding a fully blown goods market with production and consumption. We avoid that by this simplifying assumption, which, in an industry involved in rapid technical change, will be behaviourally quite adequate.

the individual endowments, the joint profile becomes the input of a knowledge production function.

In Equation (1), θ is governed by the nature of the knowledge pooling which the knowledge creation task demands. If it is possible to separate the sub-tasks in the innovation process, the agents will specialize, each agent doing some sub-tasks, and bringing the results together at the end to create the complete innovation. Here, since specialization is possible, the better econometrician will do the econometrics; the better theorist will do the theory, and so on, and thus the pooled knowledge vector is the element-wise maximum of the individual vectors: θ will approach one. By contrast, if the tasks are not separable, and both partners must be involved in every sub-task, then the weaker partner will be a bottleneck: the pooled knowledge vector will be the element-wise minimum, and θ will approach zero.[3]

On this interpretation θ can be considered a measure of the taste for dissimilar partners. Indeed if $\theta \approx 0$ then in any element m in which $v^m(j) < v^m(i)$, agent j reduces the effectiveness of i. The converse is true as well. In this case, agents will be driven to find partners similar to themselves, as in that case they provide the least drag on each other. In the extreme case they stay alone (this effect will be observed in the results below). By contrast, if $\theta \approx 1$, agents look for partners whose endowments tend to complement their own, since they can benefit from each others' strengths. Implicitly, they search for partners who are different from themselves in the sense of being good where they are bad.

The pooled knowledge vector serves as the vector of inputs to the innovation production function. To formalize this, we use a standard constant elasticity of substitution production function $\phi : \mathbb{R}^\ell \to \mathbb{R}_+$, with

$$\phi(v(i,j)) = \left[\sum_m (v^m(i,j))^\beta\right]^{1/\beta}. \tag{2}$$

Parameter β is an inverse measure of the elasticity of substitution across knowledge types, which is written as $1/(1-\beta)$. To see how it affects the type of partnerships wanted, consider an agent with asymmetric profile (i.e. marked strengths and weaknesses) and remark that ϕ is symmetric and homogeneous. In general this agent would like to find a partnership such that the joint profile is more evenly distributed and so a higher isoquant can be reached. How desirable this is varies with the degree of substitutability across knowledge types. When β is small the agent is really eager to find a partnership and many will suit him. By contrast, when β is close to 1, being evenly skilful becomes less important as substitution between knowledge types gets easier. So less networking should be observed.

[3] In the literature there is some disagreement about whether innovative success increases or decreases with the cognitive distance of the partners. Mowery et al. (1998) find a U-shaped relationship; Nooteboom (1999) argues for a similar conclusion. At the same time though, van Alstyne and Brynjolfsson (1996,1997) and Peretto and Smulders (2002) and others assume that the relationship is monotonic, either increasing or decreasing. With our formalism, this effect is parametrized, and explicitly linked to the nature of the innovation task, and includes the possibility of a relatively complex relationship due to the multi-dimensional nature of firms' knowledge.

2.2 Innovative success and experience

Finally, for knowledge creation, there is the issue of success. An innovation project may fail, and the projected new knowledge may not be created. Here history is important since project success is driven in part by familiarity of the partners, as discussed in the introduction.

Define the probability that the next collaborative attempt of i and j is a success, given they have already had t attempts, as $\pi_t(i,j) = \pi_t(j,i)$. This probability is determined by their joint history, and captures the development not only of trust but also the ability to work successfully together. We define $\gamma_t(i,j)$ as a measure of their historical success. Recent experiences will weight more heavily, so define

$$\gamma_t(i,j) = \sum_{1 \leq s \leq t} \rho^{t-s} \chi_s(i,j), \tag{3}$$

where $0 < \rho < 1$ is a discount factor, and $\chi_s(i,j) = 1$ if i and j had a successful interaction for their sth collaborative attempt and 0 otherwise. Given this definition of past success, i and j form their beliefs about successful collaboration as an a priori probability $\pi_t(i,j)$ that the collaborative attempt $t+1$ is a success. We define the success probability of the first encounter between i and j, the lower bound of the success probability as $\pi_0(i,j) = \pi_L$, for all $i \neq j$, and assume that $\pi_t(i,j)$ is an increasing function of $\gamma_t(i,j)$. Innovation is always a risky activity, and so the maximum possible success probability (if partners have an infinitely long perfect success record) is $\pi_H < 1$. As there is no a priori obvious choice for the functional form of the relationship between $\pi_t(i,j)$ and $\gamma_t(i,j)$, and because we would like the impact of $\gamma_t(i,j)$ to be gradual, a simple linear form is posited.[4] As

$$\gamma_t(i,j) \leq \sum_{1 \leq s \leq t} \rho^{t-s} = \frac{1-\rho^t}{1-\rho}, \tag{4}$$

an obvious upper bound to $\gamma_t(i,j)$ is $1/(1-\rho)$. The functional form we assume for $\pi_t(i,j)$ is then

$$\pi_t(i,j) = \pi_L + \gamma_t(i,j)(\pi_H - \pi_L)(1-\rho), \ \forall t \geq 1. \tag{5}$$

Then the expected amount of knowledge potentially produced by a cooperation between i and j can be expressed as

$$F(i,j) = \pi_t(i,j) \cdot \phi(v(i,j)). \tag{6}$$

This is simply the amount of knowledge produced in case of success multiplied by the probability that the cooperation succeeds. (Thus agents are assumed to be risk-neutral.) Operating in autarchy is assumed to have a success probability of π_H (there is no issue of learning to collaborate with myself). As a result the magnitude of $\pi_H - \pi_L$ plays a central

[4] An exponential specification was also tried, but the effect of changes in $\gamma_t(i,j)$ was too marked, though similar behaviour could be seen to emerge. The linear form, on the other hand, produces smoother changes in the aggregate statistics.

role in a firm's decision to operate in autarchy rather than collaborating. Given π_H, a larger baseline probability of failure in new born relationships (lower π_L) will obviously imply a stronger tendency for firms to conduct innovative projects in isolation. In the case where learning is absent any collaboration has a probability of success equal to that of autarchy equal to π_H in each period (past interaction is irrelevant).

Rating potential partners according to Equation 6 is formally similar to the score approach used by Klos and Nooteboom (2001). However interpretations differ. In their model the score is an increasing function of the profit the agent can potentially make as a result of the transaction and his trust in the potential partner, where trust refers to "the intention to honour agreements to the best of one's ability" (p. 514). Thus there is a subjective probability that the other party is not opportunistic which has an effect on preferences, while at the same time a transaction actually never fails (so the objective probability that the other party is opportunistic is actually zero). We abstract from the difficulties of dealing with trust and simply consider $\pi_t(i,j)$ to be a measure of agents i and j's ability to work together, which increases with the past success of the relation.

If the innovation project is successful, and new knowledge is created, it is added to each of the partners' knowledge vectors. The general intuition is that as an agent uses knowledge or is exposed to it, he will assimilate at least part of it, and thereby change the precise area of his expertise. As the argument of the production function is the joint knowledge profile, it is natural to let this joint profile also determine the type of knowledge produced.[5] It is assumed that the probability, conditional to the collaboration being successful, of the new knowledge being of type m is

$$\frac{v^m(i,j)}{\sum_m v^m(i,j)}. \tag{7}$$

If the collaboration fails, both agents get 0.

The way the success probability of a collaborative attempt is up-dated is reminiscent of the standard reinforcement learning model imported from psychology (cf. Hebb, 1949; or Luce, 1959) and popularized among others by Erev and Roth (1995). In reinforcement learning an agent chooses probabilistically within a finite set of strategies, based on a cumulative (though slowly decaying) memory of the benefits collected by strategies at each past chance. The propensity to choose a strategy is updated based on the payoff earned when that particular strategy is used, and choice among strategies is made based on their propensities. Thus strategies that have done well in the past are chosen more frequently in the future. In the present paper agents learn to collaborate on the basis of their historical success, as in reinforcement learning. The logic is different, however, as agents do not choose probabilistically among different partners nor do they actually learn from history the payoff associated with particular partners. Rather they are able to calculate accurately the success-payoff, based on current information. What is updated, depending on past

[5] We have explored other variants in which a share of the new knowledge is allocated according to the joint profile to a category common to both participants, while the remaining part is allocated according to each partner's profile. As long as the share of new knowledge allocated according to the joint profile is not negligible, the results remain qualitatively unchanged.

successes and failures, is a belief about how likely a collaboration is to be successful, and thus partners with whom an agent has had success in the past are likely to be viewed favourably in the future.

2.3 Pair formation and equilibrium

We draw on the literature on matching problems for our basic model of pair formation. Because we consider a single population of firms rather than two populations (of jobs and workers for instance) matching here is a roommate problem, rather than a standard marriage problem. A one-sided, roommate, matching problem is defined as follows (see Gale and Shapley, 1962). Each individual $i \in S$ has a strict preference ordering $\succ_i$ over all the individuals in $S - \{i\}$. All preferences are complete and transitive. Let $\succ = \{\succ_i, i \in S\}$ denote the profile of the preference orderings of all the individuals in S. We generalize the standard roommate problem to include the possibility of self-matching. This is done in a straightforward way: the preference ordering is over the entire set S.

Definition 1 *The pair $(S, \succ)$ is called a roommates matching problem.*

A generalized matching is a partition of S into q singletons and $(n - q)/2$ pairs of roommates, that is to say a bijection $\mu : S \to S$ such that $\mu(\mu(i)) = i$ for all $i \in S$.

Definition 2 *A matching μ is stable in $(S, \succ)$ if there is no $(i, j) \notin \mu$ such that both $j \succ_i \mu(i)$ and $i \succ_j \mu(j)$.*

Put another way, stability is characterized by the non-existence of blocking pairs. In the particular problem examined here, the preference profile $\succ$ is generated by the expected output of a pairing $F : S^2 \to \mathbb{R}_+$, which associates to any pair of individuals (i, j) a value that represents the expected innovative output of this pair. In the event that $i = j$ the pooled vector is simply the vector of i, and production remains defined as it was above. The profile of preference orderings $\succ$ is then defined by the following rule.

Definition 3 *For any $i \in S$, $j \succ_i k$ if and only if $F(i, j) > F(i, k)$, for all (j, k) in $S^2, k \neq j$.*

Before turning to the emergence of network structure and the associated knowledge dynamics, we discuss further in detail the market clearing mechanism present in this model. Because agents in any pair assign the same cardinal value to their match, a unique stable matching always exists, i.e. the market for alliances always possesses a unique equilibrium. We prove this by construction.

Proposition 4 *A roommates matching problem $(S, \succ)$ for which the preference ordering of any $i \in S$ is derived from $F : S^2 \to \mathbb{R}_+$, with $F(i, j) = F(j, i)$, has a unique stable matching μ.*

Proof: The algorithm to construct the stable matching is as follows. Let $S_0 = S$ and $\mu_0 = \{\varnothing\}$. There exists a pair (a_1, b_1) such that $F(a_1 b_1) = \max_{(i,j)\in S^2} F(i,j)$. Furthermore as F is an injection, (a_1, b_1) is unique. Then (a_1, b_1) must belong to any stable assignment, as b_1 is preferred by a_1 to any other partner, and this preference is reciprocal. No matching which does not involve this pair could be stable. Hence the pair (a_1, b_1) is necessarily part of a stable matching. Note that it is possible that $a_1 = b_1$. Let then $\mu_1 = \mu_0 + \{(a_1, b_1)\}$ and $S_1 = S_0 - \{a_1, b_1\}$. If $a_1 = b_1$ then $\mu_1 = \mu_0 + \{(a_1, a_1)\}$ and $S_1 = S_0 - \{a_1\}$. Again there is a single pair within S_1 maximizing the innovation function. Proceed recursively this way. Denote p the smallest integer such that $S_p = \{\varnothing\}$. Then $\mu_p \equiv \mu$ is a stable matching and $q = 2p - n$ is the number of agents that have preferred self-matching to cooperation.

In case of a tie (that is to say when individual i can achieve the same innovative output with two or more different partners), we apply an arbitrary rule to guarantee that the score function is still strict: if $F(i,j) = F(i,k)$ then $j \succ_i k$ if and only if $j > k$.

The model just developed represents a complex dynamic process. As such it is impossible to track analytically, so we use numerical experiments to examine how the performance of the system responds to different parameters. We are concerned with three parameters here: the nature of the innovation task (whether divisible or not) and consequently the nature of knowledge pooling as measured by θ; the elasticity of substitution in the production of new knowledge (β); and the effects of agents' history of interactions. These parameters are varied in the numerical experiments below, in order to understand their effects on network structure and knowledge growth and distribution.

3 Numerical experiment

We study a population of $n = 100$ firms. At the outset individual knowledge endowments are randomly drawn from a uniform distribution over [0,1], independently for every element $v^m(i)$ in each agent's knowledge vector. Each period, the market for collaborative agreements is activated and firms form pairs (or stay alone) in order to innovate. The pairing results in a stable matching where stability is defined as above (everybody is as satisfied as possible, given everyone else's preferences), and where the value of a pair is equal to the expected amount of knowledge produced by that pair. After innovation, the new knowledge is added to the firms' knowledge stocks; the firms' knowledge types change, as described previously; and so does accumulated experience. At the end of the period all pairs disband, and the process begins again in the following period. We iterate this process for 1,000 periods, recording data for the entire history of the industry. In the numerical experiments reported below we chose the initial probability of a success in an emerging relation to be $\pi_L = 0.9$, and learning gradually increases it to values close to $\pi_H = 0.99$. The three parameters we examine are θ, β and the presence versus absence of cumulative learning about other agents. Regarding pooling (θ) we consider 100 randomly generated values uniformly distributed over $[0, 1]$. Similarly β takes 100 uniformly distributed values

in $(0,1]$. We thus have a large data sample on which we apply a simple non-parametric estimation technique — Kernel regression (Yatchew, 1998 for a comprehensive presentation) — which basically amounts to local averaging.[6]

Regarding the properties of knowledge accumulation, beside the obvious effect that increasing β decreases knowledge levels (see Equation 2), we care about the allocation of knowledge. We employ the knowledge production function ϕ rather than say the sum over knowledge categories so endowments are projected into a unit in which summing makes more sense. Letting $\phi = \sum_i \phi(v(i))$, equity in allocation can be assessed by considering the coefficient of variation in innovative potential

$$v = \frac{\sqrt{\sum_i \phi(v(i))^2 / n - (\phi/n)^2}}{\phi/n}. \tag{8}$$

Large values of v indicate the coexistence of rich and poor knowledge agents, while lower values of v indicate a more even distribution. At the individual level symmetrically the specialization index s_i of individual i can be defined via the coefficient of variation in his endowments $s(i) = \sigma(i)/\bar{v}(i)$, where $\bar{v}(i)$ is the average knowledge level of i and $\sigma(i)$ the standard deviation. The larger this index is the more of a specialist and less of a generalist i is. Summing across the population produces a normalized specialization index $s = \sum_i s(i)$, for which large values indicate a population of experts, while low values indicate a population of generalists.

Regarding the network, in any period the static network consists of q isolated agents and $(n-q)/2$ disconnected pairs, as given by the stable matching μ_t. To study the properties of the dynamic network, we record the list of connections active over time. This generates a weighted graph, in which the weight of an edge indicates how frequently the two firms have interacted in the history.

Denote (S, V_t) the graph associated with the stable matching achieved at time t, with $V_t(i,j) = 1$ if $(i,j) \in \mu_t$, and $V_t(i,j) = 0$ otherwise. The weighted graph recording past interactions is denoted (S, W_t), where $W_t(i,j)$ is the frequency of activations of the connection between i and j, obtained as $W_t(i,j) = \sum_{1 \leq s \leq t} V_s(i,j)/t$. For this graph we study the properties of the distribution of collaborative links, specifically the average path length and the clustering coefficient (cliquishness in Watts and Strogatz, 1998). To move from W_t to a 0/1 graph, distances must be computed first. The distance $d(i,j)$ between two nodes i and j is the number of edges in the highest frequency path linking them. Indeed any path $i_0, i_1, \ldots, i_z$ with $i_0 = i$ and $i_z = j$ has an associated frequency $\prod_{l=1,\ldots,z} W_t(i_{l-1}, i_l)$ and a length $z \geq 1$, and a path with maximum frequency exists. The average path length is then

$$d = \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} d(i,j)$$

[6]The quality of Kernel estimation is controlled by a scale parameter: the bandwidth. The figures which are presented below are kernel estimations of a number of statistics, obtained with a uniform kernel and an optimal bandwidth obtained by cross-validation, which consists in minimizing the out-of-sample prediction error.

and simply measures how distant vertices are on average, which is a global property of the graph (its inverse is sometimes referred to as closeness centrality in social networks analysis). Denoting $\Gamma(i) = \{j \neq i : d(i,j) = 1\}$ the neighbourhood of vertex i and $n_i = \#\Gamma(i)$ the size of i's neighbourhood, the average degree of the graph is

$$\Delta = \frac{1}{n} \sum_i n_i,$$

and measures the density of the interaction structure (sometimes referred to as degree centrality in social networks analysis). Cliquishness is the share of active links between any given vertex's neighbours, averaged over the system. It is written

$$c = \frac{1}{n} \sum_i \sum_{j,l \in \Gamma(i)} \frac{X(j,l)}{n_i(n_i - 1)},$$

where $X(j,l) = 1$ if $d(j,l) = 1$ and 0 otherwise.

4 Results

We are interested in the effects of the nature of knowledge pooling on network structure and on the growth of knowledge. The network is the result of firms' abilities to combine knowledge of different sorts and evolves as firms' needs, and preferences over the population of potential partners changes over time.

In each of the figures that follow there are two panels. In the left panel, we assume that past experience with an agent has no effect on the probability of success of a collaboration with that agent. Initial or a priori success probabilities are set at $\pi_H = 0.99$. In the right hand panel, the effects of experience are a described in Section 2.2 above, in which the success probability increases with past successes starting from $\pi_L = 0.9$ and bounded above by $\pi_H = 0.99$. In each panel the results are displayed as a shaded contour plot, which should be read as a map in an atlas: darker grey scales imply higher values on the z axis. This provides a compact display of the relationship between the pooling parameter (θ), the substitutability of knowledge types in production (β), and the performance measures we are concerned with.

4.1 Network

Figure 1 displays the number of connections held by the average agent, that is, how many distinct partners an agent has, on average, over the history of the economy. The first thing to observe is that magnitudes can differ by a factor up to 5 depending on whether experience matters or not. This is because experience and the size of innovation are substitutes in the score value (Equation 6) based on which individual rankings are formed. Thus learning about a firm tends to make it more attractive as a partner over time, as a result of which there is always less networking when past experience is relevant in predicting the outcome of future collaboration. Here we observe the inertia in network formation as described

by Roijakkers (2003): when experience matters, firms tend to stick with partners they have had in the past. Also interesting is the fact that the frontier between autarchy and collaboration is not located identically in both panels. The area of the parameter space in which only autarchic innovation takes place is larger when experience with other firms matters. That is, for any value of β, when there is learning, partnerships (as opposed to autarchy) will appear only for larger values of θ. This is under autarchic innovation the probability of failure is always $1 - \pi_H$, whereas failure in joint innovation can occur with probabilities as high as $1 - \pi_L > 1 - \pi_H$. An agent will only run this higher risk of failure if he is compensated, which can only happen if having a partner adds significant value. Only when θ is high can finding a partner (different from oneself) add enough value. Thus when experience matters, isolation is the preferred mode of functioning over a larger range of parameter values.

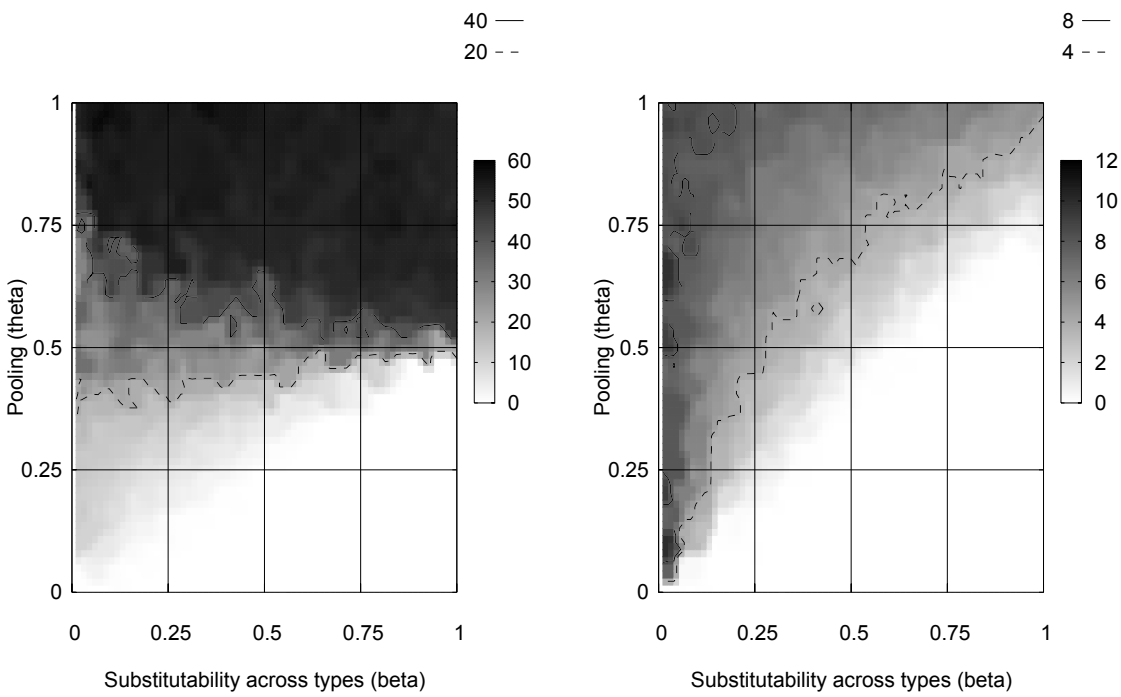

Figure 1: Average degree of the network in the (β, θ)-space (left: no learning; right: learning).

Figure 2 is a binary indicator of the connectedness of the network (0, white: disconnected; 1, black: connected). Intermediate grey levels are artefacts of the Kernel smoothing procedure. The emergence of a unique connected component takes place when θ is large enough for given β. Including experience allows the possibility of enclaves forming. These will form for random reasons, and show no pattern in (β, θ) space. But the formation of an enclave implies that that group will not be connected with the rest of the network, which shows in Figure 2 as a white island within the black region of connected networks. The

effect of β arises because as β increases, the ability to substitute between knowledge types increases, and so the need to find a partner to "fill a hole" in a firm's knowledge vector is less severe. Thus when β is large, many firms are able to innovate successful on their own, and there is no strong compulsion to find a partner. On the other hand, when β is small, it is difficult to substitute one type of knowledge for another, so if a firm has a "hole" in its knowledge profile, it is imperative to find a partner, and this implies that the graph will connect. The effect of θ works through the nature of optimal partnerships and the dynamics of knowledge. If dissimilar partners are the norm (θ is small) repeated partnering with the same partner drives firms too close together and they soon switch to another. Firms have many partners, which tends to create a connected network. By contrast, when similar partners are desirable, (large θ), sticking with the same partner is a good thing, so firms have few partners and the network will contain many isolated groups. The β and θ effects play off against each other to create the boundary between connected and disconnected networks.

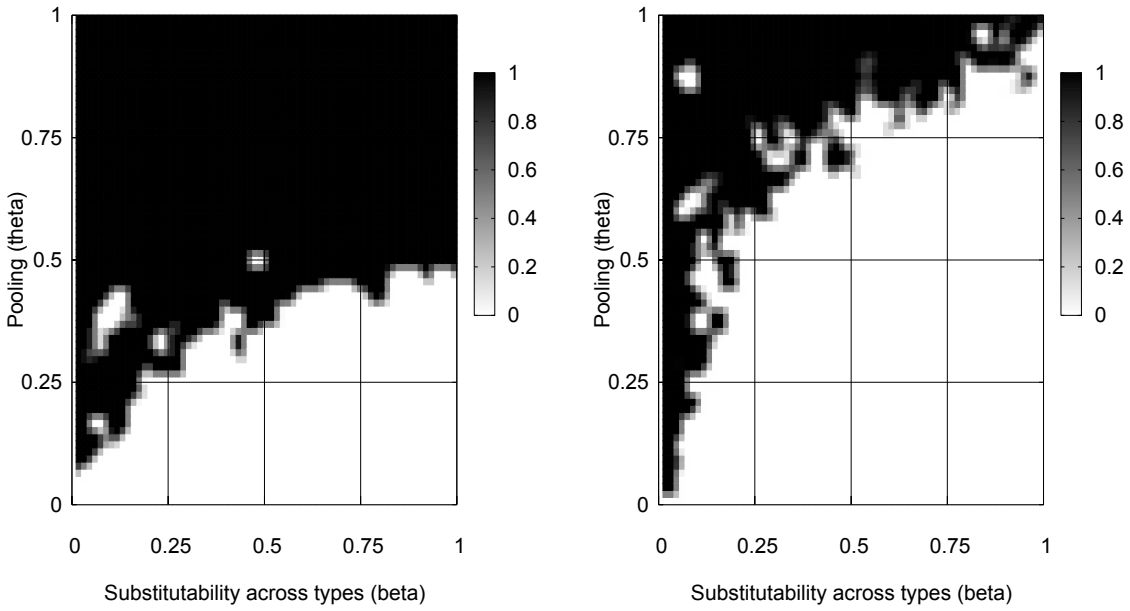

Figure 2: The emergence of connectedness in the (β, θ)-space (left: no learning; right: learning).

Looking at Figures 1 and 2 together we can see a zone of relatively low degree between the parts of the parameter space where the network is disconnected (white in Figure 2) and where it is connected (black). This suggests that within this zone, a thick path roughly along the main diagonal in (β, θ)-space, the network consists of connected sub-networks; that is, as the network evolves, small, coherent but isolated groups form.

Figure 3 depicts average shortest path lengths (distances) between agents in the network. Because a disconnected graph has some agents who are infinitely distant from each other, the averages shown in Figure 3 are computed only for the networks that are connected. The pattern here mimics Figure 2 with relatively little variation in distance for the connected networks. For the sake of comparison, a random graph of homogeneous degree Δ has path length approximately equal to $\ln n / \ln \Delta$. Without learning, in the networks that emerge from our model, the average path length within a connected graph is very irregular between 2 and 4 in a triangle-shaped zone in the middle of the panel, and then between 1 and 2 in the higher larger grey zone. In the middle triangle the degree is between 20 and 40, and between 40 and 60 in the upper one. So the prediction for a random graph would be respectively around 1.4 and 1.2. Thus the paths are slightly longer than in the equivalent random graph. When experience matters, average distances are between 2 and 4 along the left axis, and between 4 and 5 in the upper zone. As the degree is around 8 the prediction for a random network would be around 2.2. So again we have longer distances in our results, suggesting the structure is not purely random.

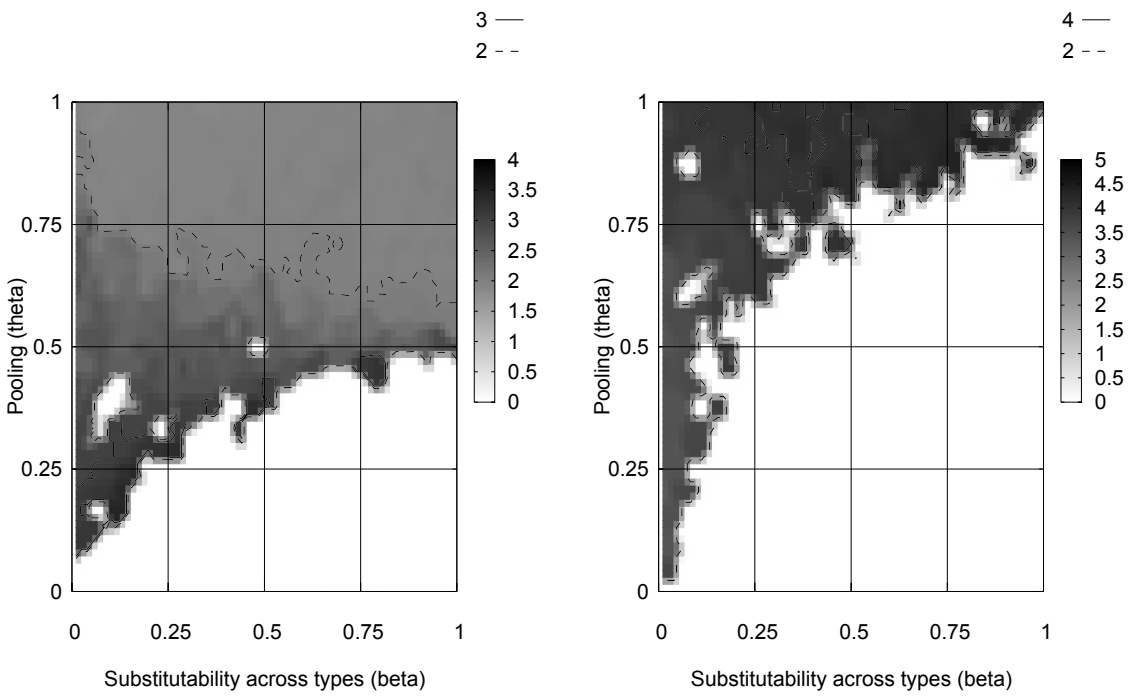

Figure 3: Average distance in the (β, θ)-space (left: no learning; right: learning).

Figure 4 displays the average share of an agent's neighbours who are also neighbours of each others (transitive triples). Thus this coefficient measures the degree of local transitivity emerging in the graph. Again there is a factor of almost 3 difference in the magnitudes between the case when experience matters and when it does not (in the graphs the z-range was compressed to render differences more visible), and the relationship to the two

parameters (θ and β) closely follows that for degree. Since path length does not change dramatically as links are added to the network (Figure 3, this implies that as firms find new partners they are not creating shortcuts but are rather reinforcing local coherence. The white region at the bottom of the graph corresponds to the region in parameter space in which firms tend to innovate as individuals rather than as part of a pair. There clustering is defined to be zero.

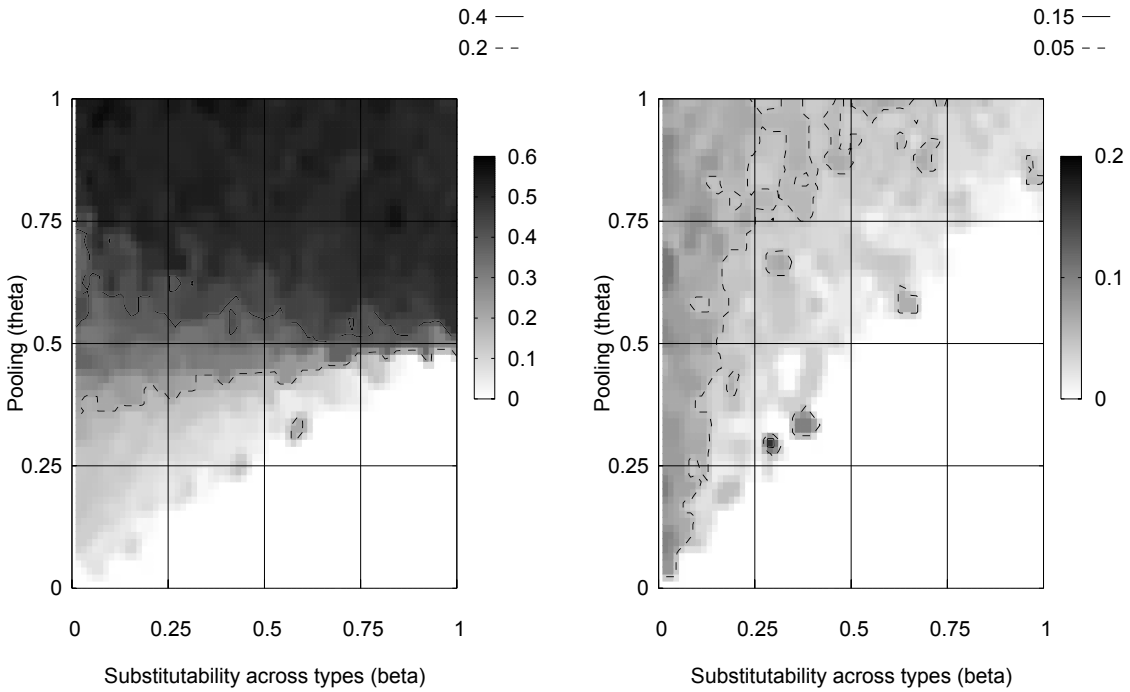

Figure 4: Cliquishness in the (β, θ)-space (left: no learning; right: learning).

This measure of cliquishness can be misleading, though, since it is driven to a very large extent by the degree of the graph. As agents acquire more links, even if they are acquired at random, the network will become more dense locally. To get a better measure of the structure of the graph, it is necessary to compare cliquishness with that of a network of known structural properties. We do this with a measure called "excess cliquishness". Excess cliquishness, as shown in Figure 5, re-scales the measure from Figure 4 to make it comparable to a random graph with the same average degree (which has cliquishness of approximately Δ/n.) Figure 5 depicts the ratio of observed over predicted cliquishness, and only values significantly larger than 1 would indicate a structure richer than a random graph.

In the left panel there is a wide middle triangle between 1 and 2, and the upper one is slightly below 1, together with peaks along the frontier where levels higher than 2 are reached. In the right panel we see relatively isolated islands of excess cliquishness reaching levels higher than 2 along the autarchy frontier. Other isolated islands corresponding to

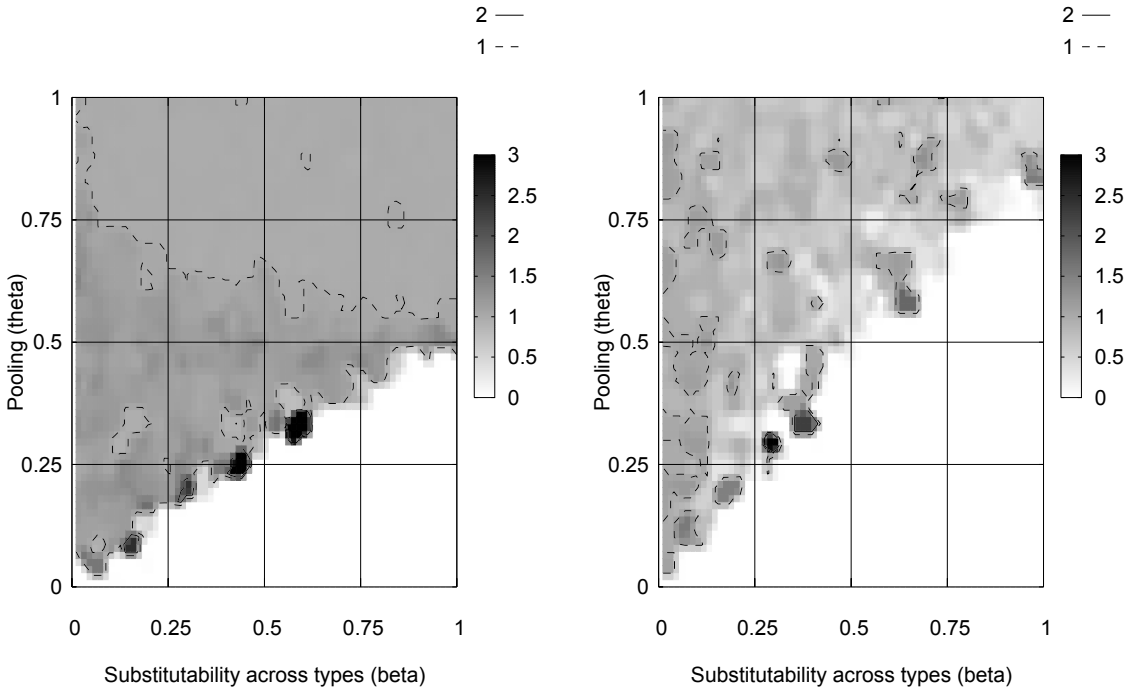

Figure 5: Normalized cliquishness in the (β, θ)-space (left: no learning; right: learning).

levels between 1 and 2 exist, mostly along the θ axis (so for low β), but the pattern is much less marked than in the left panel. So in general, the networks tend to be more clustered than random networks. The discernible effect of β, the substitutability across knowledge types, would be that highest excess cliquishness obtains for intermediate values of β. There is an effect of θ, the knowledge pooling parameter. When θ is just above the threshold at which networking begins, the optimal matching joins agents who are similar to each other. In this region a great degree of similarity between partners is tolerated. Thus a pair of agents tends to stay together for a long time, and converge toward each other in knowledge space. But partners can be too close to each other. When this occurs they search for new, but still relatively similar partners. This area of the parameter space will engender "partner-swapping". Although the magnitude of the excess clustering is relatively small except in very localized regions, there is a suggestion that cliques tend to form, which together with the existence of longer path lengths (as compared to a random situation) is evidence partly supporting the small world assumption. That it does not show up it a too marked on the other hand also should probably not come as a surprise, because in the model there is nothing that would give transitivity a value.[7]

[7]To illustrate where this value could arise, in Cowan and Jonard (2004) agents are bartering knowledge over a fixed network. There the value of transitivity is that if there is a failed double coincidence of wants between two directly connected agents and still one of them could benefit from the other, then the existence of a common acquaintance provides a route of length 2 over which knowledge can also travel if

4.2 Knowledge

Figure 6 depicts the relationship between the long run coefficient of variation of individual knowledge creation and the pair (β, θ). The general patterns are the same when experience matters and when it does not. Inequality is weakly decreasing along the β axis with a sharp peak when $\beta \approx 0$, but there is no discernible effect of θ. The effect of β is explained through the nature of the CES production function. As β increases from 0 to 1, the marginal product of the knowledge inputs falls quite dramatically, particularly near zero. This implies that when β is small, well-endowed agents will make large innovations. They will therefore tend to grow faster than less well-endowed agents. This preserves, and possibly magnifies initial differences. This is the effect of β seen in the figure.

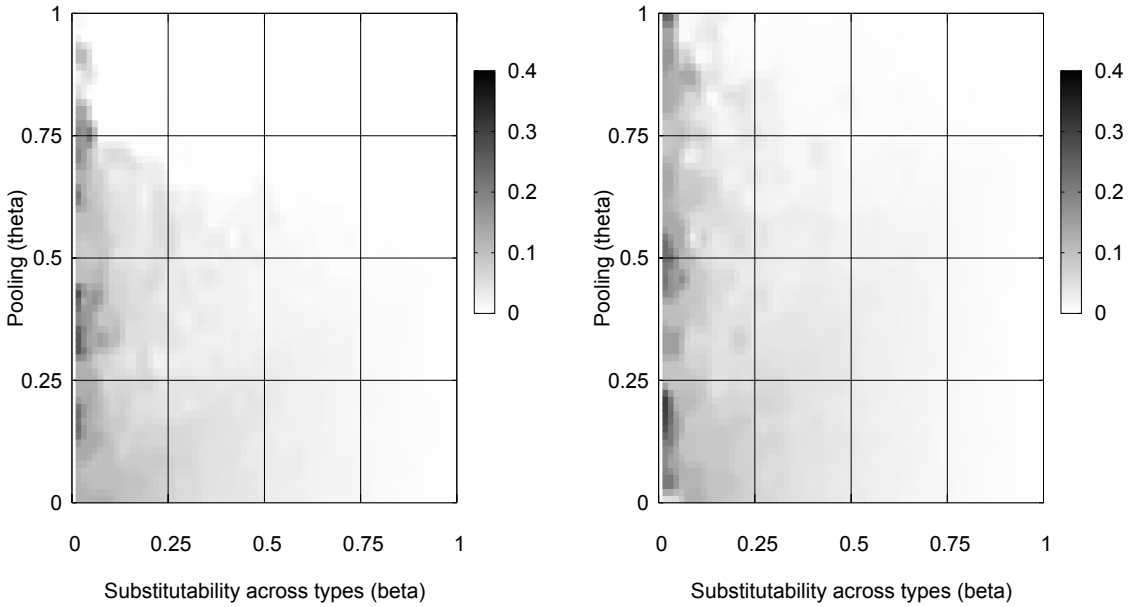

Figure 6: Coefficient of variation in the (β, θ)-space (left: no learning; right: learning).

In Figure 7 we display the average individual specialization as measured by the individual coefficient of variation on knowledge endowments.

The individual coefficient of variation has an analytical lower bound of 0 when all components of the endowment are identical (the agent is a generalist), and an upper bound of 2 when all the knowledge held by an agent is in one category (the agent is an expert). The effect of θ is clear, with the degree of specialization falling as pooling moves towards the maximum of the partners' endowments. This figure shows essentially a mirror image

the conditions are met.

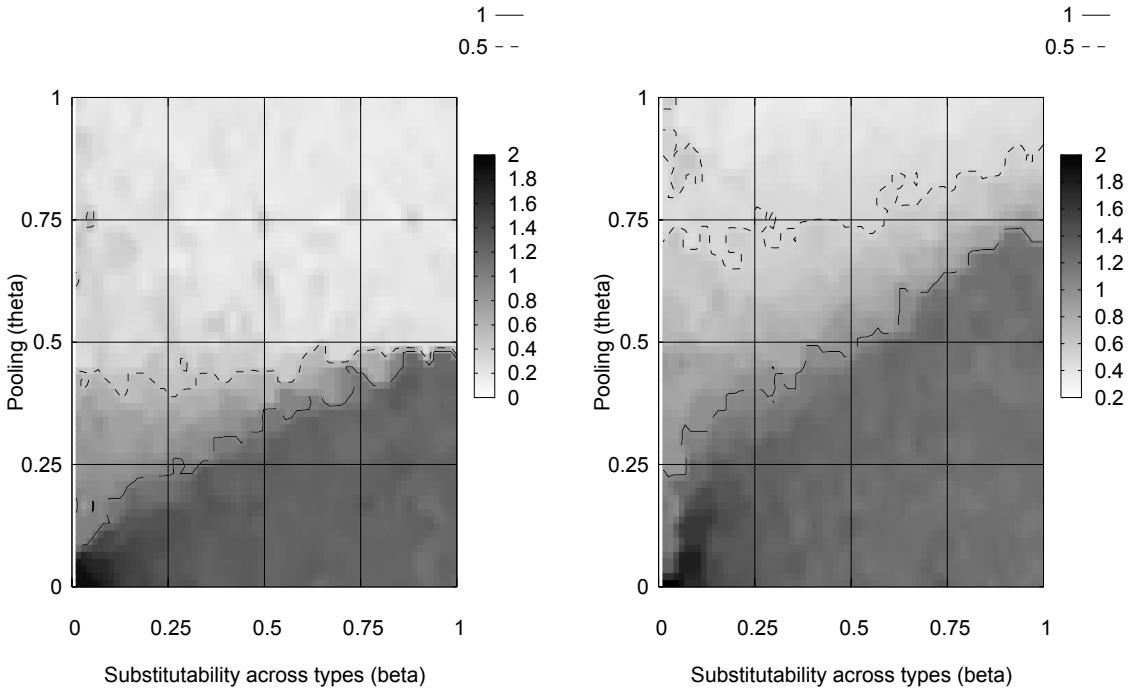

Figure 7: Specialization in the (β, θ)-space (left: no learning; right: learning).

of the pattern in Figure 1, the density of the network. Agents become highly specialized when they innovate largely as individuals. What drives this is that the type of knowledge produced is probabilistically the same as the knowledge input. Thus an agent is likely to innovate where he has most knowledge. In expected value, this process will lead to an agent innovating always in the same knowledge type, and so drive extreme specialization. When alliances form in a more systematic way (larger θ) there is quickly the emergence of generalist profiles. Sometimes an agent will innovate in his speciality, sometimes in his partner's, and an agent will have many different partners. This will both smooth the agent's profile immediately, and possibly even shift his area of expertise. This sort of mixing produces much flatter profiles, and the more partners an agent has, the more this mixing will take place.

Figure 8 is the average over the agents of the average pairwise angle between an agent and his partners. The weights in the averaging are the interaction frequencies. The larger this angle is the more different (or complementary) the expertise involved in the pairs that form. Again we have a pattern that resembles that exhibited by network degree in Figure 1. When I have a small number of partners over my history we repeatedly have "angle diminishing" interactions, since a joint innovation adds the same amount to the two of us, so decreases unequivocally the angle between us. Thus the average angle between us is lower than when I have a large number of partners, given I have a constant opportunities for partnerships. This is an unintentional effect of the number of partners I have over

the course of the history. In addition there is the intentional effect that for high θ ceteris paribus I look for people different from me (complementary, so large angle), when for low θ only pairs between similar people will form, also contributing to a low angle over history. So in this case the intentional and unintentional effects act in the same direction to produce the patterns seen in Figure 8.

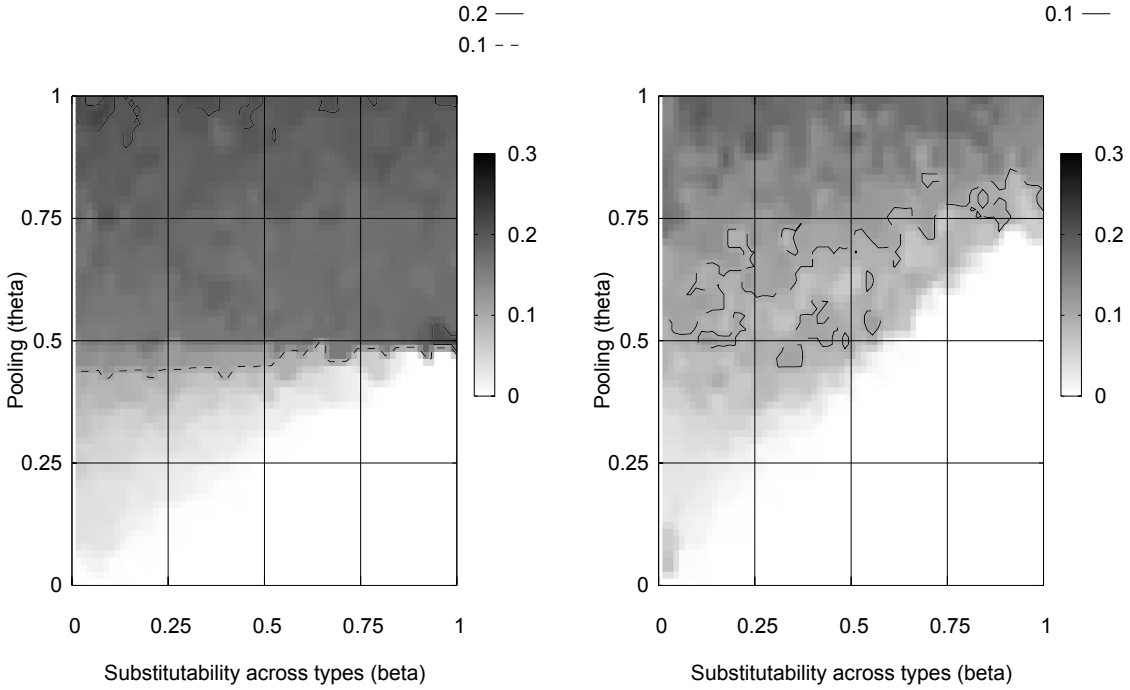

Figure 8: Dis-similarity between partners in the (β, θ)-space (left: no learning; right: learning).

5 Conclusion

In this paper we have addressed the issue of knowledge growth and network formation through the recombination of existing individual knowledge endowments in a collaborative process. The main issue concerns the tension between on one hand the reduction of uncertainty as a result of former successful interactions (learning to collaborate) and, on the other hand, the loss of mutual attraction when successful innovations bring partners' knowledge profiles closer together, a tendency which is stronger when learning is at work.

As driving forces of the model we considered two central parameters: the ease with which knowledge from different fields can be substituted for each other in the process of creating new knowledge (β the exponent in the CES function); and the ease with which pooling takes place between different knowledge agents prior to the conduct of innovative

research itself (θ in the model). The latter form of substitution rests on the degree of cognitive division of labour in the innovative process.

First when learning to collaborate is not part of the process we find that there is a critical relationship between θ and β where the process tips from a world of firms operating in isolation to a collaborative world with intensive networking. The critical θ^* is increasing with β, and the steepness of the transition between autarchy and a multilateral system is increasing with β, from very smooth for $\beta \approx 0$ to almost a step function for $\beta \approx 1$. Connectedness emerges very fast (i.e. for θ slightly higher than θ^*) for any value of β, and path lengths fall as the degree then keeps increasing with increasing θ. Cliquishness also mechanically increases with θ as a consequence of increasing degree, but when we look at the purely random counterpart of the emerging network, we see that there is more cliquishness and path lengths are longer than pure randomness would predict. This suggests a network with more structure than a random one, especially in the transition zone between autarchy and collaboration. Given the magnitude of the excess cliquishness and distances it would be difficult to claim definitively that a small world has emerged, but there is some structure to the networks that form.

Regarding the properties of knowledge accumulation, the most marked result is the decline of individual specialization as inter-individual pooling becomes easier. It may seem at odds with the general intuition that as labour division gets easier specialization should increase, but one should keep in mind the way new knowledge is allocated to individuals: the field in which novelty is created is common to both partners, and this precisely prevents specialization. In Cowan et al. (2003) a situation in which people where innovating most often in their domain of expertise was explored, and there of course as division got easier specialization increased. The contrast between these two results indicates the importance of understanding exactly how new knowledge is absorbed by economic actors.

Contrasting these results with those found when agents do increase their collaboration skills by experience, it first appears that under learning the critical relationship between β and θ is different, with autarchy demanding larger θ to vanish. This stems from the relative importance of two factors: agents' innovative success is always more likely in isolation, but the gains can be lower. Apparently the increased risk of failure is not compensated by potentially larger innovations jointly made, so all else equal, agents are more likely to innovate as individuals. In addition, in the region of the parameter space where collaboration does take place, the degree of the network is much smaller (by roughly a factor 5) which means that fidelity via collaborative learning is a strong force, even though repeated interactions can exhaust the innovativeness of a pair. The emergent network is completely connected only for significantly larger θ-values, and there is a thick frontier-zone where degree is low, the graph is disconnected and there is some excess cliquishness. In this region the network evolves towards a more structured state, individual agents concentrating their collaborations on a smaller number of strongly interconnected partners. In this narrow zone the innovation process is separable enough to allow partners to retain some differentiation in their knowledge profiles, but still joint enough to permit mutual learning and transfer. Rather than small worlds here we tend to see small disconnected groups relatively densely connected.

We have modelled a knowledge system with different levels of substitutability between agents and between fields, and examined the resulting patterns of networking and knowledge production. Looking at the effects of the two parameters, we see a strong similarity in the results patterns for the properties of knowledge and for network structures. This suggests very strongly that the two processes: changes in knowledge stocks, and emergence of network structures, are strongly linked, and that it would, in general, be a mistake to try to understand one without examining the other. In particular, we observed a strong effect of θ, the parameter capturing issues of how firms pool their knowledge, and how the joint innovative process works. This implies that to understand network formation it is necessary to know the detail of the processes by which firms jointly innovate. Different network structures emerge as the process of pooling knowledge, or dividing tasks changes. The model presented has, for reasons of simplicity and parsimony, has been very stylized in its treatment of knowledge and innovation. Nonetheless it very strongly suggests that the interplay between the decomposition of tasks in knowledge creation, the creation of trust through repeated interaction and the general properties of innovation as knowledge recombination is central in understanding how networks, and industry structures more generally, emerge and evolve.

References

[1] Coriat B. and Guennif S. (1998), "Self-interest, Trust and Institutions", in Lazaric N. and Lorenz E. (Eds.), *Trust and Economic Learning*, Edward Elgar.

[2] Cowan R. and Jonard N. (2004), "Network Structure and the Diffusion of Knowledge", *Journal of Economic Dynamics and Control*, 28, 1557-1575.

[3] Cowan, R., N. Jonard and J.-B. Zimmermann (2003), "La Dynamique Conjointe de la Connaissance et des Réseaux", *Revue d'Economie Industrielle*, 103, 253-274.

[4] Dodgson M. (1996), "Learning, trust and inter-firm technological linkages: some theoretical associations", in Coombs R., Richards A., Saviotti P.P. and Walsh V. (Eds.), *Technological Collaboration, the Dynamics of Cooperation in Industrial Innovation*, Edward Elgar.

[5] Dutta, B. and M. Jackson, (2003), "On the Formation of Networks and Groups", in Dutta, B. and M. Jackson (Eds.), *Networks and Groups: Models of Strategic Formation*, SpringerVerlag, Heidelberg.

[6] Dyer, J.H. and K. Nobeoka (2000), "Creating and Managing a High Performance Knowledge Sharing Network: The Toyota Case", *Strategic Management Journal*, 21, 345-67.

[7] Erev, I. and A. Roth (1998), "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria", *American Economic Review*, 80, 848-881.

[8] Gale D. and Shapley L. (1962), 'College admissions and the stability of marriage', *American Mathematical Monthly*, 69, 9-15.

[9] Galison, P. (1999), "Trading zone: Coordinating action and belief", In M. Biagioli (Ed.), *The science studies reader*, 137-160, New York: Routledge.

[10] Garcia-Pont, C. and N. Nohria, (2002), "Local versus Global Mimetism: The dynamics of alliance formation in the automobile industry", *Strategic Management Journal*, 23:307-21.

[11] Grandstand, O. (1996), "Towards a Theory of the Technology-Based Firm" *Research Policy*, 27, 465-489.

[12] Grandstand, O. and S. Sjolander (1990), "Managing Innovation in Multi-Technology Corporations", *Research Policy*, 19, 35-60.

[13] Hagedoorn J. (2001), "Inter-firm R&D partnerships. An overview of major trends and patterns since 1960", in: J.E. Jankowski, A.N. Link and N.S. Vonortas, *Strategic research partnerships: Proceedings from a NSF workshop*, Washington DC, National Science Foundation, 63-92.

[14] Hebb, D.O. (1949), *The Organization of Behavior: A Neuropsychological Theory*, New York: John Wiley Sons.

[15] Kirman, A. and N. Vriend, (2001), "Evolving Market Structure: An ACE model of price dispersion and loyalty" *Journal of Economic Dynamics and Control, 25*, 459-502.

[16] Klos, T. and B. Nooteboom (2001), "Agent-based computational transaction cost economics", *Journal of Economic Dynamics and Control, 25*, 503-526.

[17] Luce, R.D. (1959), *Individual Choice Behavior*. New York: Wiley.

[18] Mowery, D.C., Oxley, J.E. and B.S. Silverman, (1998), "Technological Overlap and Interfirm Cooperation: Implications for the Resource-based View of the Firm", *Research Policy*, 27, 507-523.

[19] Nooteboom B. (2000), *Learning and Innovation in Organizations and Economics*, Oxford University Press.

[20] Nooteboom B. (2004), "Learning and governance in inter-firm relations", *Revue d'Economie Politique*, 114, Janvier-Février.

[21] Olsson, O. (2000), "Knowledge as a Set in Idea Space: An Epistemological View on Growth", *Journal of Economic Growth*, 5, 253-276.

[22] Peretto, P. and S. Smulders (2002), "Technological Distance, Growth and Scale Effects", *Economic Journal*, 112, 603-624.

[23] Plourabouê F., Steyer, A. Zimmermann, J.B. (1998), "Learning Induced Criticality in Consumer?s Adoption Pattern: a neural network approach", *Economics of Innovations and New Technology*, vol. 6 p. 73-90.

[24] Powell, W.W. (1990), "Neither markets nor hierarchies: Network forms of organization" in B.M.Shaw and L.L. Cummings (eds.) *Research in Organizational Behavior,* vol. 12, 395-336.

[25] Roijakkers, N. (2003), *Inter-firm cooperation in high-tech industries: a study of R&D partnerships in pharmaceutical biotechnology*, PhD thesis Maastricht University, Maastricht, Datawyse/Universitaire Pers Maastricht.

[26] Sako M. (1991), "The role of trust in Japanese buyer-supplier relationships", *Ricerche Economiche*, 45, 449-474.

[27] Schumpeter, J.A. (1934), *The Theory of Economic Development.* Harvard University Press.

[28] Teece, D. and G. Pisano (1989), "Collaborative Arrangements and Global Technology Strategy: Some Evidence from the Telecommunications Equipment Industry" in R.A. Burgelman and R.S. Rosenbloom (eds.), *Research on Technological Innovation, Management and Policy*, 4, 227-256, Greenwich, CT: JAI Press.

[29] van Alstyne, W. and E. Brynjolfsson (1996), "Widening Access and Narrowing Focus: Could the Internet Balkanize Science?", *Science*, 271, 1479-1480.

[30] van Alstyne, W. and E. Brynjolfsson (1997), "Electronic Communities: Global Village or Cyberbalkans?" Working Paper MIT Sloan School.

[31] Yatchew, A. (1998), "Nonparametric Regression Techniques in Economics", *Journal of Economic Literature*, 36, 669-721.

[32] Watts D. and Strogatz S. (1998), "Collective Dynamics of Small-World Networks". *Letters to Nature*, 393, 440-442.

[33] Weitzman, M. (1998), "Recombinant Growth", *Quarterly Journal of Economics*, 113, 331-360.