

R. Cowan · N. Jonard · J.-B. Zimmermann

Evolving networks of inventors

Published online: 4 January 2006
© Springer-Verlag 2006

Abstract In this paper we develop a model of R&D alliance formation. Pairs of firms combine their knowledge in an attempt to innovate. Whether this attempt is successful depends in part on whether the pair has been successful in the past: accumulated experience teaches a pair of firms how to innovate together, but at the same time increases the similarity of their knowledge stocks. A tension exists between the desire for a familiar partner, and desire for a partner with complementary knowledge. How this tension is resolved depends on the nature of the innovation process itself, and the elasticity of substitution of different types of knowledge inputs in knowledge production. From the alliance-innovation process, a variety of networks form. In different parts of the parameter space we observe isolated agents, a dense, connected network, and small worlds.

Keywords Innovation · R&D collaboration · Networks · Knowledge · Computational experiment

JEL Classification L14 · Z13 · O3

1 Introduction

One of the long-standing legacies of Schumpeter is the view that innovation plays a central part in shaping the evolution of a dynamic economy. Also pivotal is Schumpeter's idea that innovation consists largely of the recombination of existing

R. Cowan (✉)
MERIT, University of Maastricht, P.O. Box 616,
6200 MD Maastricht, The Netherlands
E-mail: r.cowan@merit.unimaas.nl

N. Jonard
CNRS, CREA Ecole Polytechnique, 1 Rue Descartes,
75005 Paris, France
E-mail: nicolas.jonard@shs.polytechnique.fr

J.-B. Zimmermann
CNRS, GREQAM, EHES, 2 Rue de la Charité,
13002 Marseille, France
E-mail: Jean-Benoit.Zimmermann@univmed.fr

knowledge. The practical force of this idea is that if agents have access to more and a wider variety of knowledge or information, innovation and thus growth will be fostered.¹ Recently, changes in the technological landscape have made this issue more important than ever before. It has been argued that technologies both being used and being produced involve technological expertise that covers a much broader range of disciplines than has hitherto been the case.² What this implies is that the types of knowledge necessary to innovate and compete successfully can lie outside a firm's main area of expertise. One strategy to address this challenge is to look outside the firm's boundaries. Collaboration with other firms or other institutions more generally can serve to transmit knowledge, either codified or tacit. It is no surprise, then, that we observe a rapid increase in the number of inter-firm strategic alliances in recent decades.

Nevertheless, collaboration is risky in the sense that it is marked by uncertainty relative to both the skills of the partner and his reliability (Powell 1990, p. 318, provides a discussion of those risks). Successful collaboration demands mutual knowledge and sharing of routines, representations and ways of thinking, that is, a form of proximity that the experience of collaboration permits to build (Garcia-Pont and Nohria 2002). Partners have to learn to collaborate. The repetition of interactions creates information that permits partners to reduce uncertainty and increase predictability regarding each others' behaviour.³ Also, through the repetition of interactions a common language develops, an intermediary-level knowledge specific to the partners, whose availability increases the efficiency of collaboration (see Galison 1999, for a study of the field of experimental physics, and how the dialogue between theorists, experimenters and instrument makers was made possible by the emergence of a jargon specific to their purpose).

This discussion suggests that in a variety of ways, after a successful interaction the two parties will find each other more attractive than they did before. Successful interactions increase partners' understanding of each others' motives, increase the awareness of each others' skills, and generate more common tacit knowledge which improves communication. Empirically, firms with a history of partnering are more likely to have alliances than those without (Powell et al. 1996), and two firms that have worked together in the past are more likely to choose each other as partners than to choose new firms (Roijakkers 2003).

At the same time, repeated collaboration tends to increase partners' similarity, possibly reducing their mutual attractiveness. Nooteboom (2004) notes that "... ongoing interaction will yield a reduction of cognitive distance". Mowery et al. (1998, p. 517) find that "... technological overlap between joint venture partners after alliance formation is greater than their pre-alliance overlap". (See also Dyer and Nobeoka 2000.) But increased similarity means that partners can have less to contribute to each other. In the extreme, if they become identical in what they know there is no reason for them to collaborate, no matter how much they trust each other or how smoothly they can interact.

While the strategic alliance tends to be a bilateral relationship, firms can have, over time, more than one partner. Thus in an industry in which strategic alliances

¹ On an application of that idea to a growth model, see for instance Weitzman (1998).

² See the discussion of that trend in Powell et al. (1996), Grandstand and Sjolander (1990), Grandstand (1996), and Teece and Pisano (1989).

³ On that issue see Coriat and Guennif (1998), Sako (1991), Dodgson (1996) and the model by Klos and Nooteboom (2001).

are common, networks arise: collections of firms, each of which has some non-market relationship with a small number of firms in the industry. There has been considerable empirical interest in the structure or architecture of the networks that arise in different industries in different periods (see for example Powell et al. 2005 for a recent contribution).

In addition to this empirical interest, there is now a growing literature in economic theory on network formation, but the majority of it treats the problem in game-theoretic terms, looking for stable structures that emerge from agents' one-time strategic decisions about which links to form. In general in that literature it is difficult to provide an exhaustive description of the stable structures (a notable exception being Goyal and Joshi 2003), the studies tend to focus on the stability of some specific architectures (the star, the wheel, the complete network for instance) and a tension between stability and efficiency is often identified (for a recent survey, see Dutta and Jackson 2003). The results rest upon strong assumptions about what agents know (that is, everything) and very poorly address the adaptive, path-dependent nature of network formation and operation. By contrast, the model developed in this paper continues work using another approach, centered on the idea that agents are continually forming and breaking links with each other, modifying their own characteristics, and that this ongoing activity is what underlies the networks that we observe. Agents choose with whom to form a link in order to achieve some immediate goal which modifies their properties. The repetition of these actions results in the formation of a network that evolves over time.⁴

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. We abstract from many features of alliance formation and focus only on the effects that follow from the goal of knowledge production.⁵ Firms combine 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 other. By repeated alliance formation and dissolution a network emerges and continuously evolves. How this structure and the nature of the economy's knowledge stock change in response to parameters governing innovation are the issues that we explore.

2 The model

Consider an industry in which each period a finite number of firms form alliances, with the goal of creating new knowledge. An alliance having formed, the partners pool their knowledge to create a joint stock and, by combining elements of that joint stock, seek to innovate. If the project succeeds, new knowledge is created and added to each partner's knowledge stock. Then the alliances are dissolved and the process is repeated.⁶ Formally, the population of firms is denoted $S = \{1, \dots, n\}$ (n is

⁴For related approaches to adaptive network formation, see for instance Cowan et al. (2003), Kirman and Vriend (2001), Plouraboué et al. (1998), and Weisbuch et al. (2000).

⁵See Oliver (1990) or Narula (1999) for discussions of the motives underlying alliance formation.

⁶The processes of knowledge pooling and production, described in detail below, are similar to those we use in Cowan et al. (2004b).

even). Each $i \in S$ is characterized by a knowledge endowment of $\ell \geq 2$ types of knowledge represented as a vector $\alpha(i) = (\alpha_l(i); l = 1, \dots, \ell)$, where $\alpha_l(i)$ represents the amount of knowledge of type held by i . Firms are thus treated as located in an ℓ -dimensional knowledge space.

2.1 Knowledge production

There are many ways to characterize knowledge, none of them without its pitfalls. The approach proposed here is simple and flexible. It obeys the following intuitive requirements. First, when innovation is jointly conducted, it is natural to expect that the post-innovation knowledge stocks held by the partners are larger than their pre-innovation stocks. Second, after innovation, the similarity of the knowledge profiles of the partners (that is the relative distance between them in the underlying knowledge space) has fallen.

Formally, when a partnership, ij , forms, i and j first combine their knowledge into a joint vector $\alpha(ij)$. This is done in each category $l = 1, \dots, \ell$ through

$$\alpha_l(ij) = (1 - \theta) \min \{ \alpha_l(i), \alpha_l(j) \} + \theta \max \{ \alpha_l(i), \alpha_l(j) \}. \quad (1)$$

The joint knowledge vector becomes the input to the knowledge production process.

In Eq. 1, θ expresses 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 agents will specialize, each agent doing some sub-tasks, and bringing the results together at the end to create the complete innovation. The better econometrician will do the econometrics, the better game theorist will do the game theory, and the joint knowledge vector will consist of the maximum level of knowledge of each type. This is captured by setting θ equal to 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: joint knowledge is the minimum of each type, and θ will approach zero.

The motivation for θ has been the separability of the innovation task, which maps nicely into the range $0 \leq \theta \leq 1$. However, there is a second interpretation of θ , which connects explicitly with the characteristics of partnerships. Do I look for partners who are different from me, or similar to me? We can address this by asking whether an agent prefers joint or autarchic innovation. If θ is small, joint knowledge is driven by minimum values. If j is worse than i in any knowledge type, then the joint profile looks worse than i 's own profile. Similarly if i is worse than j anywhere, the joint profile is worse than j 's profile. A weaker partner pulls me down. In the extreme, if θ is zero, I will only be interested in partners who are in no way worse than me. Since this is reflective, only identical firms can be partners. By contrast, if θ is large, no partner can make me worse off (since if you are worse than me anywhere, we use my knowledge there). Again this is reflective, thus the most natural partner is one who is strong where I am weak, and vice versa. By this reasoning, θ indicates a taste for dissimilar partners. In the literature there is some disagreement about whether innovative success increases or decreases with the distance between partners, i.e. whether a firm prefers similar or dissimilar partners. Mowery et al. (1998) find a U-shaped relationship, a conclusion also supported by

Nooteboom (2000). In the models by van Alstyne and Brynjolfsson (1997) and Peretto and Smulders (2002) on the other hand it is assumed 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. Thus there is the possibility of a relatively complex relationship due to the multi-dimensional nature of firms' knowledge.

The joint knowledge vector serves as the input to the innovation process. To formalize this, we use a standard constant elasticity of substitution production function

$$\phi(\alpha(ij)) = \left[\sum_l \alpha_l^\beta(ij) \right]^{1/\beta}. \quad (2)$$

The parameter $0 < \beta \leq 1$ is an inverse measure of the elasticity of substitution across knowledge types, which is defined as $1/(1-\beta)$. When β is very small it is difficult to substitute between different types of knowledge, and the CES function approaches a fixed-coefficients production function: output is determined by the minimal input. Here, if an agent has a weakness in one knowledge type, he has strong incentives to find a partner who is strong there, in order to create a flatter input profile. By contrast, if β is large, it is relatively easy to substitute between different types of knowledge. If an agent has a weakness in one type, this can be offset if he is strong somewhere else. The incentive to find a partner is thus weaker. This suggests that incentives to find partners decrease with β , and consequently so should the amount of networking we observe.

2.2 Innovative success and experience

Innovation is an inherently uncertain activity. A project may fail, and the anticipated new knowledge may not be created. It also seems reasonable to suppose that firms will not undertake projects with extremely low probabilities of success. Thus, as an initial step, we bound the success probability of any project between $\underline{\pi}$ and $\bar{\pi}$ ($0 < \underline{\pi} < \bar{\pi} < 1$). However, the success probability of a partnership can change over time and here history is important, because project success is driven in part by familiarity of the partners, as discussed above.

Consider period t has just elapsed, and firms are to make decisions for the new period $t+1$. Formally, for all $1 \leq s \leq t$ and all $i \neq j$, define the binary variable $\eta_s(ij)$ as taking the value 1 if there has been a successful collaboration between i and j , at time s and 0 otherwise. Then define

$$\gamma_t(ij) = \sum_{1 \leq s \leq t} \eta_s(ij) \rho^{t-s}, \quad (3)$$

where $0 < \rho < 1$ is a discount factor, to be the measure of i and j 's historical success. $\gamma_t(ij)$ is a discounted sum of past experiences. A history of joint success will improve a pair of firms' ability to work together, and so increase the probability of future successes. But experiences further in the past weigh less heavily in future success, and so are discounted. Now define $\pi_t(ij)$ as the probability that a collaborative attempt by i and j in period $t+1$ is a success, conditional on their history up to t . Success, as measured by $\pi_t(ij)$, should increase in $\gamma_t(ij)$ while

remaining bounded by the minimum and maximum success probabilities, $\underline{\pi}$ and $\bar{\pi}$. There is no obvious choice for the functional relationship between $\pi_t(ij)$ and $\gamma_t(ij)$,⁷ so to maintain a gradual impact of $\gamma_t(ij)$ on $\pi_t(ij)$ we use a simple linear form. Noting that

$$\gamma_t(ij) \leq \sum_{1 \leq s \leq t} \rho^{t-s} = \frac{1 - \rho^t}{1 - \rho}, \quad (4)$$

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

$$\pi_t(ij) = \underline{\pi} + \gamma_t(ij)(\bar{\pi} - \underline{\pi})(1 - \rho). \quad (5)$$

On this formulation, the first time two firms attempt to innovate together, because they have no history, their success probability is $\underline{\pi}$. By contrast, when a firm attempts to innovate alone, since there is no issue of learning to cooperate with oneself, we assume that its success probability is the maximum, $\bar{\pi}$.

Finally the expected amount of knowledge produced by a collaboration between i and j at time t can be expressed as

$$F(ij) = \pi_t(ij) \cdot \phi(\alpha(ij)). \quad (6)$$

This represents the quantity produced by a successful venture, multiplied by the probability of success.⁸

If the partnership succeeds in innovating, the new knowledge is added to the partners' respective knowledge stocks. 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. We here assume that there is some path dependence in knowledge production, so output will resemble input. To simplify we assume that one type of knowledge is produced, but it resembles the input probabilistically:⁹ the probability of the new knowledge being of type m is

$$\frac{\alpha_m(ij)}{\sum_l \alpha_l(ij)}. \quad (7)$$

2.3 Pair formation and equilibrium

We draw on the literature on matching problems for our basic model of pair formation. We have a single population of agents, and wish to allow for the

⁷ An exponential specification was also tried, but the effect of changes in $\gamma_t(ij)$ 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.

⁸ Rating potential partners according to their expected values is formally similar to the score approach in Klos and Nootboom (2001). In their model the score of i is increasing with my trust in i 's intention to honour agreements to the best of his ability (p. 514). This subjective belief however does not correspond to an objective risk, as in their model transactions never fail. We abstract from this sort of difficulty by simply viewing $\pi_t(ij)$ as a measure of i and j 's ability to work together.

⁹ This assumption regarding the distribution of new knowledge is not critical. For example, one might suppose that each partner creates slightly specific knowledge – partly associated with the joint profile, and partly associated with his idiosyncratic profile. Provided the weight on the joint profile is not trivial, the results are qualitatively identical.

possibility of autarchic innovation, or in this context, self-matching. Thus we have a generalization of the Roommate matching problem (Gale and Shapley 1962). Every agent ranks every agent in the economy (including himself) as a potential partner. Based on these rankings partnerships form, and the population is partitioned into q single agents and $(n-q)/2$ pairs. We use μ to denote such a matching, where $\mu(i)$ is the partner of i .¹⁰ A matching is stable if there is no pair of agents who would block it. Formally, writing $j \succ_i k$ to express that j is preferred by i to k , a matching μ is stable if there is no pair $ij \notin \mu$ such that $j \succ_i \mu(i)$ and $i \succ_j \mu(j)$. That is, there is no pair ij who are not matched under μ but who prefer each other to their respective partners in μ .

As our agents are interested in knowledge creation, a natural way to rank prospective partners is by the expected amount of knowledge a partnership would produce: i prefers j to k if the partnership ij produces more knowledge (in expectation) than ik . Formally $j \succ_i k$ if and only if $F(ij) > F(ik)$, a calculation any agent is able to make to create a complete transitive ordering.¹¹

In general, in a roommate matching problem, stable matchings do not always exist, and when they do they are not guaranteed to be unique. Our model, however, is a special case (though still relatively general) and a unique stable matching always exists due to the fact that both parties to a partnership assign it the same value (see Proposition 1 in Cowan et al 2004a). The general idea is that of all potential partnerships, there is one that yields the highest payoff (and both partners agree what the payoff is). This partnership must be part of any stable matching. A simple recursion of this observation generates the unique stable matching.

3 Computational experiment

We study a population of $n=100$ firms each endowed with a five-category knowledge vector. At the outset, individual knowledge endowments are independently and identically distributed over the unit interval. Each period, firms form pairs (or decide to stay alone) to conduct innovative activities. This list of pairs and singletons constitutes a stable matching in which the value of a pair (or singleton) is equal to the expected amount of knowledge produced by that pair (or singleton). Any new knowledge created within the partnership is added to both firms' existing knowledge. In addition, firms record their experience with their partners. Partnerships dissolve, and the next period begins. We iterate this process for 1,000 periods.

We are interested in understanding the effects of knowledge pooling (θ), substitutability in the production of new knowledge (β), and whether or not history, or learning about a partner, matters in success probabilities. We construct a 100×100 grid of equally spaced points in the (β, θ) -space (the unit square) and record data for the entire history of the industry, under the two learning regimes (history does or does not affect success probabilities) for which we set $\underline{\pi} = 0.9$ to $\bar{\pi} = 0.99$.

We examine two aspects of the emergent structure: the properties of knowledge accumulation and distribution; and the properties of the network. Our characterization of knowledge is as a multi-dimensional resource, thus to generate

¹⁰ A matching is thus a bijection $\mu : S \rightarrow S$ such that $\mu(\mu(i))=i$ for all $i \in S$.

¹¹ In case of a tie we apply an arbitrary rule to guarantee that the score function is still strict: if $F(ij)=F(ik)$ then $j \succ_i k$ if and only if $j <_i k$.

statistics, it is useful to reduce this vector to a scalar. Since knowledge production is the assumed goal of our firms, the scalar we use is a firm's innovative performance $\phi(\alpha(i))$. By simply treating a firm's knowledge stock as the quantity the firm could produce in isolation, an accessible presentation of the results is obtained.

The results on network structure are slightly more involved. 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 partners have interacted. Formally, in each period t denote A_t the adjacency matrix associated with the stable matching μ_t , that is $A_t(i, j) = 1$ iff $j = \mu_t(i)$. The weighted network recording past interactions is then denoted B_t , where $B_t(i, j) = \sum_{1 \leq s \leq t} A_s(i, j) / t$. To move from B_t to a proper adjacency matrix again, distances must be computed first. Define the distance $d(i, j)$ between two firms i and j as the number of edges in the highest frequency path linking them. Any path i_0, i_1, \dots, i_ξ with $i_0 = i$ and $i_\xi = j$ has an associated frequency $\prod_{l=1, \dots, \xi} B_t(i_{l-1}, i_l)$ and a length $\xi \geq 1$. There is a finite number of such paths, and thus a path with maximum frequency exists: its length ξ is defined as the distance $d(i, j)$. Two agents are neighbours if the distance between them is one, and the adjacency matrix unfolds. Using this reduction to an unweighted graph permits us to use standard descriptive statistics to characterize network structures.

The results are presented in two-panel figures, where each panel is a shaded contour plot. Such a plot reads as a map in an atlas: darker grey scales imply higher values on the z axis. Contours are added to make patterns clearer. This provides a compact display of the relationship between the substitutability of knowledge types in production (β), the pooling parameter (θ), and the performance measures we are concerned with. Left panels correspond to situations where past experience has no effect on future success (the initial or a priori success probability is $\bar{\pi} = 0.99$). Right panels display effects of the same parameters when past success increases the probability of future success, with a success probability increasing from $\underline{\pi} = 0.9$ to $\bar{\pi} = 0.99$ as past successes accumulate.

3.1 Knowledge

Firms are primarily interested in knowledge creation, and form partnerships precisely for that purpose. At the aggregate level, this leads us to ask how aggregate knowledge grows, whether knowledge is evenly distributed over firms, whether firms become specialists, and finally whether firms all become similar. Regarding the first question, any pattern in aggregate knowledge growth is driven entirely by the parameter β . Indeed, as β increases from 0 to 1 the marginal product of the knowledge inputs falls quite dramatically, particularly near zero. Thus aggregate knowledge as measured by total innovative performance simply increases as β falls.

3.1.1 Knowledge distribution

Over the history of the economy, each agent creates a certain amount of new knowledge. As agents begin this history with similar knowledge levels, final levels

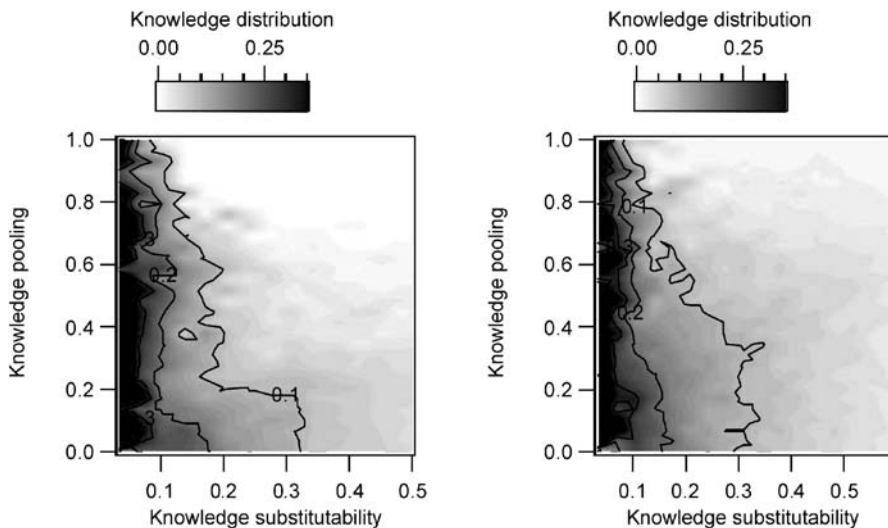


Fig. 1 Coefficient of variation of innovative performance in the (β, θ) -space (*left*: no learning; *right*: learning)

can be used to examine the distribution of knowledge generated by innovation. Figure 1 shows the relationship between the coefficient of variation of final knowledge levels and the pair (β, θ) .

The general patterns are the same when experience matters and when it does not. Inequality in knowledge distribution is decreasing along the β axis with a sharp peak when β is close to 0, with only a marginal effect of θ . The effect of β is explained through the nature of the CES production function. As noted above, a small β implies a very high marginal product. Thus when β is small, well-endowed agents will make large innovations. They will therefore grow faster than less well-endowed agents, which will preserve, and often magnify initial differences.

3.1.2 Specialization

A generalist firm can be thought of as one having roughly equal amounts of all types of knowledge. A specialist, by contrast, will have an identifiable area of expertise – having noticeably more knowledge of one (or a few) types than the others. To measure specialization at the firm level, we compute the coefficient of variation of the firm's endowment. For a single firm this ranges from 0 (for the perfect generalist) to 2 (for the perfect expert, with only one type of knowledge). Figure 2 reports the population average of individual coefficients of variation at each (β, θ) pair.

The effect of θ is clear, with the degree of specialization falling as pooling moves towards the maximum of the partners' endowments. Agents become highly specialized when they innovate largely as individuals, which happens increasingly as θ decreases. What drives this results 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

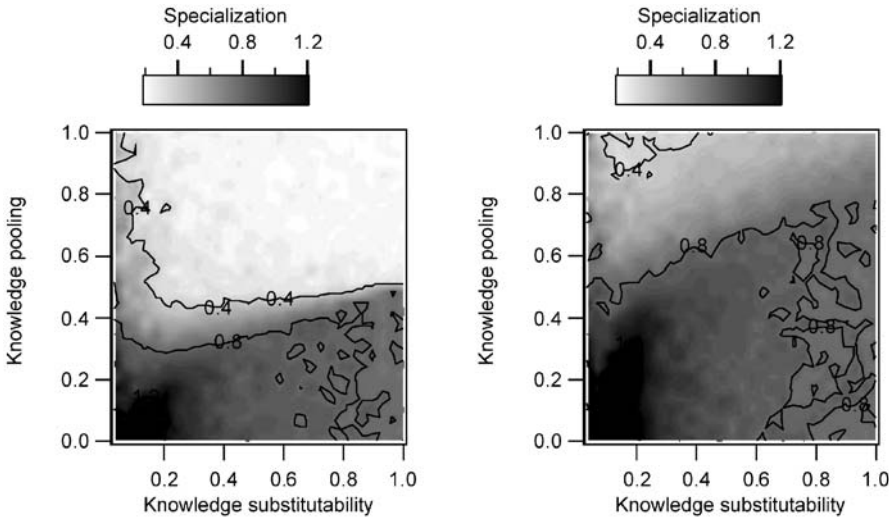


Fig. 2 Average agent specialization in the (β, θ) -space (*left*: no learning; *right*: learning)

specialization. When alliances form in a more systematic way (larger θ) generalist profiles quickly emerge. Sometimes an agent will innovate in his specialty, sometimes in his partner's, and for large θ , an agent will have many different partners. This will both smooth the agent's profile, 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. The effect of β is comparatively much weaker: as we move from right to left, initial differences in knowledge profiles are increasingly amplified, yielding the strongest specialization when β is close to 0.

3.1.3 The homogenization of firms

Quantifying convergence in knowledge profiles of firms can be done by computing the average angle between a firm and its partners in knowledge space. (This average is weighted by the frequency of interaction.) The larger this angle is, the more different (or complementary) is the expertise in the pairs that have formed. Figure 3 shows the population average of those weighted means in the (β, θ) -space.

The general intuition is that dissimilarity will be large when firms look for partners different from themselves (complementary, i.e. forming large angles) which happens for high θ . For low θ by contrast only pairs of similar people will form, contributing to a low angle over history. Thus dissimilarity is expected to fall when θ falls. This is the effect we observe in Fig. 3. In addition there is another, more mechanical effect. Given firms have constant opportunities for partnerships, those with a smaller number of partners will have lower average angle to them than those with a large number of partners. Number of partners is also expected to fall when θ falls. So in this case the strategic and mechanical effects act in the same direction to produce the patterns seen in Fig. 3.

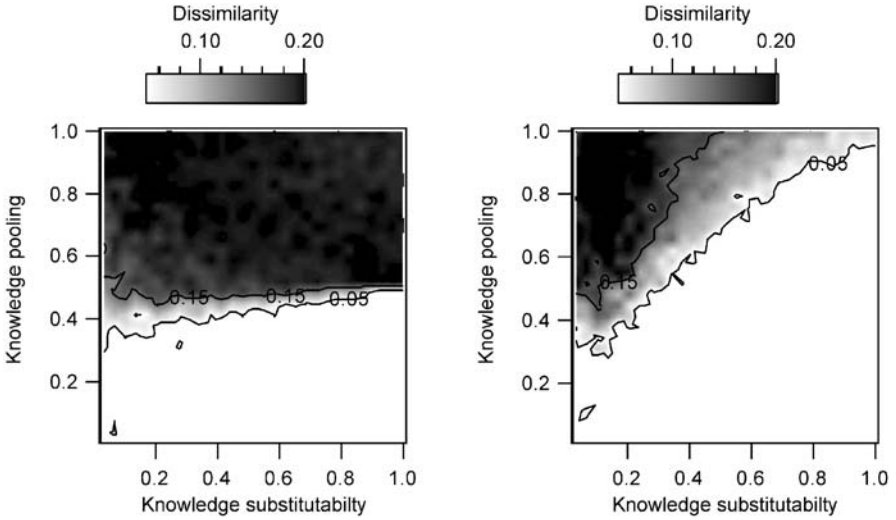


Fig. 3 Weighted average dissimilarity between partners in the (β, θ) -space (*left*: no learning; *right*: learning)

3.2 Network

Three statistics are used to describe the networks that emerge from the process of collaborative R&D. Defining Γ_i as the set of neighbours to whom i is directly connected and letting n_i be the number of those neighbours ($n_i = \#\Gamma_i$), the average degree of the graph is

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

and measures the density of the interaction structure. Characteristic path length is defined as

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

and simply measures how distant vertices are on average. Finally, neighbourhood clustering is the share of active links between the neighbours of any given vertex, that is, the proportion of existing over possible triangles in a firm's neighbourhood. Average clustering is then written as

$$\frac{1}{n} \sum_i \sum_{j, l \in \Gamma_i} \frac{\omega(j, l)}{n_i(n_i - 1)},$$

where $\omega(j, l) = 1$ if $j \in \Gamma_l$ and 0 otherwise.

3.2.1 Degree distribution and economic performance

Figure 4 shows the number of connections held by the average firm, that is, how many distinct partners a firm has on average over the history of the economy.

The first observation is that including the effect of experience on success changes the order of magnitude by a large factor. When experience matters, success in a partnership increases the relative attractiveness of those two firms to each other, and this can be a strong source of inertia in partnership formation. The process is described, for instance, by Gulati and Gargiulo (1999): firms tend to stick with partners they have had in the past to reduce search costs and the risk of opportunism in inter-firm ties.

Both panels show a frontier between autarchic and joint innovation: for low values of θ (implying that knowledge is pooled on the minimum) firms prefer to innovate in isolation. When θ approaches 0, and knowledge pooling uses the minimum level of the two partners, a firm cannot be made better off by having a partner. If my partner is worse than me in any category, my knowledge production falls, since the minimum of our knowledge serves as the input. Thus for low values of θ , firms prefer isolation. The strength of this effect depends, though, on a firm's ability to compensate for its own weakness with its own strength, that is, on the elasticity of substitution in production $1/(1-\beta)$. If it is difficult to substitute internally (small β), the need for a partner to fill a hole in my own knowledge is strong, and it may be worth accepting a partner who reduces the input in one category because he increases it in another. Thus the threshold value of θ , below which no partnerships form, increases in β .

The effect of θ on the density of the network operates through the nature of optimal partnerships and the dynamics of knowledge. If dissimilar partners are the norm (as is the case when θ is large), repeated partnering with the same partner drives firms too close together and they soon switch to another. Firms thus have many partners. By contrast, when similar partners are desirable (θ is small), sticking with the same partner is a good thing, thus firms tend to have few partners.

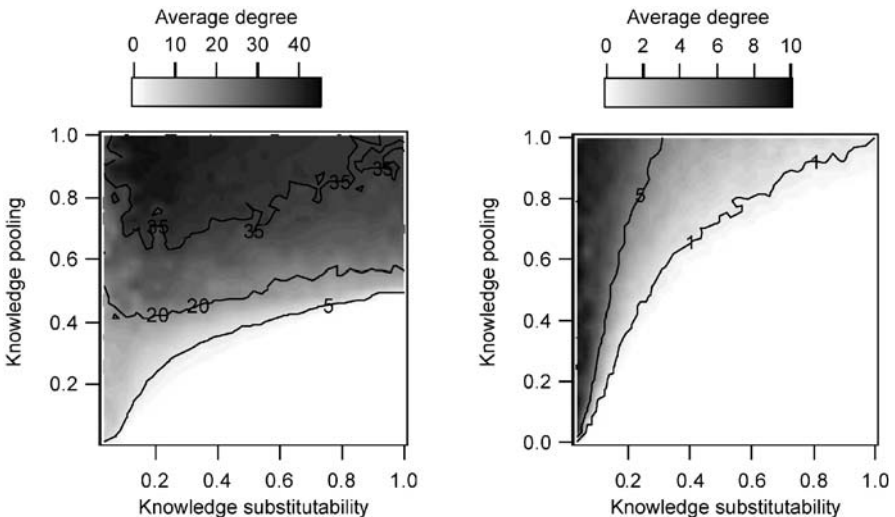


Fig. 4 Average degree of the network in the (β, θ) -space (left: no learning; right: learning)

Table 1 Correlation between degree and performance at the firm level

		β						
		0	0.17	0.33	0.50	0.67	0.83	1
No learning								
1		0.875	-0.054	-0.191	-0.030	-0.019	-0.372	-0.028
	0.83	0.935	0.073	0.093	-0.181	-0.121	0.161	-0.407
	0.67	0.836	0.745	0.442	0.496	0.107	-0.112	0.104
θ	0.50	0.738	0.684	0.401	0.478	0.416	0.076	δ
	0.33	0.704	0.353	-0.008	-0.413			
	0.17	0.449	-0.139					
	0							
Learning								
1		0.731	0.007	0.103	0.195	0.319	0.014	-0.265
	0.83	0.531	0.448	0.278	0.443	0.156	-0.254	
	0.67	-0.207	0.299	0.391	0.245			
θ	0.50	0.327	0.322	0.176				
	0.33	0.571	0.221					
	0.17	0.291						
	0							

This effect explains the increased density of the network as θ increases. The corollary of the increase in density is that as θ increases, networks are more likely to be connected.

Finally, comparing the two panels, we see that when experience matters autarchy is more attractive: it takes place in a larger area of the parameter space. The explanation is in terms of the likelihood of innovative success. We have assumed that when experience does not matter, success probabilities are high, so there is no penalty (from that source) in forming a partnership. By contrast, in a partnership, success probabilities are by definition lower than for autarchic innovation. Thus this penalty has to be overcome before a firm will consider joint innovation.¹²

Switching focus from the macroscopic properties of networks to economic performance, we can ask whether the firms who engage in many partnerships are better off than the ones who do not.¹³ We thus calculate the correlation coefficient between degree and performance as measured by innovative potential. The results are summarized in Table 1, in which θ values decrease from top to bottom, to mimic the figures.

Over most of the parameter space where networking takes place there is a (strong) positive association between degree and performance. The firms with largest innovative value are also those with larger degree. Partnerships are thus essentially a valuable asset. There are two regions where negative correlations

¹² This result depends entirely on the assumption that success probabilities are high when experience does not matter: it would be reversed had we assumed that success probabilities are low when experience does not matter.

¹³ Note that being involved in more alliances rather than fewer does not imply higher costs, the cost of forming an alliance being assumed negligible. Adding a cost to creating and maintaining partnerships would render autarchy preferable over a larger parameter region.

occur, though. The first is along the autarchy frontier. In this region firms that have relatively “good” endowments tend to innovate alone, those with weaknesses in their own stock are forced to look for partners. Here, partnership signals a weakness that needs be overcome. The second region is in the top right triangle in the table, an area where both θ and β are high. From Fig. 2 we can see that this corresponds to the region in which firms tend to be generalists. Here, the negative correlation is driven by firms that have fewer than average partners over history. Because of the convergence caused by innovation, a long-lived pair must consist of two specialized firms (otherwise they converge quickly and part company). When specialists come together, they gain a lot from the partnership and so will make relatively large innovations, particularly compared to a pair of generalist firms. Long-lived pairs of specialists will create more knowledge than pairs of generalists (which are common in this region) and this creates a negative correlation between density and innovation performance.

3.2.2 The emergence of local and global structure

Figure 5 displays the average distance between agents in the network. Because a disconnected graph has some agents who are infinitely distant from each other, the averages are computed only for networks that are connected. Thus this figure also indicates when a single connected component emerges, namely when θ is large enough for given β . Comparing the two panels, this critical θ is much larger when experience with partners matters. As discussed above in Section 3.2.1.

In both panels the contour line indicating an average degree of 5 (from Fig. 4) has been added. Though the parameter region over which a unique component forms is much larger when experience does not affect success probabilities, both panels contain a region where the degree is approximately 5, but the network is

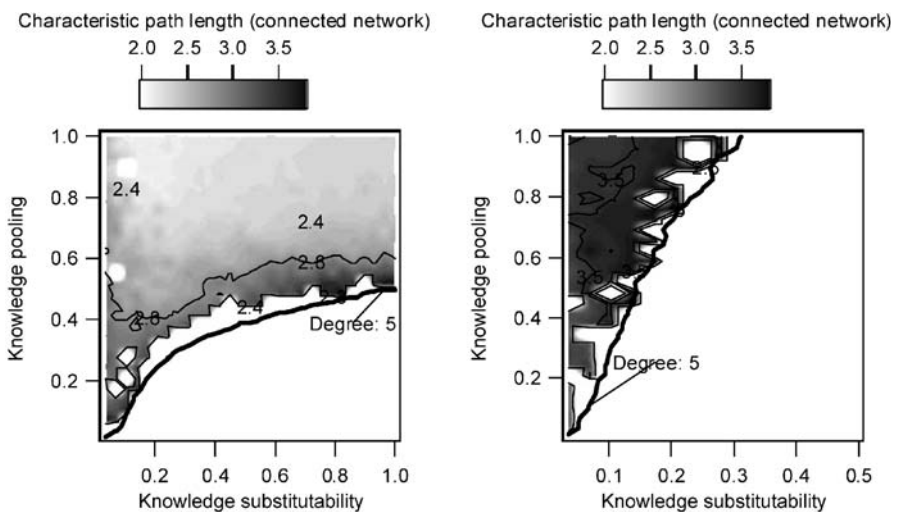


Fig. 5 Characteristic path length in the (β, θ) -space (*left*: no learning; *right*: learning)

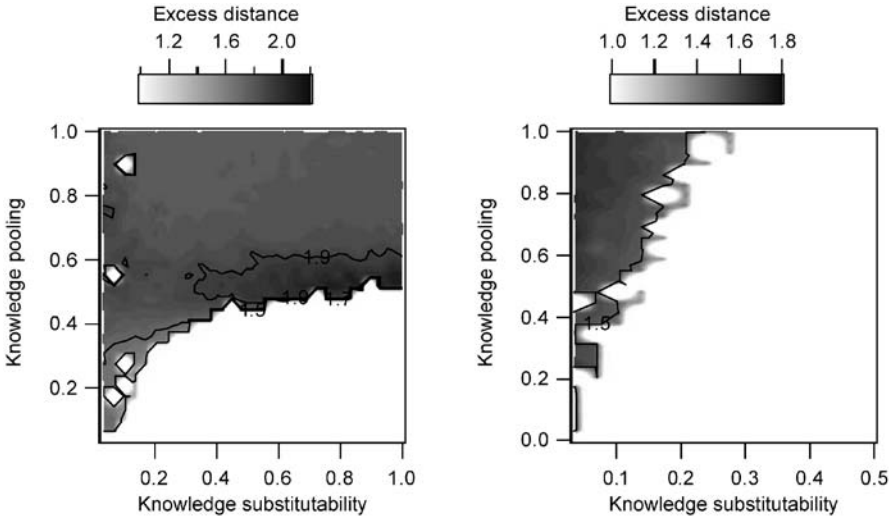


Fig. 6 Excess characteristic path length in the (β, θ) -space (*left*: no learning; *right*: learning)

disconnected. This region lies near the autarchic innovation border, and here the network consists of connected sub-networks. The pattern in the (β, θ) -space that Fig. 5 shows is essentially symmetric to the one observed in Fig. 4: as the average degree increases the average distance falls mechanically.

To get a finer idea of the extent to which structure is emerging, a useful benchmark is a random graph of the Erdős and Rényi (1960) type, in which a constant (average) degree is assumed for each vertex. In such a graph, when average degree is Δ , the expected path length is well-approximated by $\ln n / \ln \Delta$.

In Fig. 6 the ratio of the observed characteristic path length to that of the Erdős and Rényi benchmark is shown. In both learning regimes path lengths are slightly longer than those in equivalent random graphs. The networks that emerge from the processes we model are not random, but have a greater prevalence of localized connections.

Consider now clustering as a measure of local order in the network. Figure 7 shows the population average clustering. The relationship between clustering and the two parameters θ and β is very similar to that for degree. Since path length does not change dramatically as links are added to the network (Fig. 5), this implies that as firms find new partners they are not creating shortcuts to distant agents (which would reduce path length), but are rather reinforcing local coherence. (The white regions at the bottom of the graphs correspond to the region in parameter space in which firms tend to innovate as individuals, yielding zero clustering by definition.)

Again clustering 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 becomes denser locally. Thus it is necessary to compare clustering also with the Erdős and Rényi benchmark, where clustering is approximately Δ/n . Figure 8 shows the ratio of observed clustering over the benchmark. Values significantly larger than 1 would indicate a structure richer than a random graph.

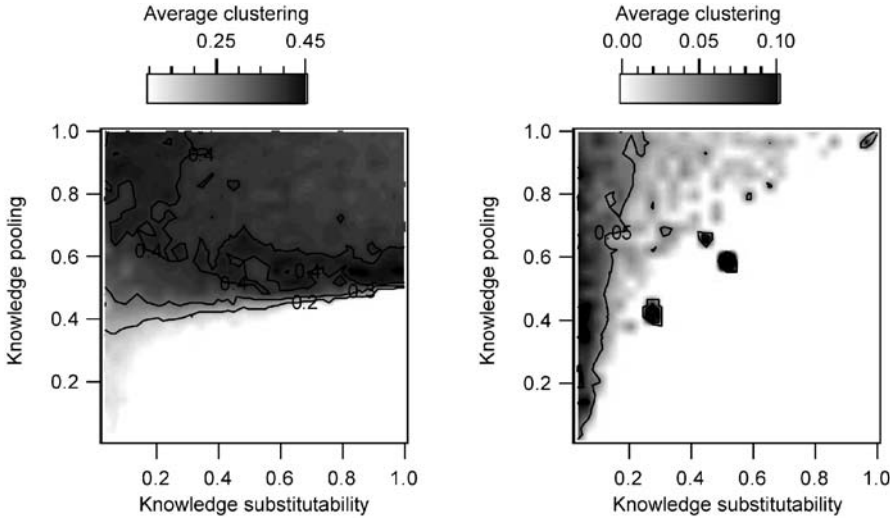


Fig. 7 Clustering in the (β, θ) -space (*left*: no learning; *right*: learning)

In the left panel, over almost the entire parameter space there is more clustering in our networks than there is in random graphs. In the right hand panel, clustering is lower than or similar to that of random graphs in the upper left corner of the parameter space, and larger elsewhere. The interesting feature is that in both panels, excess clustering is present, and strong along the autarchy frontier. In this region, θ is just above the threshold at which networking begins, stable matchings

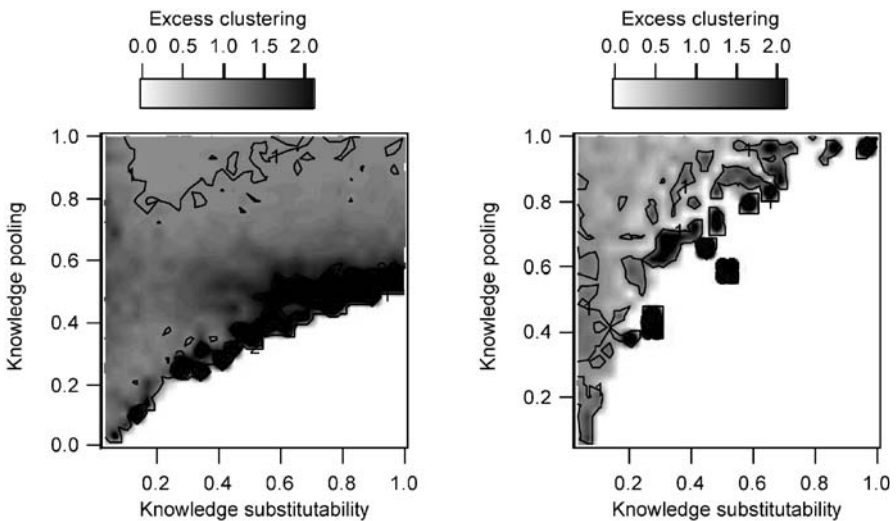


Fig. 8 Excess clustering in the (β, θ) -space (*left*: no learning; *right*: learning)

unite agents who are similar to each other. Here 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. It is likely that this area of the parameter space will engender partner-swapping. There is, then, a suggestion that cliques tend to form here, which together with the existence of longer path lengths (as compared to a random situation) is evidence partly supporting the small world conjecture.¹⁴ By extension, since the desire for similarity falls as θ increases, partnerships dissolve more rapidly, and when they do, firms search for very different partners. This takes them away from firms with whom they have already partnered, and also away from partners of partners.¹⁵

4 Conclusion

In this paper we have modelled the processes of knowledge growth and network formation. Both result from firms' decisions to form bilateral alliances in which they pool their existing knowledge and use it to create new knowledge. For a single firm, repeated interaction with the same partner has a positive effect by increasing the probability of successful innovation, but also has a negative effect through generating similarity in knowledge stocks. Familiar partners, by having similar knowledge profiles, will have small innovations with high probability; unfamiliar partners, with dissimilar profiles, will have larger innovations with lower probability. The resolution of this trade-off determines the nature of the networks that emerge, and this depends both on substitutability of inputs in knowledge production, and on whether the innovation task is decomposable into independent sub-tasks.

Specifically, two central parameters drive the model: the ease with which knowledge from different fields can be substituted in the knowledge production function (β), and whether the pooled knowledge vector lies closer to the maximum or the minimum of the partner's vectors, (θ), which is determined by the extent to which the innovative task is decomposable.

The most striking result regarding knowledge accumulation is the decline of specialization as separability increases. Separability implies that a division of labour is possible in innovation, and one might expect that this will make it possible for firms to specialize. The contrary result that we find arises from the assumption that, probabilistically, the type of knowledge created mirrors the joint vector of knowledge inputs to innovation. When innovation is thus separable, firms look for complementary partners, and their joint knowledge vector will be relatively flat. On average, this flattens firms' own stocks as the alliance-innovation

¹⁴ There is no explicit value from transitivity in the model, so we should not expect too much clustering in general.

¹⁵ Firms move close to their partners in knowledge space, and this "close to" relationship is transitive. So as a firm is close to its partners, it will also be (slightly less) close to its partners' partners. In that case the search for partners who are different from oneself will prevent the creation of closed triangles, and cliquishness will be small.

process proceeds.¹⁶ When experience with partners influences the probability of success, the decline of specialization with θ is less marked. Partnerships are more persistent because the experience of collaboration creates additional value by lowering the risk of failure, and thus an ongoing partnership can develop its own area of expertise. A similar effect exists at the population level. When innovation separability is strong, firms have many partners. Convergence following a successful innovation drives a firm from one to another partner, since complementary rather than similar partners are desirable. This partner churning, and period by period convergence between existing partners, has the effect of generating a population of homogeneous firms. Care must be taken in this interpretation, however, since even if there is homogenization at the aggregate level, it remains that at the level of the individual firm, similarity between a firm and its partners falls as task separability increases.

Our second interest lies in the emergent networks themselves. Given the recent literature on the subject, a natural question to ask is whether we observe small worlds: a collection of densely connected sub-groups joined by clique-spanning ties? (On that issue, see for instance Baum et al. 2003 and the references therein.) First, we find that there is a critical frontier which partitions the parameter space into two regions, one of in which agents always innovate in isolation, and one in which networking takes place. This critical frontier $\theta^*(\beta)$ is increasing, and exists both in the presence and absence of learning about partners. Above the alliance-autarchy frontier, the number of distinct partners a firm has over time increases with θ , and except when the number of partners gets very high, firms with more partners tend to accumulate more knowledge over the history of the economy.

As the innovation task is easier to separate (as θ increases) firms seek partners different from themselves, and as a consequence have more of them. In a fairly natural way then characteristic path length falls and average clustering rises. However we find more clustering and a higher characteristic path length than exists in a comparable random graph. This suggests that our networks have more structure. We can be more precise. As we move away from the frontier of autarchic innovation, while local structure remains in the form of excess clustering, the networks come more and more to resemble random graphs. Below the frontier, agents spend their histories in isolation, so the network cannot really be said to exist.

However, in the region around the autarchy-alliance frontier, more structure is present. There, what we observe are densely connected subgroups, with sparse connections between them (sometimes too sparse to connect the entire graph, but dense enough to create connected components). Thus small worlds do emerge from the processes we have modelled. In the narrow region where they are present, 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.

Two properties of innovation are central in this model: the ability of a firm to compensate for its weakness with its own strengths, which is driven by the nature of substitution in knowledge production; and the way in which firms are able to combine their knowledge stocks, which is driven by separability in the innovation

¹⁶In Cowan et al. (2003) a situation in which firms were innovating most often in their own domain of expertise was analysed, and there specialization increased as task division got easier.

process. What we observe is a strong similarity between the effects of these properties on knowledge variables and on network variables. This suggests very strongly that the two processes, of knowledge generation and network evolution, are strongly linked. It would, in general, be a mistake to try to understand one without examining the other. In addition, we observed a strong effect of θ , the parameter capturing issues of how firms pool their knowledge, and how the innovative process works. Thus to understand network formation it is necessary to know the details of the processes by which firms jointly innovate. For reasons of simplicity and parsimony, the model presented has been very stylized in its treatment of knowledge and innovation. Nevertheless, it makes the point that the interplay between the decomposition of tasks in knowledge creation, collaborative learning through repeated interaction, and the general properties of innovation as knowledge recombination are central in understanding how innovation networks, and industry structures more generally, emerge and evolve.

Acknowledgement Jonard and Zimmermann acknowledge support from the French Ministry of Research (ACI Systèmes complexes en Sciences Humaines et Sociales, SCo0061).

References

- Baum JAC, Shipilov AV, Rowley TJ (2003) Where do small worlds come from? *Ind Corp Change* 12:697–725
- Coriat B, Guennif S (1998) Self-interest, trust and institutions. In: Lazarić N, Lorenz E (eds) *Trust and economic learning*. Edward Elgar, London
- Cowan R, Jonard N, Zimmermann J-B (2003) La dynamique conjointe de la connaissance et des réseaux. *Rev Econ Ind* 103:253–274
- Cowan R, Jonard N, Zimmermann J-B (2004a) On the creation of networks and knowledge. In: Gallegati M, Kirman A, Marsili M (eds) *The complex dynamics of economic interaction. Lecture notes in economics and mathematical systems*. Springer, Berlin Heidelberg New York, p 531
- Cowan R, Jonard N, Zimmermann J-B (2004b) Networks as emergent structures from bilateral collaboration. MERIT Research Memorandum, RM2004-017
- Dodgson M (1996) Learning, trust and inter-firm technological linkages: some theoretical associations. In: Coombs R, Richards A, Saviotti PP, Walsh V (eds) *Technological collaboration, the dynamics of cooperation in industrial innovation*. Edward Elgar, London
- Dutta B, Jackson M (2003) On the formation of networks and groups. In: Dutta B, Jackson M (eds) *Networks and groups: models of strategic formation*. Springer, Berlin Heidelberg New York
- Dyer JH, Nobeoka K (2000) Creating and managing a high performance knowledge sharing network: the Toyota case. *Strateg Manage J* 21:345–367
- Erdős P, Rényi A (1960) On the evolution of random graphs. *Publ Math Inst Hungarian Acad Sci* 5:17–61
- Gale D, Shapley L (1962) College admissions and the stability of marriage. *Am Math Mon* 69:9–15
- Galison P (1999) Trading zone: coordinating action and belief. In: Biagioli M (ed) *The science studies reader*, Routledge, New York, pp 137–160
- Garcia-Pont C, Nohria N (2002) Local versus global mimetism: the dynamics of alliance formation in the automobile industry. *Strateg Manage J* 23:307–321
- Goyal S, Joshi S (2003) Networks of collaboration in oligopoly. *Games Econom Behav* 43:57–85
- Grandstand O (1996) Towards a theory of the technology-based firm. *Res Policy* 27:465–489
- Grandstand O, Sjolander S (1990) Managing innovation in multi-technology corporations. *Res Policy* 19:35–60
- Gulati R, Gargiulo M (1999) Where do inter-organizational networks come from? *Am J Sociol* 40:619–652

- Kirman A, Vriend N (2001) Evolving market structure: an ACE model of price dispersion and loyalty. *J Econ Dyn Control* 25:459–502
- Klos T, Nootboom B (2001) Agent-based computational transaction cost economics. *J Econ Dyn Control* 25:503–526
- Mowery DC, Oxley JE, Silverman BS (1998) Technological overlap and interfirm cooperation: implications for the resource-based view of the firm. *Res Policy* 27:507–523
- Narula R (1999) Explaining strategic R&D alliances by European firms. *J Common Mark Stud* 37:711–723
- Nootboom B (2000) Learning and innovation in organizations and economics. Oxford University Press, Oxford
- Nootboom B (2004) Learning and governance in inter-firm relations. *Rev Econ Polit Janvier-Février*: 114
- Oliver C (1990) Determinants of inter-organizational relationships: integration and future direction. *Acad Manage Rev* 15:241–265
- Peretto P, Smulders S (2002) Technological distance, growth and scale effects. *Econ J* 112:603–624
- Plouraboué F, Steyer A, Zimmermann JB (1998) Learning induced criticality in consumer? Adoption pattern: a neural network approach. *Econ Innov New Technol* 6:73–90
- Powell WW (1990) Neither markets nor hierarchies: network forms of organization. *Res Organ Behav* 12:395–336
- Powell WW, Koput KW, Smith-Doerr L (1996) Inter-organizational collaboration and the locus of innovation: networks of learning in biotechnology. *Adm Sci Q* 41:116–145
- Powell WW, White DR, Koput KW, Owen-Smith J (2005) Network dynamics and field evolution: the growth of inter-organizational collaboration in the life sciences. *Am J Sociol* 110:1132–1205
- Roijakkens 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
- Sako M (1991) The role of trust in Japanese buyer-supplier relationships. *Ric Econ* 45:449–474
- Teece D, Pisano G (1989) Collaborative arrangements and global technology strategy: some evidence from the telecommunications equipment industry. In: Burgelman RA, Rosenbloom RS (eds) *Research on technological innovation, management and policy*, vol 4. JAI, Greenwich, CT, pp 227–256
- van Alstyne W, Brynjolfsson E (1997) Electronic communities: global village or cyberbalkans?. Working Paper, MIT Sloan School
- Weisbuch G, Kirman A, Herreiner D (2000) Market organization and trading relationships. *Econ J* 110:411–436
- Weitzman M (1998) Recombinant growth. *Q J Econ* 113:331–360