

Center



Discussion Paper

No. 2005–44

INNOVATION, LEARNING AND CLUSTER DYNAMICS

By Bart Nooteboom

April 2004

ISSN 0924-7815

Innovation, learning and cluster dynamics

Bart Nooteboom

Tilburg University

b.nootboom@uvt.nl

submitted April 2004

Abstract

This chapter offers a theory and an analytical framework for the analysis of cluster dynamics, i.e. the innovative performance and evolution of clusters. It develops three types of embedding: institutional embedding, structural embedding (network structure), and relational embedding (type and strength of ties). The analysis is conducted from a perspective of both competence (learning) arising from relations and governance of relational risk, which includes risk of lock-in and risk of spillover. A basic proposition is that innovative clusters face the challenge of combining exploration and exploitation. Hypotheses are specified concerning differences between networks for exploration and exploitation, and concerning combinations and transitions between them. Arguments are presented that in some important respects go against the thesis of the 'strength of weak ties'. Some empirical evidence is presented from recent studies.

Key words: innovation, organizational learning, clusters, industrial districts, networks

Areas of study: geography, regional science, economics, business, industrial organization, organization

Introduction

This chapter focuses on cluster dynamics, in two meanings: the evolution or development of clusters, and their contribution to innovation. There are three reasons for this. First, although much has been written about clusters, we still know little, theoretically and empirically, about how clusters develop and evolve. Second, while it is often claimed that clusters yield innovation, it has not been made clear how that works, in terms of firms acting in clusters, and it is not clear that clusters always do contribute to innovation. Indeed, this chapter will argue that sometimes they inhibit rather than promote innovation. Third, shifts of activity due to globalization demand attention to the effects on the location and structure of clusters. Recent studies show that the famous Italian industrial districts are becoming locally dis-embedded, and shift some activities, especially in production, to emerging, proximate, lower wage countries such as Rumania. These issues have a certain urgency from the perspective of public policy. Policy makers have caught on to the fashion of cluster thinking, or perhaps we should say ideology, and contemplate active cluster policy to promote innovation. If the 'buzz' of clusters is based on myth more than facts, this can have serious adverse effects. Some studies have broached issues of cluster development under globalization and innovation (Asheim and Isaksen 2002, Boschma and Lambooy 2002, Oinas and Malecki 2002, Zuchella 2003). The purpose of the present chapter is to extend that work.

It is of dubious validity to claim that regional features have a direct effect on innovation. The challenge is to show how they affect the activities of firms, in their conduct of innovative activities, in processes of invention, innovation, spillover, division of labour, allocation of resources and collaboration. For this purpose, the notion that firms are locally 'embedded' is

in need of clarification, at least partly in terms of linkages between institutions and firms, and between firms among each other. I do not exclude the possibility that some regional features impact directly on firm behaviour, without their being traceable to ties between organizations. Indeed, I will argue that there is something like regional culture that affects firm behaviour fairly directly, but we have to be explicit in showing how it does so. Much embedding, on the other hand, entails ties between organizations, which also have to be made explicit, at least in part, to explain regional effects. Here, good use can be made of the sociological and management literature on networks of firms.¹

While local embedding may contribute to innovation, and it remains to be seen how exactly that works, in terms of the operations of firms, an escape from local embedding may also be needed for innovation. Oinas and Malecki (2002) proposed that in the study of regional systems of innovation we should recognize the need for linkages outside a region, and it may be better to speak of ‘spatial systems of innovation’. Embedding, in the sense of linkages between activities, need not always be tied to location, and may also occur in ‘communities’ that are to some extent virtual, with communication at a distance. This issue of local and other types of embedding is a central theme of the present chapter.

It will be argued that processes of innovation and learning have different stages, with different characteristics. In particular, in early development there may be a relatively greater need for local embedding, while later development requires disembedding, as suggested by Asheim and Isaksen (2002).

The chapter proceeds as follows. First, it briefly considers the definition of clusters and innovation networks, and proposes three kinds of embedding: *institutional embedding*, *structural embedding* (structure of ties between firms) and *relational embedding* (type and strength of ties). Second, for an analysis of innovation, insights are derived from the literature on innovation and learning, in particular the notion of *cognitive distance* (Nooteboom 1999) and the distinction between *exploitation* and *exploration* (March 1991). Exploitation refers to the efficient employment of current assets, including intangible assets such as capabilities, while exploration refers to the development of new capabilities. Innovation typically starts with exploration and then moves on to exploitation. A crucial problem is how one next moves out again into exploration. Here, clusters may fail and may get locked into exploitation. Third, next to issues of competence, in innovation, inter-firm relations also entail issues of governance, i.e. the management of relational risk. Effects of localization may lie more in governance than in competence. Fourth, an analysis is given of the requirements for exploitation and exploration networks, from a perspective of both competence and governance, concerning the three types of embedding: institutions, network structure, and strength of ties. Some empirical illustrations are provided. Finally, conclusions are drawn concerning the development of clusters and the implications for cluster policy.

Clusters, networks and embedding

Cooke and Huggins (2002) defined clusters as follows: ‘Geographically proximate firms in vertical and horizontal relationships involving a localized enterprise support infrastructure with shared developmental vision for business growth, based on competition and cooperation in a specific market field’. Apparently, clusters and industrial districts are more or less synonymous. The concept of a network is more general, and does not necessarily entail local embedding, a shared objective, or a specific market. Thus, a cluster is a network but not necessarily vice versa. This chapter focuses on clusters in the context of innovation.

¹ E.g. Granovetter 1973, Coleman 1988, Burt 1992, Krackhardt 1999, Nooteboom & Gilsing 2003.

Presumably, the notions of ‘innovation clusters’ and ‘regional systems of innovation’ are close synonyms.

As noted above, Oinas and Malecki (2002) proposed that in the study of regional systems of innovation we should recognize the need for linkages outside a region, and it may be better to speak of ‘spatial systems of innovation’. These considerations raise the question whether local embedding and geographical proximity should be retained as defining characteristics of clusters. If they are, we may have to say that in their development clusters are transformed into other types of networks. To avoid this definitional issue, I will at times speak of innovative networks, rather than clusters, which may or may not be strongly locally embedded.

As indicated in the introduction, the notion of ‘embedding’ requires specification. If we are to explain the causality of regional characteristics, it seems that we have to make explicit how regional level variables affect firms, and this is likely to operate in large measure through ties between firms and other organizations, such as schools, universities, intermediaries of many kinds, bodies of public administration, and between firms among each other, in relations of supply, demand, alliances and other forms of collaboration. Here, use can be made of the sociological and business network literature.

I propose three kinds of embedding: *institutional embedding*, *structural embedding* and *relational embedding*. Institutional embedding regards the impact of regulation and norms of conduct, taxes, subsidies, legal system, infrastructure, schooling, research, labour market, etc. Structural embedding derives from the social network literature. Structural features of networks are size (number of participants), density (actual number of direct ties as a ratio of the maximum possible number), centrality (of which there are several forms), and stability of structure (rate of entry and exit). Relational embedding appears in the social network literature in the notion of the ‘strength of ties’, but is developed in more detail in the literature on alliances or inter-organizational relations (IOR’s). In other words, I propose that an adequate understanding of clusters requires a combination of geography, social networks and inter-organizational relations.

Innovation and cognitive distance

For innovation, diversity is a crucial condition, to produce Schumpeterian ‘novel combinations’, as demonstrated, in particular, in evolutionary economics (Nelson and Winter 1982). Diversity is associated with the number of agents (people, firms) who are involved in a process of learning or innovation by interaction. Next to the number of agents involved, a second dimension of diversity is the degree to which their knowledge and skills are different. This entails the notion of *cognitive distance* (Nooteboom 1992, 1999). Note that here cognition is seen in a broad sense, including not only rational evaluation but also emotion-laden value judgements, and mental heuristics of attribution, inference and decision making. The notion of cognitive distance derives from a social constructivist view of knowledge, according to which perception, interpretation, understanding and value judgment entail mental constructions on the basis of mental categories that are developed in interaction with the physical and social world. As a result, different people, and different organizations, that have developed their cognition along different paths of development, in different conditions, will perceive, interpret and evaluate the world differently.²

A central task of organizations is to sufficiently reduce cognitive distance, in an organizational *focus*, including epistemic as well as moral categories, to enable the achievement of joint purpose. Note that such cognitive categories serve for guiding cognition

² For an elaboration, see Nooteboom (2000).

in the narrow, epistemic sense of attention, perception, and interpretation, but also for setting behavioural values, in a moral order, to facilitate collaboration, constrain opportunism, build trust, and limit and resolve conflicts of interest (governance). Such categories tend to be internalized, to a greater or lesser extent, by people, as part of tacit knowledge, assimilated in socialization and habituation.

Existing cognitive structures constitute *absorptive capacity*. On the level of organizations, this was recognized by Cohen and Levinthal (1990). Here, absorptive capacity includes organizational capabilities to assimilate information, internally distribute it, and implement knowledge in design, development, production and marketing. It depends, among other things, on R&D. The notion of absorptive capacity is crucial in the analysis of spillover, which plays a central part in the analysis of clusters and regional innovation systems. Spillover depends on absorptive capacity.

Absorptive capacity, connected with organizational focus, not only enables but also constrains organizational cognition, yielding organizational myopia, which needs to be compensated by engaging in outside relations with other organizations, with different, complementary foci, at some cognitive distance. This yields a new purpose for inter-organizational alliances, next to the usual considerations, known from the alliance literature.³ Firms need to make a trade-off between organizational identity in the form of a clear focus and wide scope of internal competencies. A wide scope, with a wide focus, entails limited identity. A limited scope, with a narrow focus and stronger identity, can be compensated by alliances.

The notion of organization as a focusing device is relevant for two reasons. First, it is needed to understand the functioning of firms in networks. Second, it may have implications for the notion and the role of clusters. Perhaps a defining characteristic of a cluster is that it, also, entails a shared culture with corresponding cognitive focus, in an epistemic and moral order. This is certainly facilitated by local embedding and geographical proximity, but those may not be necessary. Perhaps there is a viable and fruitful combination of geographical distance complemented by frequent meetings to build and maintain the shared focus.

In processes of learning and innovation, in interaction between firms, cognitive distance, now in terms of difference in organizational focus, yields both an opportunity and a problem. The opportunity lies in diversity: the novelty value of a relation increases with cognitive distance. However, mutual understanding (absorptive capacity) decreases with cognitive distance. If learning performance from interaction is the mathematical product of novelty value and understandability, the result is an inverse-U shaped relation with cognitive distance. Optimal cognitive distance lies at the maximum of the curve.⁴

The analysis of optimal cognitive distance has several implications for cluster dynamics. One is that firms should seek optimal distance for innovation. Another implication concerns the duration of relationships. Between firms, cognitive distance may be reduced to the extent that they have engaged in continued interaction, especially when that interaction was exclusive. In other words, their foci start to overlap, in a shared epistemological and normative framework. This reduces the novelty value of a partner's cognition, with a reduction of innovation performance. This suggests that while familiarity breeds trust (Gulati 1995), it may also reduce learning potential, so that for the purpose of learning ties should not be too strong in terms of duration. In sum, next to optimal cognitive distance there is also something like an optimal duration of ties for learning: long enough to build mutual understanding and trust, but not so long as to run out of innovative steam.⁵ The point has important implications for cluster dynamics, and for any cluster policy. Too durable, local

³ For an elaboration, see Nooteboom (1999, 2004)

⁴ For an empirical test of optimal cognitive distance, see Nooteboom et al. (2005).

⁵ For an empirical test of optimal duration, see Wuyts et al. (2005).

embedding, particularly when it is cut off from outside contacts, may reduce cognitive distance too much. It may be good for trust but bad for learning. As will be argued in more detail later, this yields one way in which clusters may inhibit rather than promote innovation.

Note that the claim of a negative effect of duration of a relation was qualified by the condition that the relation is exclusive, i.e. the firms involved have no contacts outside the relation, pertaining to the subject of collaboration. When, on the other hand, both sides of a relationship tap into outside, non-overlapping networks, they may be continually re-charged with novel insights that keep their relation vibrant. This connects with Burt's (1992) notion of the advantages in bridging 'structural holes': the relation just described provides such a bridge. This illustrates why clusters, to the extent that they entail stable relations, need outside ties for ongoing innovation.

In sum, regional variables that are relevant for innovation, in cluster dynamics, include variables that affect absorptive capacity of firms, such as educational facilities, R&D in firms, R&D in public organizations and the transfer of outcomes to firms. They also include variety and cognitive distance in the cluster, the duration of linkages and outside linkages that replenish variety.

Exploration and exploitation

The economic success of regions and clusters requires success in both exploitation and exploration (March 1991). Exploitation, i.e. the efficient employment of current assets and capabilities, is needed to survive in the short term. Exploration, i.e. the development of novel capabilities, is needed to survive in the long term. Thus, to survive in the short and long term, firms must combine the two, somehow. That is a paradoxical task. Exploitation often requires the maintenance of a stable organizational structure, in division of labour, with unambiguous terms and clear standards, in a narrow organizational focus, while exploration requires the reverse: loosening of structure for novel reconfigurations, shifting meanings and deviation from existing standards, in a wide focus.

A key problem is how exploration may be based on experience in exploitation, and how to ensure that the outcome of exploration will be exploitable. How do exploitation and exploration build on each other? What path of development can we think of that maintains exploitation while at the same time yielding exploration?

Here, use is made of the 'cycle of discovery' proposed by Nooteboom (2000). It plays a key role in this chapter, in yielding a model for cluster development. In particular, this model will show how clusters can both support and inhibit innovation, in different stages of the cycle of development.

The cycle of development proposes several stages of development, in which there is an alternation of variety of content and variety of context. First, variety of content (of a concept or practice) that emerges from exploration is reduced, in *consolidation* into a *dominant design*, as suggested in the innovation literature. As a result of reduced uncertainty, demand increases, and new producers jump on the bandwagon. Related industries and existing distribution channels go along, and adapt, from fear of missing the new boat. The new technology/product/market combination develops into a dominant design or 'dominant logic' (Bettis and Prahalad 1995) of organization, including network structure and 'industry recipes' (Spender 1989), with pressures to conform, in 'organizational isomorphism' (DiMaggio and Powell 1983). New entrants exert pressure on price, and for the sake of efficient production increase of scale, division of labour and associated specializations emerge.

So far, this is nothing new, from the perspective of innovation theory. However, the question next is how one gets away from the dominant designs in technology and organization, in a next round of (radical) innovation. This is a crucial point in cluster

dynamics: how does a cluster get away from the possible inertia as the outcome of previous development, in the efficient use of past innovation? Here, the proposal is that for exploration that yet maintains exploitation, one next needs to open up to a new variety of contexts of application, in *generalization*. These novel contexts of application may be sought voluntarily, in an expansion of activity. Voluntary new applications of established capabilities appear to be based on an instinctive drive that among psychologists is known as a principle of 'overconfidence'. In economics, there is also a pressure to extend the market, as growth in the original market stagnates. With the entry in new markets, one needs to adapt products and organization, which requires knowledge from outside. Clusters may need MNC's as a vehicle for this reach outside existing boundaries. This is in line with Boschma and Lambooy's (2002) analysis of developments in Italian industrial districts, where they identified the role of MNC's as 'bridging enterprises', to carry activities into international markets and to access outside sources of knowledge, and with Asheim and Isaksen's (2002) analysis of how Norwegian clusters had to make a shift from local to global operations.

However, new conditions of market, technology and institutions may also be imposed from outside one's familiar niche. An illustration of this, in the development of multi-media, is that publishers finally went along in digitalization and electronic distribution of text, invading their market, for fear of losing their position (Gilsing 2003). For clusters, this may arise from an invasion of multinationals.

A novel context is needed for three reasons. The first is that established capabilities arose and consolidated in a given niche, and therefore perform well there, and are taken for granted, so that new conditions of technology, demand, infrastructure and institutions are needed to gain new *insights in limits* of validity. The second reason is to build *insight in novel goals and motivation* for change, resulting from such misfits, in the novel context. The third reason is to yield *insight into potential novel content* of practice, for which inspiration is found in the novel context.

First, to maintain exploitation as much as possible, there is an attempt to make minor, incremental adjustments to established practice, in what is called *differentiation*. Insight for this may come from previous experience, in novel selections from familiar repertoires, which are retrieved in an attempt to improve fit in the novel context. Next, when this fails, experiments are conducted with novel elements, adopted from the novel context, which seem to be successful where familiar practice fails, in hybrids of old and new elements, in *reciprocation*. This yields an opening up to new variety of content. The function of this is two-fold. First, it still allows for ongoing exploitation, albeit in new forms. Second, it allows for experimentation with new elements, to test their potential, without sacrificing existing basic design principles.

Next, when such potential emerges, there is more willingness to make more radical changes in architecture, when that is needed for the novelty to realize its full potential, in *accommodation*. Here, rigidities of established structures, which may have offered an advantage for exploitation, become a liability. Emerging novelties cannot achieve their potential under the systemic limitations imposed by existing structures, practices and ways of thinking. If the cluster or network is unable to cope with this, it may need to be broken up, so that different elements have more scope to adapt, in different ways, to new conditions. Accommodation, then, leads to a new beginning, under radical uncertainty, in search of novel dominant designs, in consolidation, and we are back at the beginning of the cycle.

In sum, in the efficient exploitation of previous innovations, clusters are in danger of getting stuck in inertia. To escape from this, they need to step outside, or they must allow and indeed invite entry into the cluster, to generate novel insights into limits of current practice, new needs and opportunities, even if, or precisely because, this may lead to a break-up of the cluster, in a next radical innovation.

Effects of embedding

The opportunities and problems for combining exploitation and exploration depend on institutional, structural and relational embedding. The difficulty of combining exploitation and exploration, in differentiation and reciprocation, depends on three structural features of the exploitation system of the cluster, which determine its rigidity:

- a. The *complexity* of division of labour, defined as the *number* of component activities and the *density* of direct ties of dependence between them. Structure is *simple* when complexity is low.
- b. The *modularity* of the system, on the basis of clear and stable constraints on activities, along such ties of dependence, in the form of *standards*, needed to maintain systemic integrity. The opposite would be ambiguity and variability of constraints, by which activities need to be continually coordinated.
- c. The *tightness* of constraints, i.e. the scope for variety in contributions from component activities. Structure is *loose* when tightness is low.

Exploitation is *systemic* when it has features a and b (complex and tight), and *stand-alone* in the opposite case (simple and loose). In case of feature b (modularity), component activities can be autonomous, and can be replaced, as long as they satisfy the constraints on interfaces in the position they take in the structure. If exploitation entails a systemic structure, simultaneous exploration is constrained by the many and tight constraints on component activities. Exploration would soon yield a breaking of constraints on interfaces, yielding many unknown repercussions in the dense structure of dependencies, such as change of content of linked activities, which may in turn trigger change elsewhere, possibly resulting in wide ranging architectural change (Henderson and Clark 1990). Management, or cluster policy, would rightly be wary of accepting that risk and cost, unless there were a clear and proven potential of the novelty that would justify them.

One possibility is to see to it that exploitation structure is not systemic but modular or stand-alone, even if this entails loss of efficiency, to maintain options for exploration by reconfiguration. Here, it matters what options for reconfiguration are at hand. Here, perhaps, we encounter the notion of ‘Jacobs externalities’ (Boschma and Lambooy 2002). In urban regions with a large variety of different activities, and a rich, varied, complex infrastructure, with a wide scope of spillovers, new ideas, and activities that become complementary in new ways, there is more scope for new exploration. When exploitation is irremediably systemic, there are two basic options to combine exploitation and exploration: separation in place and separation in time.

With separation in place, within a region, exploitation would take place in one part, and exploration elsewhere, in different cluster, perhaps, which is at least temporarily exempt from exploitation and may need to be protected from premature competition and spillover. This includes the classic separation between production (exploitation) and R&D (exploration). A familiar case is the pharmaceutical industry, where small biotech firms explore novel medically active substances, and when those are found and ripe for exploitation, large pharmaceutical firms take over for the lengthy process of clinical testing, and efficient large-scale production and distribution. A problem arises in ensuring that exploration is based on inspiration from exploitation and that exploration is motivated and able to apply results of exploration. This may be achieved by mobility of firms or people between the two sectors. For example, researchers in R&D organizations are encouraged to try and apply their

inventions in firms, and workers in firms are encouraged to reflect on their experience and new alternatives in an R&D environment.

With separation in time, a region's structure fluctuates between integration and disintegration.

Exploration often requires disintegration: new elements that do not fit in existing structures (of production, supply, market, distribution channels, institutions) need to shield themselves off in a niche where deviation from established structure and process is feasible. This often requires the emergence of new firms that are not imprisoned in existing structures and interests. This may entail entry from outside, or spin-offs from existing firms or networks, of entrepreneurs who escape from organizational or network prisons.

Alternatively, a region or cluster may specialize in either exploitation or exploration, and seek the other in relations with other regions with complementary specialization. This is a form of separation in place, but between rather than within regions. Here, outside linkages are especially important. In a region that specializes in exploitation, when exploitation is highly systemic, an option is to encourage entrepreneurs that cannot find the leeway to innovate within the region to 'spin off' into a region that is more oriented at exploration, and come back when results of exploration are ready for exploitation.

Risks of lock-in and spillover

Inter-organizational relations, and networks, have a competence and a governance side (Nooteboom 2004b). Competence includes, in particular, innovative competence, i.e. the ability to generate, efficiently employ and diffuse novelty, in technology, products, production, organization, distribution and other aspects of marketing. This was discussed in the previous sections. Governance refers to relational risks and ways to manage them. Generally, in the literature, the focus has been on competence, to the neglect of governance. It is important to include governance, because regional effects may lie at least as much, and perhaps more, in the management of relational risk than in the development and support of competence.

For the analysis of governance, here the focus is on risks of lock-in and spillover. In relational embedding, risk of lock-in includes the 'hold-up' risk from transaction cost economics (TCE). That is relevant here for several reasons, one of which is that it has implications for the strength of a tie, in terms of its duration and frequency of interaction. Hold-up risk results from dependence as a result of relation-specific investments, defined as investments that have value only (or largely) in a specific relation. Specific investments yield switching costs: when the relation breaks, the investments have to be made anew in a new relation. For such investments to be made, the relation should be expected to last sufficiently long, and be sufficiently intensive, in frequent interaction, to recoup specific investments. Lock-in yields a temptation for the partner to expropriate value, in opportunistic behaviour.

Specific investments include the usual types, offered by TCE: location specificity of facilities, physical asset specificity (installations, tools, instruments), human asset specificity (training), dedicated capacity, brand name specificity, and time specificity. The specificity of investments depends on the flexibility of technology: the more an investment allows for a variety of products or production processes, the less specific it is. It also depends on the availability of shared standards: the more different firms share standards, the less specific investments are. Here, in particular, lies a task for public bodies, in the provision of standards, and in the stimulation and perhaps in the policing of their use.

An example of specific investments that is of particular relevance in the present context of cluster dynamics is associated with the earlier discussion of optimal cognitive distance. One can increase mutual understanding, and thereby increase optimal cognitive distance and raise

innovative performance, but this entails an investment that may be largely or partially specific to a relation, so that by the logic of transaction cost theory the relation would need to last sufficiently long to make that investment worth while. Note that now we have the elements of an important trade-off in cluster dynamics. Here we find that a certain stability of relations may be needed to promote mutual understanding needed to utilize opportunities form cognitive distance. In the previous section we found that when relations become too durable, they may run out of innovative steam, when those relations are exclusive.

Investment in relation-specific trust may also constitute a relation-specific investment, in the sense of TCE. This is important especially in the present context of cluster dynamics, since under the uncertainty of innovation, contracts, as a means of governance, are difficult to specify. Hence, for governance one must fall back on other instruments of governance, such as reputation and relation-specific trust.⁶

Lock-in may also arise from structural embedding. Here one is locked into a community by constraining coalitions of members of the community. There might also be network-specific investments, which can be used within but not outside the network. Some of these may indeed be related to the network level (structural embeddedness) rather than to the level of individual relations (relational embeddedness). Network-specific investments may also arise in finding out ‘who is who’ in the network, and in getting embedded in local reputation systems.

Lock-in may also arise from institutional embedding. Here, one is locked into the regional ‘focus’ of shared understandings and moral precepts or customs, local obligations of loyalty and conformity, and lack of cognitive distance, in ‘group think’.

Now I turn to spillover risk. Linkages with other actors (firms, other organizations, individuals) yield access to variety of knowledge, and this is the positive side of spillover, emphasized in cluster studies. However, firms may also see spillover as a risk that knowledge that is part of one’s ‘core competence’, which constitutes competitive advantage, may be used in competition, either by a direct contact (relational embedding), or indirectly, elsewhere in the network, through a sequence of direct contacts (structural embedding). Note that the assessment of spillover risk requires a trade-off between knowledge adopted by others and knowledge gained from them. The risk is potentially serious only when there is a net loss rather than gain.

In relational embedding, spillover risk depends on how tacit or documented knowledge is, with the latter spilling over more easily than the former. It also depends on the absorptive capacity of potential competitors, i.e. their ability to effectively understand and implement knowledge spillovers. That depends on the ‘cognitive distance’ between actors, i.e. differences in their ability to perceive, understand and evaluate relevant phenomena. Finally, spillover risk depends on the speed with which knowledge changes: if it is obsolete by the time it has spilled over and has been absorbed and imitated by potential competitors, spillover risk drops out.⁷ In structural embedding, spillover risk depends on density of the network, and the centrality of one’s position.

In sum, regional variables that are relevant for relational risk include the type of industry and technology, in particular their implications for specificity of investments, tacitness of the knowledge involved, and the speed at which that knowledge changes. They also include the availability and use of technical standards, and the density of networks. They also include the features that affect absorptive capacity, indicated before.

Instruments of governance

⁶ For a more detailed analysis of instruments for governance, see Nooteboom (1999, 2004).

⁷ Ibid.

Relational risks require governance to limit them and to create trust. Counter to TCE (Williamson 1993), I hold that trust can go beyond calculative self-interest, in loyalty and *benevolence*, and yet be viable in markets, although I acknowledge that such trust should not be unconditional, and is subject to limits.⁸ In the notion of trust, we need to distinguish between ‘competence trust’, in the ability of people and firms to satisfy expectations, and ‘intentional trust’, in the commitment of people to perform to the best of their abilities, and not to engage in opportunistic behaviour. A survey of instruments for the governance of intentional risk, as ‘sources of collaboration’, is given in Table 1. Here, a distinction is made between macro and micro, and between self-interested and other-directed sources of collaboration.

 Table 1 about here

The distinction between macro and micro sources of collaboration, in Table 1, is also known as the distinction between ‘universalistic’ or ‘generalized’ sources versus ‘particularistic’ sources, made by Deutsch (1973: 55), and between impersonal and personalized sources made by Shapiro (1987). The first arise partly from institutional embedding, in laws, norms, values, standards, and agencies for their enforcement, and partly in the structural embedding of relations. The former yields ‘institution-based trust’. This kind of trust requires that we trust those institutions to support trustworthiness of people and organizations. Structural embedding includes opportunities for coalitions and reputation mechanisms. The ‘micro’ sources arise in specific relations, in relational embedding, and are often personalized.

The table further distinguishes between self-interested and altruistic or ‘other-directed’ sources of co-operation. The self-interested sources are associated with the notions of deterrence and ‘calculus-based trust’ (McAllister 1995, Lewicki and Bunker 1996). In the present reconstruction, this includes *opportunity control* and *incentive control* (Nooteboom 1999). Opportunity control entails that the space of feasible action is constrained. Incentive control affects the choice of opportunities, in the space of feasible actions. Within organizations, opportunity control entails control by hierarchy, and in inter-firm relations it entails control by contract. Contracts are only useful to the extent that one is able to adequately specify them and monitor conformance to them. Even under the best of institutional conditions, legal ordering cannot be closed, including all relevant future contingencies. This is problematic especially in innovation, with its unknowable future contingencies of contract execution. Also, in innovation knowledge is sometimes highly tacit, which would also inhibit the specification of contracts.

In incentive control, partner B behaves well towards A because he is dependent on A for one or more of the following reasons: A has a unique, difficult to replace value to B, B faces switching costs as a result of relation-specific investments, partner A holds a hostage from B, or B has to protect his reputation. Reputation mechanisms depend on the structure of the network, in particular network density, and on the presence of intermediaries who take up positions of centrality and act as selectors of gossip (to test legitimacy of complaints), and as amplifiers and broadcasters. This is especially important when contracts are not feasible, as in innovation. Intermediaries have several other roles to play in the governance of inter-firm relations, such as to aid in mutual understanding, guarding spillover, providing intermediation or arbitration in conflicts, building relation-specific trust and helping to end relations with a minimum of conflict.⁹ The latter is important in view of the need to maintain sufficient flexibility of relations for the sake of innovation, as discussed before. The notion of hostage is

⁸ For a detailed discussion of the types, sources and limits of trust, see Nooteboom (2002).

⁹ For further discussion, see Nooteboom (2002), and, in the context of learning regions, Nooteboom (2003).

also taken from TCE. In business, hostages often take the form of information or knowledge that is sensitive, in the sense that it could cause great damage when leaked to competitors. It can also take the form of cross-participation, or the borrowing of staff, with the threat of poaching them.

Now we turn to the other-directed sources of collaboration, in trust that goes beyond calculative self-interest. On the macro level, they lie in established, socially inculcated norms and values (macro). They include pressures of allegiance to groups one belongs to, and values and norms inculcated by socialization into those groups. Of course, one can never be sure *ex ante* to what extent a stranger without reputation has actually internalized such norms and values.

On the micro level of specific relationships, trust may be based on empathy. This entails that one knows and understands how partners think and feel. It is connected with mutual openness, and acceptance of control by others, which are crucial for the build-up of trust (Zand 1972). While trust can go beyond calculative self-interest, it has, and should have, limits. Empathy allows one to assess strengths and weaknesses in competence and intentions, to determine limits of trustworthiness under different conditions. Identification-based trust goes further: it entails that people think and feel in the same way, sharing views of the world and norms of behaviour. This may lead to affect- and friendship-based trust. Routine-based trust, proposed by Nooteboom (1999), entails that when a relation has been satisfactory for a while, awareness of opportunities of opportunism, for oneself and for the partner, is relegated to ‘subsidiary awareness’ (Polanyi 1962). One takes the relation for granted and does not continuously think about opportunities to gain extra advantage. As Herbert Simon has taught us long ago, routinized behaviour is rational in view of bounded rationality, since it allows us to focus our limited capacity for attention and rational evaluation on matters that are new and have priority. Routines are rational also in the sense that they are based on proven success in past behaviour. On the other hand, their lack of awareness creates the problem that they may no longer be adequate when conditions change. However, when results or perceived events exceed certain tolerance levels, triggered by emotions routines are often summoned back from subsidiary into focal awareness, to be subjected to rational scrutiny. Empathy, identification and routinization may be enhanced by joint membership of clubs, such as sports clubs, community centres, Rotary club, etc.

In sum, regional variables that are relevant for instruments of governance include legal systems, norms and values of conduct, as part of regional culture, intermediaries that support reputation mechanisms, arbitration and intermediation, building of relation-specific trust and the ending of relations (institutional embedding). They also include the effect that network structure has on reputation mechanisms (structural embedding), and whatever social conditions that may affect the building of relation-specific trust (relational embedding).

Synthesis: structure and strength of ties

Now I turn to details of structural and relational embedding, integrating the analytical results from previous sections. What features of embedding are relevant, from the perspective of both competence and governance, for an analysis of differences between networks for exploitation and exploration?

First, consider the structure of ties in a network, in *structural embedding*. From a perspective of competence, recall that cognitive diversity has two dimensions: the number of actors and ties between them, and cognitive distance in the ties. The first is determined by the *size* and *density* of the network. As analysed in a previous section, for innovation large size and density allow for more access to different sources of information, but by the same token also increase possibilities of spillover. A third feature of network structure is network

stability, which has implications for how variety develops in time, in entry and exit of new members. As analysed in a previous section, this is important to maintain variety for the sake of exploration. High stability may be good for mutual trust and efficiency of exploitation, but bad for exploration. Another well-known feature of structure is *centrality*, of which there are several types. Here, I focus on degree centrality, which is the degree to which some nodes have more direct ties than other nodes do. An extreme case of centrality is a hub-and-spoke structure. Centrality may be needed for the coordination of activities. A central position yields power, but possibly also constraints on behaviour, in view of the many possibly divergent interests it is involved in (Krackhardt 1999), and in its task of centralized coordination it may suffer from information overload. In view of the latter, structure may need to be hierarchical. As noted before, from a perspective of governance, size and density affect possibilities of lock-in by coalitions, reputation mechanisms, and shared norms of ethical behaviour.

Next I turn to tie strength, in *relational embedding*. Here, seven dimensions are proposed.¹⁰ The first four arise from considerations of competence and the last three from considerations of governance.

1. *scope*, defined as the range of activities involved in the tie. Does it involve only knowledge on the location and relevance of knowledge, anywhere in the network, or also the actual exchange or joint production of new knowledge (cf. Hansen 1999)? Does it involve knowledge only on a small number of issues, or on a wider range of issues, concerning technology, markets, organization, and reputation of players in the network?
2. *investment in mutual understanding*, needed to build mutual understanding, for crossing cognitive distance. To the extent that this investment takes time and is specific, ties need to entail sufficient ...
3. *frequency* and/or
4. *duration* of interaction.

While investment, frequency and duration facilitate learning, they also facilitate spillover. As argued earlier, long duration of a tie may lead to identification, which enhances mutual understanding and trust, but may reduce learning potential, particularly if the tie is exclusive, i.e. in the areas of collaboration (in the scope of the relation) there are no direct ties with others.

From the perspective of governance, ties require instruments for the management of relational risk of lock-in and spillover, specified in Table 1. This yields the following three dimensions of tie strength:

5. *opportunity control*, by contract
6. *incentive control*, by mutual dependence, reputation or hostages.
7. *trust and mutual openness*, beyond control.

All dimensions of governance depend on *institutional embedding*. This is one important area where regional variables impact on inter-firm relationships. Contracts depend on the availability of a legal infrastructure of laws and non-corrupt judiciary and police. As indicated in Table 1, trust may be relation-specific, on the basis of empathy, identification, affect and routinization, but also taps into shared values and norms, in a given community. As analysed before, reputation depends not only on the structure (density) of the network, but also on the availability of other social groups that facilitate, filter and guide gossip, such as professional

¹⁰Adopted from Bogenrieder and Nootboom (2004) and Nootboom and Gilsing (2003).

and industry associations, clubs, and the like. Such institutions as well as personal acquaintances may also fulfil other roles of intermediaries for the building of trust, as analysed before. Trustworthiness of (potential) partners may be attributed not only on the basis of experience in transactions, but also from chance meetings and observations in a rich variety of social settings that are most readily available in local embedding.

As noted before, local embedding may be needed more for governance than for competence. This is related to the notion of optimal cognitive distance. As indicated before, cognition here is a wide notion, including mental categories concerning both morality, in ways of dealing with each other, and cognition in the narrower sense of substantive knowledge and skills, concerning technology, markets, etc. For governance one may need more proximity in morality and for competence one may need more distance in substantive knowledge. This effect is reinforced by the fact that morality is often more tacit, and requires more face-to-face interaction, gossip and chance interactions, in local ‘buzz’, than substantive knowledge does. This important point is elaborated below.

Networks for exploration and exploitation¹¹

Differences between networks for exploration and networks for exploitation may now be specified in terms of the features of embedding set out above. First I turn to network structure and subsequently to the strength of ties.

In networks for exploration, there is uncertainty concerning future dominant designs, in both technology and organization, which yields structural uncertainty concerning the configuration of future networks for exploitation. One needs access to actors who might offer complementary knowledge, but one does not know what elements of knowledge will turn out to be relevant when a dominant design develops. Also, one does not know what actors will survive by that time. Therefore, the network has to be *dense*. Later, it will be argued that dense structure is also needed for a reputation mechanism. Here, we start to diverge from the thesis of the ‘strength of weak ties’ proposed by Granovetter (1973) and Burt (1992), according to which structure should not be dense, and ties should be weak. They assumed, implicitly, that one knows:

1. What knowledge will be relevant
2. Who has what knowledge
3. Who will survive to provide direct or indirect knowledge

And that:

4. One is able to absorb that knowledge

In exploration, however, one does not yet have such knowledge and absorptive capacity, and therefore one has to hedge relational bets. One does not yet know what ties will turn out to be redundant, since one does not know who will develop what knowledge and what the configuration of relevant elements of knowledge will be. One has to maintain direct linkages even if they may later turn out to be redundant, to keep options of access open, covering for the risk that some ties will drop out and thereby eliminate indirect access to other sources. Even if a tie is already known to be redundant for access to a known source of knowledge, it may be needed to assess, understand and absorb that knowledge. More precisely, if A remains linked to both B and C, even if there is also a link between B and C, this may help A to understand C by comparing what A understands from C with what B understands from C. In other words, a dense structure enables firms to ‘triangulate’ among their multiple sources and

¹¹ This section is largely based on Nooteboom & Gilsing (2003).

thus better assess their value, and to better absorb knowledge from them. This role of third parties for the sake of competence is extended by roles of intermediaries for governance, discussed before.

The argument against redundant relations, from Granovetter and Burt, was that their set-up and maintenance yield excess costs. However, relevant costs are only those of relation-specific investments in mutual understanding, since other, more generic investment would be useful also in other ties. In exploration, in contrast with exploitation, specific investments other than in mutual understanding are often limited in size, in activities such as prototyping rather than large outlays for efficient production, marketing, distribution, and servicing. Furthermore, in exploration costs are less of an issue, since competition focuses on form, in connecting complementary competencies in the fast development of prototypes, rather than on the price of a ready product, as in exploitation.

In sum, we need a careful trade-off between costs and benefits of redundancy. In exploration:

- the relation-specific costs of setting up and maintaining ties may not be high, or at least not as high as in exploitation,
- such costs may not have priority,
- and redundancy may be needed to hedge structural bets,
- bets on knowledge content,
- for triangulating knowledge content and reliability, and
- for aiding the absorption of knowledge.

To maintain the variety of cognition needed for exploration, network *stability* is expected to be generally low, allowing for entry and exit. Exploration is facilitated by volatility of interaction, allowing for chance meetings, to discover interesting potential partners. Here, local embedding may be needed for reasons of competence. Under conditions of radical innovation, with uncertainty concerning what elements will emerge and survive in what configuration, *centrality* is likely to be low, especially in stand-alone technology. Centrality might yield an obstacle, from attempts to maintain the power invested in an established centralized position.

Now I turn to the strength of ties. In exploration, uncertainty is diffuse and wide ranging, so that interaction entails many issues, including technology, organization and perhaps also future market demand, the availability of competent suppliers and so on, and as a result ties tend to be strong in the dimension of *scope*. It was already noted above that building mutual understanding might require a *relation-specific investment*, which requires sufficient *frequency of interaction* and/or *duration*, to make such investment worthwhile. However, since knowledge changes fast, in exploration, the economic life of the investment is short, so that it should be recouped in a short time, in frequent contacts, and duration, though it should be sufficient, *need* not be long. How long duration should be depends, among other things, on the size of specific investment for mutual understanding, which depends on the depth and level of specialization of knowledge, and the degree to which it is tacit. Duration *should* not be too long, for two reasons. The first reason is that it would prevent novel architectures of configurations. This is particularly relevant under systemic conditions, where innovation often takes the form of frequent and rapid architectural change. Here, one might think of the car industry, for example. The second reason is that too durable relations may yield identification that goes so far, in an excess of familiarity, as to reduce innovative potential. However, this depends on how exclusive the relation is. If A and B have a tie, on a certain subject, and both A and B also have other ties, on the same subject, to different nodes, apparently unafraid of

spill-over risks, then their mutual value as sources of knowledge may be replenished from those outside contacts, so that a long duration does not necessarily kill learning potential.

From a governance perspective, in exploration the use of *opportunity control* by contracts tends to be problematic. In exploration much knowledge is tacit, which complicates the specification of contracts. Uncertainty about contingencies, even in the very near future, may also preclude their detailed specification. Since change is rapid, the content of contracts would have to change frequently, which makes them less cost-effective. In view of new and not yet dispersed knowledge, it would be difficult to monitor and assess conformance to contracts. In exploration, governance is likely to be based on *incentive control*, with a balance of mutual dependence, hostages in the form of sensitive information, a reputation mechanism, and/or on *trust and mutual openness*. A reputation mechanism is especially strong here, in view of the uncertainty about possible future configurations of relations. Since it is impossible to assess who may and who may not in the future yield an important connection, one has to be careful in *all* relations. As analysed before, a reputation mechanism requires *density* of relations, and is facilitated by *institutional embedding*. The institutional basis for trust typically lies in professional values, norms, and standards, guarded by professional associations, which also play an important role in reputation mechanisms. Typically, in exploration trust initially is competence trust, in professional knowledge and skill, and this establishes a basis for intentional trust to develop, on the basis of pre-existing professional empathy. Here we find a second argument for *frequency* of interaction, as needed for the build-up of trust, in empathy, identification and routinization (Table 1). Such relation-specific, personalized trust entails, and requires a great deal of *mutual openness* (Zand 1972). It is known from the trust literature that trust is stimulated by mutual dependence. When one cannot do without each other, one simply has to develop trust in collaboration. As indicated, in exploration such mutual need is high, to search for complementary knowledge, in the race for a viable prototype.

In sum, the hypothesis is that in exploration ties tend to be strong in terms of scope, frequency and trust/mutual openness, of some strength in terms of relation-specific investments, depending on the complexity and tacitness of knowledge, and some duration, depending on how systemic the technology is. They are generally not strong in control, and duration should be limited, particularly when relations are exclusive.

A potential problem now is that density of the network, investment in mutual understanding, frequency of interaction, and trust and openness may yield a high risk of spillover. However, at this stage, in exploration, with large uncertainty on what dominant design will emerge, and to what products it will lead, in what markets, knowledge often is 'pre-competitive', so that spillover risk may be limited. Also, it may be difficult to assess who will in future turn out to be a potential competitor.

Restricting relations for fear of spillover would soon entail no relations at all. Finally, knowledge may change so fast as to eliminate serious spillover risk. Another potential danger is that the network becomes too tight and stable, with too durable relations between members of an in-crowd, in a tight 'clan' (Ouchi 1980), which reduces diversity in terms of both people involved and cognitive distance, and yields stagnation. To counter this, as discussed above, ties should not last too long, especially when technology is systemic and innovation is frequent and architectural, and network stability should not be large, offering a certain volatility of network membership, for the sake of novel combinations. This is where the thesis of the strength of weak ties comes into its own.

At several points, the analysis has implications for local embedding. In particular, while reputation mechanisms, instability and volatility of network structure, wide scope of communication, frequency of interaction, and trust could all occur at a geographical distance, they are greatly facilitated by proximity and local embedding.

By hypothesis, in a network for exploitation, conditions are more or less the reverse of those that apply to a network for exploration. First we turn to network structure. Dominant designs have emerged, and technological and market uncertainty have decreased. Knowledge becomes more codified and stable, absorptive capacity increases, and knowledge gets widely diffused. With new entry into the emerging market, competition shifts to competition on price, so that considerations of efficiency become crucial. These pressures on cost yield a need to:

- utilise economies of scale, and this can be achieved because due to decreased uncertainty on the part of customers the market has enlarged. As a result, there is increase of scale, a shakeout of producers, and resulting concentration.
- search more widely for the cheapest sources of supply, which is enabled by reduction of uncertainty and emerging standards. Thus there are both needs and opportunities to loosen activities from their local embedding and to extend the network beyond old boundaries.
- Eliminate redundant ties, which is now possible due to increased certainty about network structure, the location and relevance of knowledge, and the ability to absorb it.

Thus there is a *requirement* for a less *dense* structure, which is *enabled* by the fact that now one can identify what competencies are and will remain relevant, who has those competencies, and who is likely to survive in the industry, and there is less need for redundant ties to triangulate knowledge and help in its absorption. Due to the extension of the network, reduction of cognitive distance, codification of knowledge, and slow-down of knowledge change, spillover risk increases. Due to diffusion of knowledge and stabilization of the network, routinization of established practice, and the emergence of standards, interaction becomes less intensive and shifts from developmental to transactional. The increased codification of knowledge furthers diffusion without the need for relation-specific investments in mutual understanding. Investments shift to large-scale production, distribution systems, and brand name, which are all long-term, and increase in size and economic life. In view of such large and often sunk investments, with a long economic life, and to maintain efficient division of labour, network structure is likely to be *stable*. Under systemic conditions, exploitation may require considerable *centrality*.

Concerning the strength of ties, implications of these investments for the *duration* of ties depend on the extent to which they are *relation-specific investments*, which depends on the flexibility of technology: more generic or flexible technology entails that investments are less relation-specific. In increased division of labour for the sake of efficiency, there is an increase in specialization, so that ties entail more specific knowledge, on a narrower *scope* of issues. There is less need for relation-specific trust, and a basis arises for institution-based trust. Reduced uncertainty and codified, diffused knowledge on a more narrow range of issues enable the specification of contracts and the monitoring of compliance, entailing a shift from *trust* and *incentive control* to *opportunity control*. Increased specialization, reduced scope and reduced need for trust reduce *frequency* of interaction, i.e. interaction in the exchange or joint production of new knowledge (purely in terms of transactions, there may be very frequent 'just-in-time' deliveries from suppliers).

In exploitation, the extended reach of the network, into new markets of outputs and inputs, more formality of control, and lesser dependence on reputation mechanisms and trust, entail both opportunities and pressures for local disembedding.

The hypotheses for different conditions for networks for exploration and exploitation are summarized in Table 2.

Table 2 about here

Within these categories of exploration and exploitation, there is still considerable variability of parameters, particularly in exploitation, as a function of contingencies of technology, market, and institutions: systemic/stand-alone, tacitness of knowledge, speed of knowledge change, flexibility of technology, economic life of investments, economies of scale, entry barriers, legal institutions, institutions for trust, etc. These vary with both industry and location.

Empirical illustrations

An example of separation in place, with an ongoing transfer of activities from exploration networks to exploitation networks, is the pharmaceutical industry, where Gilsing (2003) found the following structure. Biotech firms take up an intermediary position, straddling exploration networks with universities and networks with pharmaceutical companies for the transfer from exploration to exploitation. On the whole, their ties with universities conformed to the characteristics of exploration networks, and their ties with pharma companies conformed to the characteristics of exploitation networks. In the exploration network of universities-biotech, Gilsing found high network density, high frequency of interaction, and high specific investment in mutual understanding. However, counter to expectations, he found that ties were fairly strong in control, quite strong in duration, and weak in scope. This deviated from expectations, but could be explained on the basis of the underlying logic, as follows. While knowledge in the *process* of conducting scientific research is highly tacit, the knowledge *output* that was exchanged between players was highly codified, which opened opportunities for contracting. In contrast with the development and commercial use of technology, scientific research is much less concerned with issues of organization, production, commercialization and distribution, so that scope could be low. Long duration of ties was understandable from the fact that mutual understanding required high specific investment, but still appeared to yield a problem of insufficient flexibility and variability, for the sake of novel combinations. However, it turned out that a core network of durable relations in the Netherlands was complemented with a periphery of more variable ties to universities abroad. Note that in the theoretical analysis durable ties were seen to be problematic for exploration only if they are also exclusive. Durable ties need not limit the diversity needed for learning if the nodes involved tap into non-overlapping outside sources, in shorter and more variable ties. In other words, here we find a hybrid network, with a core and a periphery that have compensating strengths and weaknesses.

Perhaps this finding can be generalized, as follows. If a core network of stable relations is needed, for exploitation, or to recoup large specific investments, or to build and maintain trust, the potentially negative effects for exploration may be eliminated by tacking on a peripheral network that feeds the core network with diversity. There may be a lesson here for clusters, to maintain local embedding while tapping into outside sources of diversity.

An example of separation in time, with transformations from exploration to exploitation networks, and back again, Gilsing (2003) found a clear case in the emergence of multi-media, in the Netherlands. However, in the transition there was an intermediate stage, which did indeed show decreased density, increased stability, ties beyond the original local setting, increased centrality, decreased scope, and some increase of contractual control, as hypothesized, but still found considerable trust and frequency of interaction. This could be

explained by the fact that the provision of new web-based information services still needed considerable fine-tuning of mutual fit, in a systemic coherence of elements of the overall system (hard-ware, software and content).

Back to the cycle of discovery

Now we can give a more detailed account of transitions in the cycle of discovery. In the stage of consolidation, one can expect a transition from exploration to exploitation networks, summed up as follows. In network structure a reduction of density, in the elimination of redundant ties, emergence of centrality for the sake of coordinating specialized production activities, and increasing stability, to maintain systemic structures of production. Concerning strength of ties, scope would decrease, in division of labour and specialization, duration would increase, to ensure stability and the recouping of increased and more durable specific investments, frequency of knowledge exchange would decrease, as a result of standardization, contractual control would increase, due to reduced uncertainty, and trust would decrease, due to more arms-length contacts and extension of the network beyond local boundaries, in search of a widened market and cheapest sources of inputs.

Next, in the stage of generalization and attendant differentiation, in new contexts of application, the expectation is that the core exploitation network in the home niche is complemented with a peripheral network that taps into novel contexts. This entails a reduction of centrality, reduced stability, and increased size of the total network. We would expect emerging features of exploration in the peripheral network, and increased frequency of interaction throughout the network, to cope with increasing problems of coordination between centre and periphery.

Next, in the stage of reciprocation, the expectation is that the overall network begins to loosen up in loosely coupled parts, and may break up altogether. A crucial contingency here is the extent to which the old exploitation system needs to be maintained along with exploration, and the extent to which it is systemic. In the multi-media case discussed above, technology was systemic, and development occurred outside the old exploitation system. At the periphery, structure becomes dense, and new investments are required for local understanding, in new niches. Local commitment, at the periphery, may exceed loyalty to the centre. Stability decreases further, ties become weaker in control, and need to become stronger in trust, locally and between centre and periphery. If that is not feasible, in the maintenance of the old exploitation structure, the network is likely to break up, in a separation between core and periphery, to allow for allopatric speciation.

Finally, in accommodation, novel opportunities begin to prove their worth, new networks are formed, and move towards consolidation. Existing networks, in the old centre, come under pressure, and may have to break up, to go along with emerging novel dominant designs.

Note that in both separation in place and separation in time we may meet hybrid networks, with a relatively stable and tight core and a looser peripheral network. In fact, as described, the process of transformation from exploitation networks to exploration networks may entail a break-up and separation in place. It is very difficult, though not inconceivable, to accomplish separation in time while maintaining the overall network, due to contradictions in the combination of a narrow focus with a wide focus and the corresponding mix of incongruent organizational cultures, stability and instability, formal control with informal trust, short and durable ties, in local embedding and outside linkages.

In sum, the cycle of discovery, thus specified in terms of networks for exploitation and exploration, with attendant features in terms of structure and strength of ties, may serve as a model of cluster development.

Conclusions for clusters and cluster policy

This chapter started from the claim that in order to understand learning regions the causal effects of regional features on innovative performance should be elaborated in terms of relations between firms and organizations that intervene between regional features and regional performance. For the analysis of such causality, it elaborated the notion of embeddedness, in three dimensions:

- relational embedding: strength of ties between firms, with seven dimensions
- structural embedding: structure of ties, with four dimensions
- institutional embedding: regional variables that affect conditions of competence and governance in relations between firms, partly through the dimensions of structure and strength of ties

Innovative performance and development of clusters were analysed in terms of the combinations and transitions between exploration and exploitation. According to a ‘cycle of discovery’, exploitation emerges from exploration, on the basis of consolidation, and exploration arises from exploitation, in generalization, differentiation, reciprocation and accommodation. The combination of exploitation and exploration, in the same place and at the same time, is problematic to the extent that exploitation is systemic, i.e. entails a dense, tightly coupled structure of many elements. Under such conditions, there are two basic structural forms for combining exploration and exploitation: separation in time and separation in place.

Exploitation and exploration networks were analysed and contrasted in terms of the dimensions of structure and strength of ties. Some of the analysis contradicts the thesis of the ‘strength of weak ties’. In exploration there are good arguments for density of ties, and for strength of ties in four or five of the seven dimensions of tie strength. For exploitation, density is less and ties tend to be strong, depending on further contingencies, in dimensions where they were weak in exploration.

In sum, cluster development can require fairly drastic changes in the features of embedding, to allow for transitions between exploitation and exploration. The emergence and early consolidation of novelty may require protection from premature competition. Exploration requires local embedding mainly for the sake of governance by reputation and trust based on institutionally embedded morality and close interaction. On the other hand, dis-embedding and break-up of structure, and an escape from established interests, may be needed to proceed from exploitation to exploration. Inertia of a cluster, in the maintenance of established structure and closure to the outside, for the sake of exploitation, can seriously inhibit exploration. In the combination of exploitation and exploration, a trade-off needs to be made between sufficient durability of relations to call forth specific investments in mutual understanding and in the building of trust, and sufficient flexibility to enable new variety for innovation. In search of optimal cognitive distance, a trade-off must be made between distance for the sake of novelty and proximity for the sake of understanding, and between durability to encourage specific investments in mutual understanding and trust, and flexibility or non-exclusiveness, with external linkages, to maintain variety and novelty.

In view of the analysis, public cluster policy is problematic. Overall, the perspective for such policy seems dubious. The purpose, structure and performance of clusters are connected with cluster-specific, local conditions, issuing from a history of development, in complex constellations of variables, as indicated in this chapter. This can yield problems for the transplantation of a successful form from one institutional context to another. A cluster may arise as a compensation for local weaknesses that do not arise elsewhere. The much-lauded

development of networks or clusters in Italy can be attributed, at least in part, to a lack of reliable legal institutions, and a climate of corruption. Similar conditions presently apply in China. Then there is no opportunity for generalized institutions-based trust, and one has to fall back on personalized trust in specific relations. According to Pagden (1988), in southern Italy this goes back to a systematic breakdown of institutions, in the 18th century, as a deliberate policy of the Spanish Habsburg emperor, ruler of the kingdoms of Naples and Sicily, to prevent coherent opposition and sedition.

Thus, one should beware of ambitions for a generic blueprint for clusters that can be applied anywhere. Clusters yield solutions to specific problems/opportunities in specific contexts. As a result, public policy, if such policy is viable at all, should probably retreat from the design of cluster structure to the facilitation of processes of cluster development, as a function of local conditions, in 'giving a nudge here and there'¹², to yield triggers or remove obstacles. This requires an understanding of how clusters may develop and adapt to changing internal and external conditions.

Even that, however, requires caution. Policies appropriate for one stage of cluster development may be opposite to those for another stage. For example, in an early stage of exploration one may need to support escape from established structures, and to shield embryonic novel ventures from the competitive power of established dominant designs and interests, and allowing for and indeed stimulating locally embedded support. For consolidation, on the contrary, one may need to stimulate standards, efficient division of labour, for the purpose of exploitation. After consolidation one may need to shift policy towards the stimulation and facilitation of local disembedding and outside reach. When new radical innovation arises from outside, one may need to encourage the break-up of localized clusters, and to eliminate entry barriers, to allow for novel architectures of old and new components. It is not easy for public bodies to identify which policies are appropriate at what time in which clusters. And when they are able to do this, the delay between design and implementation of policy may yield counter-productive effects, since by the time of implementation requirements may have reversed.

¹² This phrase is due to Bo Carlsson, at a seminar at the Erasmus University Rotterdam on 5th December 2003.

References

- Ashheim, B.T. and A. Isaksen (2002). 'Regional innovation systems: The integration of local "sticky" and global "ubiquitous" knowledge', forthcoming in Journal of Technology Transfer.
- Bettis, R.A. and C. K. Prahalad. (1995). 'The dominant logic: retrospective and extension', Strategic Management Journal, 16/1: 5-14.
- Bogenrieder, I. and B. Nooteboom (2004). 'Learning groups: What types are there?', forthcoming in Organization Studies.
- Boschma, R.A. and J. G. Lambooy (2002). Knowledge, market structure and economic co-ordination: The dynamics of Italian industrial districts, University of Utrecht, Faculty of Spatial Sciences, Department of International Economics and Economic Geography, P.O. Box 80.115, 3508 TC Utrecht.
- Burt, R. (1992). 'Structural holes: The social structure of competition', in N. Nohria and R. Eccles (eds.), Networks and organizations: Structure, form and action. Boston: Harvard Business School Press: 57-91.
- Cohen, M. D. and D. A. Levinthal (1990). 'Absorptive capacity: A new perspective on learning and innovation', Administrative Science Quarterly 35: 128-152.
- Coleman J.S. (1988). 'Social capital in the creation of human capital', American Journal of Sociology, 94: 95-120.
- Cooke, Ph. and R. Huggins (2001). 'High-technology clustering in Cambridge (UK)', forthcoming in A. Amin, S. Goglio and F. Sforzi (eds.), The institutions of local development
- Deutsch, M. (1973). The resolution of conflict: constructive and destructive processes, New Haven: Yale University Press.
- Dimaggio, P. And W. Powell (1983). 'The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields', American Sociological Review 48: 147-160.
- Gilsing, V.A. (2003). Exploitation, exploration and co-evolution in innovation networks, PhD thesis Erasmus Research Institute for Management, Rotterdam.
- Granovetter, M.S. (1973). 'The strength of weak ties', American Journal of Sociology, 78: 1360-1380.
- Gulati, R. (1995). 'Familiarity breeds trust? The implications of repeated ties on contractual choice in alliances', Academy of Management Journal, 38: 85-112.
- Hansen, M.T. (1999). 'The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits', Administrative Science Quarterly, 44/1: 82-111.
- Henderson, R. M. and K. B. Clark (1990). 'Architectural innovation: The reconstruction of existing product technologies and the failure of established firms', Administrative Science Quarterly, 35: 9-30.

- Krackhardt, D. (1999). 'The ties that torture: Simmelian tie analysis in organizations', Research in the Sociology of Organizations, 16: 183-210.
- Lewicki, R.J. and B.B. Bunker (1996). 'Developing and maintaining trust in work relationships', in R.M. Kramer and T.R. Tyler (eds), Trust in organizations: Frontiers of theory and research, Thousand Oaks: Sage Publications: 114-139.
- March, J. (1991). 'Exploration and exploitation in organizational learning', Organization Science, 2/1, 101-123.
- McAllister, D.J. (1995). 'Affect- and cognition based trust as foundations for interpersonal cooperation in organizations', Academy of Management Journal, 38/1: 24-59.
- Nelson R. R. and S. Winter (1982). An evolutionary theory of economic change, Cambridge: University Press.
- Nooteboom, B. (1992). 'Towards a dynamic theory of transactions', Journal of Evolutionary Economics, 2: 281 - 99.
- (1999). Inter-firm alliances: Analysis and design, London: Routledge.
- (2000). Learning and innovation in organizations and economies, Oxford: Oxford University Press.
- (2002). Trust: forms, foundations, functions, failures and figures, Cheltenham UK: Edward Elgar,
- (2003). 'Problems and solutions in knowledge transfer', in D. Fornahl and T. Brenner (eds.), Cooperation, networks and institutions in regional innovation systems, Cheltenham: Edward Elgar: 105-125.
- (2004a). Inter-firm collaboration, learning and networks; An integrated approach, London: Routledge.
- (2004b). 'Competence and governance: How can they be combined?', Cambridge Journal of Economics, forthcoming.
- and V.A. Gilsing (2003). Density and strength of ties in innovation networks: A competence and governance view, draft paper, Erasmus University Rotterdam.
- , W.P.M. van Haverbeke, G.M. Duysters, V.A. Gilsing and A. van den Oord, 'Optimal cognitive distance and absorptive capacity', Research paper, Eindhoven Center for Innovation Studies (ECIS), Eindhoven, the Netherlands
- Oinas, P. and E.J. Malecki (2002). 'The evolution of technologies in time and space: From national to regional to spatial innovation systems', International Regional Science Review, 25/1: 102-131.

Ouchi, W.G. (1980). 'Markets, bureaucracies, clans', Administrative Science Quarterly, 25/1: 129-143.

Pagden, A. (1988). 'The destruction of trust and its economic consequences in the case of eighteenth-century Naples', in D. Gambetta (ed.), Trust, the making and breaking of cooperative relations, Oxford: Blackwell: 127-141.

Polanyi, M. (1962). Personal knowledge, London: Routledge.

Shapiro, S.P. (1987). 'The social control of impersonal trust', American Journal of Sociology, 93: 623-658.

Spender, J. C. (1989). Industry recipes, Oxford: Basil Blackwell.

Williamson. O. E. (1993). 'Calculativeness, trust and economic organization', Journal of Law and Economics, 36: 453-86.

Wuyts, S., M.G. Colombo, S. Dutta and B. Nooteboom (2005). Empirical tests of optimal cognitive distance, Journal of Economic Behaviour and Organization, forthcoming.

Zand, D.E. (1972). 'Trust and managerial problem solving', Administrative Science Quarterly 17/2: 229-239.

Zuchella, A. (2003). Geographic co-location and global value chains: Cluster dynamics and strategic innovations in cluster-based firms, University of Pavia, Italy.

Table 1 Sources of collaboration

	macro; universalistic	micro; particularistic, relation-specific
self-interest		
opportunity control	contracts, legal enforcement	hierarchy, managerial 'fiat',
incentive control	reputation	dependence: unique partner value, switching costs, hostages
altruism	values, social norms of proper conduct, moral obligation, sense of duty, bonds of kinship	empathy, routinisation, benevolence identification, affect, friendship

source: adapted from Nooteboom (2002)

Table 2 Networks for exploration and exploitation

network features	exploration	exploitation
network structure:		
density	high	low
stability	low	high
centrality	low	often high
strength of ties:		
scope	wide	narrow
investment in mutual understanding	high	low
duration	limited*	often long
frequency of interaction	high	low
opportunity control	low	high
incentive control	medium to high	low to medium
trust/openness	high	generally low

* especially when technology is systemic

Source: Nooteboom and Gilsing (2003)