February 13, 2001
10.000 words (net)
13.000 words (incl. Abstract, acknowledgement, references & footnotes)

Anders Malmberg and Peter Maskell

THE ELUSIVE CONCEPT OF LOCALIZATION ECONOMIES

– Towards a Knowledge-based Theory of Spatial Clustering

Paper for the *'Industrial Clusters' Revisited: Innovative Places or Uncharted Spaces?* session,
AAG Annual Conference, New York, 27 february – 3 March 2001.

Anders Malmberg, Professor, PhD
Department of of Social and Economic Geography, Uppsala University
P.O. Box 513, S-751 20 Uppsala, Sweden, Phone +46 18 471 25 34 Fax +46 18 471 74 18 Mob +46 704 250 175
E-mail: Anders.Malmberg@kultgeog.uu.se

Peter Maskell, Professor, Dr.Merc.
Danish Research Unit for Industrial Dynamics (DRUID),
Department of industrial Economics and Strategy (IVS), Copenhagen Business School (CBS)
Howitzvej 60, DK-2000 Frederiksberg C, Denmark, Phone: +45 3815 2881, Fax: +45 3929 2226,
E-Mail: MASKELL@CBS.DK

Abstract
A number of potential advantages of industry agglomeration—or spatial clustering—have since long been identified in the research literature, notably related to shared costs for infrastructure, the build up of a skilled labour force, transaction efficiency, and knowledge spill-overs leading to firm learning and innovation. We identify two shortcomings of existing research on the clustering phenomenon.

First, the abundance of theoretical concepts stands in sharp contrast with the general lack of work aiming to validate these mechanisms empirically and the contradictory evidence found in recent empirical work in the field. Secondly, there is a lack of a unified theoretical framework for analyzing spatial clustering.

In an attempt to remedy the second shortcoming, this paper investigates the nature of the cluster from a knowledge creation or learning perspective. We argue for the need to establish a specific theory of the cluster where learning occupies center stage. The basic requirements for such a theory of the cluster are discussed. Two main components of such a theory are identified: it must explain the existence of the cluster on the one hand and its internal organization on the other.

1. Introduction: economic geography in a knowledge-based economy

The process of globalization has gradually shifted the basis of industrial competitiveness from static price competition towards dynamic improvement, benefiting firms that are more able to innovate and create knowledge than their competitors. The competitive edge of an increasing number of firms is no longer primarily obtained by cost-reduction but mainly by the generation of entrepreneurial rents (Spender, 1994) through innovation, defined in broad Schumpeterian terms as new ways of organizing production processes; accessing new, distinctive markets in new and unconventional ways; or producing new, improved or redesigned goods or services with a significant contribution to the perceived customer benefit (Schumpeter, 1934). The emerging knowledge-based economy is characterized, among other things, by growing importance of economic transactions focused on knowledge itself; rapid qualitative changes in goods and services; and the incorporation of the creation and implementation of change itself into the mission of economic agents (Carter, 1994).

The creation of knowledge is usually seen as a process requiring dedicated investments. At the level of the individual firm, knowledge creating investments are often associated with research and development activities and the adoption of leading-edge technology. Equally important, however, is the investment in 'low-tech' learning and innovation (Maskell, 1998), which takes place when firms, also in fairly traditional industries, learn and innovate while handling and developing mundane day-to-day operations like resource management, logistics, production organization, marketing, sales, distribution, industrial relations, etc. (Malerba, 1989).

Presumably, spatial clustering always contributed to economic growth through enhanced learning processes, but the shift towards a knowledge-based economy has certainly amplified our interest in understanding the nature of this proposition. This had led some scholars to propose that the localized cluster is *the* territorial configuration most likely of enhancing learning processes. Today Silicon Valley and Hollywood are probably the most celebrated examples of successful market-led clusters, but even small nations are also sometimes looked upon in this light (Porter, 1990).

In the context of economic geography, the concept of agglomeration has to do with the spatial concentration of people or economic activity. This phenomenon has attracted research interest over extended time periods. Throughout the 20th century, a literature proliferated which, taken together, contributed to our understanding of why industry agglomerations emerge, and in what ways location close to similar or related firms contributes to the competitiveness of an individual firm.[1]

The concept of agglomeration may have two different meanings in this context (Estall and Buchanan, 1961). One is related to the phenomenon that people and economic activity in general tend to concentrate in cities or industrial core regions. The advantages gained by such behaviour are often referred to as *urbanization economies* (cf. Hoover, 1937, Dicken and Lloyd, 1990). The other is related to the phenomenon that firms within the same or closely related industries tend to gather at certain places. Those mechanisms leading to such behaviour are correspondingly referred to as *localization economies*. It is the latter aspect of agglomeration, which makes up the main focus of the rest of the article. Industry agglomeration, localization and spatial clustering are used synonymously in the literature, to denote the phenomenon that similar or related firms and industries tend to assemble (concentrate, agglomerate, co-locate, cluster) in particular places. Here, we mainly use the concepts localized clusters and spatial clustering to denote the phenomenon of agglomerations of similar and related industries.

After a period of relative neglect during the 1970s and 1980s, the 1990s saw something of boom of research efforts devoted to analyzing and explaining spatial clustering (e.g. Sabel 1989: 18, Porter 1998: 78, Storper 1993: 434, Amin and Thrift 1994: 13, Scott 1998: 398)[2]. We find scholarly interest in this phenomenon across a wide range of academic

[1] Among the "classics" in this field of research can be mentioned Marshall (1890), Weber (1909/1929), Hoover (1937, 1948), Myrdal (1957), Hirschman (1958), Ullman (1958), Jacobs (1961), Chinitz (1961), Pred (1966, 1977) and Greenhut 1970. More recent contributions have been signed by geographers like Scott (1983, 1988, 1998), Amin and Thrift (1992, 1994) and Storper (1997) alongside with business strategists and economists like Porter (1990, 1994, 1998) Krugman (1991a, 1991b, 1991c) and Enright (1998).

[2] Accounts of this literature are found in Harrison (1992), Norton (1992), Malmberg (1996), Baptista (1998), Storper (1995), Bianchi (1998), Yeung (2000) and Hanson (2000).

disciplines and traditions. Of course, individual scholars and disciplines pursue different discourses and analyze the role of geographical space in economic process as part of disparate wider agendas. Some frame their analysis in the context of specifying the role of local in a globalizing world economy, others relate such discussions to a broader transformation of capitalism from fordist to more flexible forms of production organization, while yet others approach the role of localized clusters from the point of view of general business strategy, or attempt to bring economic geography into the core of mainstream economics.

There are several reasons to take the issue of spatial clusters seriously. One is that spatial clustering is at the very core of what research in economic geography is all about. Analysis of spatial clustering brings to the fore concepts like proximity, space, place and milieu – all focal points for research in economic geography. There is a lot to learn about the role of space and place in economic processes by trying to pinpoint the driving forces that make for agglomeration in space of similar and related economic activity. Consequently, a solid theory of the cluster is certainly a good starting point for broader theory formulation in relation to the spatial economy.

Second, this task has obvious policy relevance today as well. Throughout the OECD world (and beyond, as a matter of fact) cluster-based policies have in recent years increasingly been seen as *the* main option in the field industrial and regional policy. As an important element of these policies we find a doctrine saying that regions should specialize industrially and promote the dynamics of agglomeration in order to gain or sustain competitiveness and prosperity.

Taken together, this means that spatial clustering is important enough to justify an attempt to track its conceptual underpinnings and empirical status, even if it means that sometimes different and perhaps not always compatible accounts are ‘lumped together’ in an eclectic way into something defined as ‘spatial clustering research’. We do maintain, however, that there is some justification in doing so. The point of departure of this article, therefore, is that the individual additions to this literature can still be read as contributions to a common research agenda: that of understanding the role of space and place in industrial development in general, and the phenomenon of spatial clustering in particular.

The aim of the following two main sections article is fairly modest: to review and summarize the most important arguments making up what could be labeled ‘agglomeration theory' or 'clustering theory’, and subsequently to direct our attention to the somewhat discomforting fact that the theoretical mechanisms identified to this day by and large evade successful empirical validation. Even though the phenomenon of spatial clustering is indeed one of the main ‘classical issues’ of economic geography, and despite the fact that considerable research efforts has been made over the last decade, we have to acknowledge that its causes and effects remain elusive.

Furthermore, and perhaps more discomforting, we believe that research in economic geography and related disciplines has so far failed to come up with something close to a satisfactory theory of the localized cluster. Remedying this latter shortcoming is precisely what we are aiming at in the final major section of the article. Here we propose a way of structuring our understanding of how the cluster partakes in knowledge creation. In dealing with this issue, we raise a set of questions regarding the way in which economic performance is related to space in general, and to the role of localized learning in particular.

2. A review of earlier research on spatial clustering

Existing literature provides two types of knowledge on spatial clustering. One source of knowledge is to be found in ideographic, historical work on the origin and development of various types of localized clusters. Another is the more theoretical work that tries to identify the mechanisms which give economic and other advantages to the individual firm located in a cluster. We will look briefly into the first type of literature before turning to the latter.

Genealogical approaches to the existence of the cluster

One part of the literature on industry agglomeration is predominantly ideographic, trying to capture the historical origin and evolution of localized clusters. Ever so often, one finds that localized clusters of similar and related activity have deep historical roots. One way to analyze the emergence of a localized cluster is thus simply to try to trace its roots backwards in history. If a certain amount of simplification is allowed, we may argue that these accounts tend to highlight three different mechanisms.

The ‘origin’ of the cluster in such accounts, i.e. the event or action which triggered subsequent developments, sometimes turns out to be related to some more or less traditional factor of location, i.e. some natural or social asset that at a certain point in time turned out to be a very important location factor for a particular type of economic activity.

Equally often, however, there seems to be a considerable amount of chance involved. Some person did, for some reason, get the idea to engage in a certain type of economic activity. The literature is mostly silent of the mechanisms causing such more or less random historical events which sometimes trigger the development of clusters, except for the fact that most new firms start out at the place of residence of their founder, presumably because the entrepreneur is well known locally and can therefore more easily get access to the capital and other resources necessary to get started. (Kahlin 1961: 132).

Once an activity has started in a particular place, two more mechanisms come into play, contributing to subsequent developments. The first is that a successful economic activity in one place is often followed by other similar or related activities. If one firm turns out to be successful, this seems ever so often to be a sufficient reason for similar firms to emerge in the neighbourhood. Althin described this in a nice way:

> People find that, if one is successful, why should not also others be able to succeed, as they have closely seen their neighbour or maybe even their former employer, from whom they have broken loose, do. In this way emerges, through the power of the good example, one firm after the other within the same industry close to one another in the same place. (Althin 1948: 96 (our translation)).

Another mechanism contributing to the development of clusters is related to the fact that firms, once they have been rooted in a place, rarely re-locate. Ross (1896) observed this already more than a century ago, and the observation has then been repeated and elaborated upon over time:

> The power of a locality to hold an industry (...) greatly exceeds its original power to attract. The new locality must not only excel the old, but it must excel it by margin enough to more than offset the resisting power of the matrix. (Ross 1896: 265)

We may thus argue that the emergence of spatial clusters of similar and related economic activity is related to three factors: they often originate in a series of events leading to the start of a new firm at the place of residence of the entrepreneur; they develop through spin-offs and imitation within the local milieu; and they are sustained by various forms of inertia, meaning that firms rarely relocate once they have been reproduced in a place.

This is of course not the end of the story. It is certainly possible to add to these three, another three or four elements that would contribute to portray not only the origin and rise of a cluster, but also its saturation, crisis and possible rejuvenation. Such an account would in addition to pioneering, imitation and spin-offs, include concepts like a deepening division of labour between local firms, the creation of a local culture, supporting infrastructures and institutions adopted to the proliferation industry, the establishment of the place as a brand of the industry and subsequent attraction on resources from the outside (people, capital, firms). Such a story would, ultimately also include a phase of consolidation – at some stage some firms tend to take on leading roles and this often means that they take over other firms in the cluster such that an initially small-firm based agglomeration often ends up by being dominated by a limited number of larger firms. Also, history tells us that most localized clusters, sooner or later, run into problems. Elements of petrification are often revealed at points in history when technological or other developments call for rapid restructuring. This may mean that what was once a leading centre of dynamism within a given line of business ends up as an ‘old industrial region’, facing great problems of renewal and finding itself being out-competed by firms located elsewhere. Then again, there are also examples of how crisis-ridden clusters manage to ‘reinvent’ themselves, such that they can actually retrieve some of their former greatness.

The above is not intended as a deterministic stage model of the rise and fall of localized clusters. Clusters, of course, may follow many different development paths. The aim is rather just to establish as a fact that some of our knowledge on the clustering phenomenon comes from studying industrial history, and that it is possible through such accounts to gain insight into the processes which make for the development of localized clusters. In addition, when attempting a theoretical analysis of the advantages of spatial clustering, it should be useful to keep in mind that the impact of various types of localization economies may vary considerably over time. Factors that are important in an emerging cluster might be less significant in a mature cluster.

Approaches based on cost-reduction

The more theoretically oriented part of the literature of agglomeration does usually not focus on the origin and subsequent historical development of localized clusters. Rather it aims at explaining the existence of spatial clustering

by identifying and analyzing those permanent advantages that may accrue to firms located in close juxtaposition to other similar and related firms, rather than being located in isolation. Thus, the concept of localization economies refers to these advantages. Three different mechanisms have traditionally been identified in this context.[3]

First, there are benefits to be gained from the possibility for agglomerated firms to share the cost for certain collective resources among several firms. This applies in particular to the cost of establishing the wanted infrastructure. When an agglomeration of similar or related firms is established, there is also a potential to adjust the local infrastructure, the educational system and other types of collective goods after the needs of this particular industry. The first mechanism can thus be labeled *reduced costs for producing and maintaining a dedicated infrastructure and other collective resources*.

Second, agglomeration makes for the development of a local labour market for specialized skills. The establishment of a local pool of skilled labour has been proposed as a major element of the localization economies ever since Alfred Marshall more than a century ago wrote so elegantly about the advantages of being located in an industrial district:

> Again, in all but the earliest stages of economic development a localized industry gains a great advantage from the fact that it offers a constant market for skill. Employers are apt to resort to any place where they are likely to find a good choice of workers with the special skill which they require; while men seeking employment naturally go places where there are many employers who need such skill as theirs and where therefore it is likely to find a good market. (Marshall (1890) 1920: 270)

Marshall also pointed out, that especially when the infrastructure (and thus the commuting possibilities) were badly developed, risk-considerations would naturally attract workers to clusters of similar firms thereby creating an 'obstacle to the success of any business in which special skills is needed, but which is not in the neighbourhood of others like it' (Marshall (1890) 1920: 272).

Thus, it can be argued that local labour markets function better, both from the point of view of firms and employees, if there are several similar and related firms around (cf. Krugman 1991a). *Well-functioning markets for specialized skills* can thus be added to the list of localization economies.

Third, firms in agglomerations can reduce their costs as inter-firm transactions and shipments are simplified when the distance between firms is negligible. The customer firm which can place an order with a supplier located down the street will gain an advantage in relation to a competitor which has to travel long distances to discuss a deal with its supplier. *Reduced interaction costs for co-located trading partners* may thus be the shorthand for this mechanism (Scott 1983), with is described by Porter (1998) in the following way:

> ... the proximity of companies and institutions in one location – and the repeated exchanges among them – fosters better coordination and trust. Thus clusters mitigate the problems inherent in arm's-length relationships without imposing the inflexibilities of vertical integration or the management challenges of creating and maintaining formal linkages such as networks, alliances, and partnerships. (Porter 1998: 80)

The most recent approach: Knowledge spill-overs

Towards the recent turn of the century, a far more sophisticated account has come to occupy centre stage in the discourse of the cluster, and occasioning a number of different research propositions to unfold. Here, a fourth major factor is introduced: localized clusters of similar and related firms form the basis of a local milieu that may facilitate *knowledge spill-overs* and stimulate various forms of learning and adaptation. This is the aspect of spatial clustering that has attracted the bulk of research interest during the 1990s (Porter 1990, Malmberg *et al.* 1996, Maskell *et al.* 1998, Maskell & Malmberg 1999). The insight that proximity between related firms leads to the development of relations that in different ways may stimulate various exchanges of information and knowledge is however far from new: It can be

[3] Several of these mechanisms have been proposed as factors making for urbanization economies as well as localization economies

found in Marshall (1890) and in a number of subsequent studies occasional also in economic geography. Estall and Buchanan thus emphasize that:

> ... a close relationship, almost a partnership, grows up among related firms in a given geographical area. The ability, for example, of members of the group to *meet without inconvenience* to discuss common problems and matters of mutual interest is a not inconsiderable advantage of close geographical association. (Estall and Buchanan 1961: 109, emphasis added)

The general argument is that a local industrial structure with many firms competing in the same industry or collaborating across related industries tend to trigger processes which create not only dynamism and flexibility in general, but also learning and innovation. In such an environment, chances are greater that an individual firm gets in touch with actors that have developed or been early adopters of new technology. The flow of industry related information and knowledge is generally more abundant to the advantage of all firms involved. A local culture with specific norms, values and institutions (formal and informal) makes it possible to transfer also tacit forms of knowledge from one actor to another.

Precisely when it comes to the impact of spatial clustering on the learning and innovation ability of firms, economic geography has in recent years in a very fruitful way begun to interact with other economic and social science disciplines studying firm competitiveness, learning and innovation (cf. Alvstam 1998). Analyses of how different nations exhibit distinct innovation systems (Lundvall 1992) have been adopted by geographers who have tried to show that within individual countries, there exist regional innovation systems which are partly related to the existence of agglomerations of related firms and industries (Cooke 1996, Morgan 1997, Asheim and Isaksen 1996).

Michael Porter's model of how long-term national industrial competitiveness is created has been used in a similar way. Porter (1990) argues that the characteristics of a local milieu which determine the innovative ability of firms are captured by four interrelated driving forces, related to: factor conditions, demand conditions, related and supporting industries; and firms structure, strategy and rivalry. Porter's model was originally developed in an attempt to analyse national differences in industrial competitiveness, but it has subsequently also been used as a model for explaining the competitive advantage of localized clusters (Porter, 1990, 1994, 1998, Malmberg *et al.* 1996, Enright 1998, Larsson, 1998).

In the context of research on, e.g. innovative milieus, industrial districts, innovation systems or the economics of proximity, something of "a modern approach" to the analysis of spatial clustering has gradually proliferated. This is not to deny that in several important respects, the GREMI approach to innovative milieus (Maillat, 1991, 1998, Camagni, 1995, Ratti *et al.,* 1997), is different from, for instance, Marshallian analyses of the Italian industrial districts (Beccatini, 1990, Brusco 1982, Dei Ottati, 1994, Bellandi, 1989, 1996, Gottardi 1996), the French 'proximité' tradition (Kirat and Lung, 1999, Blanc and Sierra, 1999) or econometric of cluster analyses (Swann *et al.*, 1998), just as they don't conform to the 'systemic' analysis of innovations whether national (Lundvall, 1992) or regional (Markusen *et al.,* 1986, Saxenian, 1994), nor to the managerial approach applied by Porter (1990).

Allowing ourselves a certain amount of simplification, however, we argue that a number of different schools of thought have a shared point of departure in seeing the long-term competitiveness of a firm as determined by its ability to innovate and engage on processes of continuous learning. Cultural, institutional and infrastructural factors in the local milieu affect the general climate within which firms develop. Spatial proximity between actors does in various ways make easier those knowledge spill-overs and interactions which form the basis for innovation and learning, and it is in this context the spatial clustering becomes a key focus of analysis. Most, if not all, of these approaches, have knowledge spill-overs as major research focus. Still, we would argue, to the extent that the individual studies within these divergent approaches contain a specific theoretical account of the cluster it is, again, based on *the costs of interaction.*

The cluster exists, it is implied, because the co-location of firms cut the expenses of identifying, accessing and transferring knowledge. Some studies have emphasized how firms will cut the costs of interacting if located in a cluster characterized by trust and other features of social capital (Maskell 2000) that help reduce malfeasance, induce reliable information to

be volunteered, cause agreements to be honoured, enable employees to share tacit information, and place negotiators on the same wave-length.

Other studies have, however, noted how co-location might create advantages of knowledge spill-overs even if trust-levels are insignificant, like, for instance, in Silicon Valley where "nobody knows anybody else's mother", and where no deep history or complex familial ties exist (Cohen and Fields, 1999: 2).

Yet, studies across the different approaches have run into difficulties when attempting to identify empirically and specify theoretically the localization economies that should account for the existence of the cluster. Without being unfair, we believe, one can argue that they do not contain any theory specifying how the territorial configuration of many co-localized firms in related industries would be able to create knowledge in ways not equally available, for example, to a single but larger firm, or to well-functioning (trust-based) networks of interacting firms, each placed at a different location. A common practice is to reverse the chain of causality. Analyses of innovative clusters do, for instance, seldom start, as one would expect, with identifying how knowledge is shared and technology transferred in a way that enhance the competitiveness of firms, followed by showing how the cluster emerged as a *consequence* of these benefits. Instead, the existing performance of firms already located in the cluster is assumed to be shaped by hypothetical local spill-overs. In much of the literature there is, thus, an obvious risk of ending up with models of circular causation: When the cluster-generating economic mechanisms cannot be observed, their existence is "proven" with reference to the actual existence of the cluster (cf. Krugman, 1995).

3. Shortcomings and challenges in existing research on spatial clustering

There is an obvious lack of systematic attempts to assess empirically the precise mechanisms behind and the magnitude of localization economies (cf. Sabel 1989, Malmberg 1997, Markusen 1999, Larsson 1998). While economic geographers and others have devoted considerable efforts in documenting the existence of industry agglomeration and to assess the level of inter-firm transactions in such settings, not many—and with few exceptions also less successful—attempts have been made in terms of showing differences in firm performance between those located inside or outside agglomerations (Appold 1995, Swann *et al.* 1998).

There are different explanations to this state of the art. One is that this is, after all, a relatively new, or at least relatively newly rediscovered, line of research. A second one is presumably that rigorous empirical testing presupposes a certain amount of simplification (some would say 'vulgarization') of a relatively complex theoretical argument, something that in itself is regarded as dubious by many.

Several problems can be identified in this context. First, most empirical information on industry agglomeration is based on case study material. This problem is aggravated by the fact that the selection of cases seems to be quite biased, with an analytical emphasis on high-tech industries and regional success stories (Wiig and Wood 1995). In contrast, there is a lack of systematic empirical work on what Lundquist and Olander (1998) refer to as the "grey mass" of more mundane, not so spectacular firms, industries and regions. A second problem is that many analyses are marked by a static research design. They portray inter-firm relations and patterns of interaction at one point in time, which means that ever so often attempts are made to draw conclusions about dynamic processes based on cross-sectional data (Staber 1997). The most serious problem, however, is that existing empirical results far from always give clear-cut support for the theoretical argument: the concept of localization economies indeed remains elusive.

There are at least three sets of empirical questions that seems to be of crucial importance here: How common is spatial clustering across industries and nations? In what way and on what magnitude do firms in localized clusters interact? Which characteristics of a local milieu are particularly important for the competitiveness of firms, and are firms in localized clusters more competitive than other firms? Below, some examples will be given of relatively recent attempts to answer these questions.

How common is the agglomeration phenomenon?
Much of our knowledge on industry agglomeration is anecdotal in the sense that builds on more or less detailed accounts of individual regions. Not least is there an abundance of analyses of what has been referred to as 'the holy trinity' of economic geography—Silicon Valley, the Third Italy and Baden-Württemberg—three regional success stories which during the late 1980s and early 1990s came to be regarded as archetypal or 'paradigmatic' in various respects (Winter 1997). Even though the list of examples have been extended well beyond those three regions (see

Porter 1998), the fact that so much research on the agglomeration phenomenon is based on 'stories rather than statistics' has been a source of critique (Head *et al.* 1995).

There are, however, studies which, based on aggregate industrial statistics are showing high, and rising, levels of agglomerations across a broad range of industries. Krugman (1991a) calculates Gini-coefficients for the distribution of 106 industries across US federal states and find that many industries are indeed strongly agglomerated at this level. Enright (1993) presents a statistical analysis of the forces making for this pattern (see also Ellison and Glaeser 1994). Malmberg and Maskell (1997) show that a majority of industries (ISIC 4-digit level) in four Nordic countries (Sweden, Finland, Norway and Denmark) have become more agglomerated over a 20 year period, despite the fact that manufacturing industry taken as a whole have become more dispersed over the same period. By defining a regional industry agglomeration as a local labour market area where the location quotient is larger than three (i.e. there are three times more jobs in a particular industry than it would have been if the region would have had a proportional share of national employment in the sector), Isaksen (1996) identifies 143 agglomerations in Norway and these agglomerations account for a fifth of national employment in total manufacturing industry. Head *et al.* (1995) studies the location decisions of 751 Japanese plants in the US since 1980 and find that agglomeration economies at the level of the industries play a major role in explaining the location pattern. All in all, there is support for maintaining the thesis that spatial industry agglomeration is a phenomenon widespread enough to justify further study.

Proximity matters, but to what degree are linkages local?
Localization economies are, as has already been stated, usually seen to arise because firms taking part in local exchange will benefit from lower costs. Presence in an agglomeration is held to improve the profitability of firms by reducing their costs of exchange of both goods and information (Appold 1995). According to Scott (1983, 1988) the economies of agglomeration are particularly manifest when linkages between firms tend to be small scale, varying and unpredictable, and thus where the interaction costs will rise rapidly with increasing distance. In other words, the greater the product and production flexibility, the more important are the agglomeration economies.[4]

This emphasis on the efficiency and intensity of local transaction is paradoxical, given that extensive local transactions have simply not been found in the empirical analyses that have tried to measure this. Thus, most firms have very limited buyer-supplier relations with other firms in their region of location, even in cases where there are many related firms in the same region (McCann 1995). Larsson and Lundmark (1991) analyze patterns of interaction between firms in the Kista area in northern Stockholm, and find that firms in this telecommunications and IT cluster report on few business transaction with other firms in the area. Angel and Engstrom's (1995) analysis of the American personal computer industry find no support for the proposition that computer producers and their suppliers increasingly agglomerate at the same places. Larsson (1998) shows, based on a survey to more than 300 Swedish machinery producers, that the local market (defined as a 100 km radius) is of marginal importance for the sales and purchases of these firms.

How important is the local milieu for the innovative ability and competitiveness of firms?
As we have already noted, the 1990s has meant a re-orientation in the study of spatial clustering, such that the previously dominating focus of inter-firm transactions has generally given way to an increased emphasis on knowledge spill-overs and localized interactions leading to processes of learning and innovation (See Storper 1995 or Malmberg 1997 for reviews, and Hudson 1999 for a critique of this 'learning turn'). To a degree, this turn can be interpreted as a reaction to the failure to capture in empirical analysis the local transactions previously held to explain agglomeration phenomena. When realizing that firms in a localized cluster do not conduct much business together (or when applying Storper's terminology: when the traded interdependencies between firms in agglomeration proved to be of modest magnitude), a close-hand alternative is to look for other types of (untraded) interdependencies, and then—for various reasons—knowledge spill-overs stood out as a promising hypothesis.

It has not, however, been easy to document, in empirical analyses, the existence of localized learning patterns beyond the level of case studies of individual regions. Thus, Harrison *et al.* (1996), who collected data on a large number of firms in the American metal industry, find insignificant support for the existence of localization economies in the spatial pattern of innovation and industrial competitiveness. In Larsson's (1998) study of Swedish producers of machinery, the result showed that these firms did indeed report on localized relations in the field of technological relations (as

[4] Even for presumably very 'flexibility demanding' firms in Silicon Valley or Hollywood, it would, however, be unwise to disregard the fact that not all relevant actors and holders of knowledge are to be found in the neighbourhood. Analyses should be designed to throw light on the question of which types of relations and interactions that tend to be (more) local and (more) global, respectively, rather than to put emphasis *a priori* to one or the other.

compared to the very much globalized flows of supplies and finished goods), but the correlation between localized technology relations on the one hand and innovative ability on the other, turned out to contradict the arguments put forward in agglomeration theory (Larsson 1998, Larsson and Malmberg 1999). Contrary to what was expected, it was the firms that reported on the most spatially extended patterns of technological collaboration, which seemed to be most innovative. Malmberg *et al.* (2000) show that the presence of other exporting firms in the same industry in the local milieu had an almost negligible effect on the export performance of Swedish exporting manufacturing firms. Internal economies of scale (i.e. the size of the export firm) and urbanization economies (i.e. the presence of other firms—regardless of industry—in the local milieu) had a much larger effect.

Lundquist's (1996) study of how the local milieu affects the international competitiveness of Swedish firms does not in a direct way try to test the impact of agglomerations economies, but what it does show is that firms' dependence on their local milieu varies considerably, also for highly competitive firms. Fuellhart (1999) shows that firms in the carpet industry, a highly clustered industry in the US, report on little use and availability of common sources of information compared to firms in other industries and regions. In a study by Staber (1996), no support is found for the proposition that inter-firm relations are embedded in the social structure of the local milieu or that firms in any important sense utilizes local institutional arrangements which can support collaboration and innovation. Given that this study is made in one of the regions of 'the holy trinity', the conclusion is noteworthy: "... it seems impossible to claim that Baden-Württemberg represents an example of an industrial district" (Staber 1996: 313). Herrigel (1996), who also studies Baden-Württemberg, argues that the industrial crisis of the region in the 1990s has to do with rigidities in the German productive system which can be traced back to the very same institutional arrangements which were previously regarded as the main source of competitiveness of the firms of the region. Taken together, these results put distinct question marks for several of the theoretical propositions advanced in the literature on spatial clustering.

In the concluding section of the article, we will discuss some methodological problems that might cause this predicament of empirical research on spatial clustering. We believe, however, that the causes are not just related to problems of coming up with appropriate research designs and sufficiently good data. There are also conceptual problems, to which we turn in the following section.

4. A new theory of spatial clustering: the knowledge-based approach

A main shortcoming in existing accounts of spatial agglomeration seems to be a superficial understanding of the nature of the explanatory problem at hand. In particular it is not always realized that although co-location might help diminish the costs of interaction, the best way by far reduce such costs is by joining the different activities and placing them under one common ownership:

> Internal trading changes the incentives of the parties and enables the firm to bring managerial control devices to bear on the transaction, thereby attenuating costly haggling and disruptions and other manifestations of non-cooperative behaviour. Exchange can then proceed at lower cost and with higher returns to the participants. (Teece 1980: 232)

When it comes to interaction costs the one-plant, single firm is superior to all market configurations imaginable – even to the high-trust cluster – simply because many costs of interaction are largely eliminated. It is thus obvious that a theory explaining the existence of the cluster can never be based on (interaction) cost reductions only. In consequence, a new point of departure must be taken by rephrasing the basic research question.

As we see it, a satisfactory theory of spatial clustering must be able to perform several functions.[5] It must include an explanation for *the existence of the cluster*. This means that it must specify the process or processes that impel similar and related firms to cluster at one place and – by doing so – thrive. More specifically, the theory must provide an explanation for the advantages that *many* related firms might accrue when co-located but which are not available to a hypothetical *single* firm carrying out precisely the same activities, even if at the same location, using the same suppliers,

[5] This line of thinking and the argument presented in this section is based on Maskell (2001b).

customers and workforce. Furthermore, a theory of the cluster must contain an explanation for *the internal organization of the cluster*. The theory must, finally, be dynamic in the sense that it can encompass the possibility of and reasons for the decline of formerly successful clusters.

The existence of the cluster: The advantage of being in the same place

Firms engaged in similar activities constitute the *horizontal dimension* while firms providing related activities make up the *vertical dimension* of the cluster.

Most analyses aim to transcend the traditional way of grouping firms into industries, by introducing the vertical dimension, where division of labour and complementarity are the most important dimensions of relatedness. Different stages of a production process carried out along the vertical dimension of the cluster require some kind of co-ordination as the output of one firm may be the input of another. Firms co-operate in matching their related plans in advance since '…the one that make the heads of the pins must be certain of the cooperation of the one who makes the points if he does not want to run the risk of producing pin heads in vain' (List, 1841: 150). The vertical dimension of the cluster thus consists of firms linked through input/output-relations while possessing knowledge, experience, and skills useful for undertaking dissimilar but complementary activities. As we have seen, this type of transactional links make up the core of much analysis of localization economies despite the fact that few empirical studies have been able to show that such linkages are indeed predominantly local.

On the other hand, most well known examples of industry agglomerations are obviously based on the horizontal dimension, since they are made up of several firms operating in the same industry. Industries are normally defined on the basis of similarity in the final output: i.e. firms producing the same type of goods or services are defined as belonging to the same industry. Firms producing similar output do—in principle—compete on the same market. While the firms in vertical dimension of the cluster are business partners and collaborators, the horizontal dimension consist mainly of rivals and competitors.

The dynamic effects of local competition are however relatively neglected in most theoretical accounts of agglomeration, with Porter's (1990) concept of domestic rivalry as a notable exception. However, Alfred Marshall (1890) long ago hinted at an explanation for the existence of the cluster along the horizontal dimension of the cluster. Marshall's reflection concerns the advantages of *variation* that are caused by the parallel performance of similar tasks. It is based on the conjecture that firms (i.e. owners, managers, and employees) have different perceptive powers, divergent insights and varying attitudes. Their different assessment of the information at hand results from the idiosyncratic and at least partly tacit way by which it is assembled and interpreted (Casson, 1982). Consequently, firms develop a variety of solutions as an intricate part of the their daily operation, when holding dissimilar beliefs about their chances of success if using one of several possible approaches to similar problems (von Hayek, 1937).

The variation in the cluster between and among firms doing similar things promotes the generation of ideas and guide interpretations without imposing uniformity. Even when trying hard it would be extremely difficult and often impossible for a single (multidivisional) firm to replicate internally the process of parallel experimentation taking place among independent firms doing similar things in the cluster. For as Loasby explains:

> Competing visions between firms are necessary features of an evolutionary or experimental economy. But competing visions within firms, unless very carefully managed, and limited in scope, cause trouble. (Loasby 2000: 11)

Co-located firms undertaking similar activities find themselves in a situation where every difference in the solutions chosen, however small, can be observed and compared. First, regarding *observability* spatial proximity brings with it the special feature of spontaneous, automatic observation. Just as people in a residential area simply cannot help noticing what their nearest neighbours do (regrettably, many would say), business firms often have remarkably good knowledge of the undertakings of nearby firms even if that do not make any dedicated efforts in systematic monitoring. If those neighbouring firms are in the similar business, it is the more likely that the observing firm will understand, and learn from, what it sees.

The second element is *comparability*. While it might be easy for firms to blame an inadequate local factor market when confronted with the superior performance of competitors located far away, it is less so when the premium producer lies down the street. The sharing of common conditions, opportunities and threats make the strength and weaknesses of each individual firm apparent to the management, the owners, the employees and everyone else who cares to take an interest.

Each firm in the horizontal dimension of the cluster is provided with the information of the possibilities to improve and the incentives to do so. This is also the essence of the domestic/local rivalry component of Porter's diamond model.

It is by watching, discussing, and comparing dissimilar solutions that horizontally related firms in the cluster become engaged in the process of learning and continuous improvement, on which their survival depends. Harrison C. White saw this very clearly in his account for the essence of competition[6]:

> Markets are self-reproducing social structures among specific cliques of firms and other actors who evolve roles from observing each other's behavior. I argue that the key fact is that producers watch each within a market. Within weeks after Roger Bannister broke the four-minute mile, others were doing so because they defined realities and rewards by watching what other "producers" did, not by guessing and speculating on what the crowds wanted or the judges said. Markets are not defined by a set of buyers, as some of our habits of speech suggest, nor are the producers obsessed with speculations on an amorphous demand. I insist that what a firm does in a market is to watch the competition in terms of observables (White, 1981: 518)

Successful experiments can easily be distinguished from the less successful by the knowledgeable local observers. Promising avenues identified by one firm will soon be available to others. Firms with similar capabilities in the horizontal dimension of the cluster constantly imitate the proven or foreseeable success of others while adding some idea of their own. If the firms of the cluster were to be spread throughout a large city among many unrelated businesses their ability to learn from each other's mistakes and successes would therefore be severely restricted.

It might be worthwhile to emphasize an essentially Darwinian feature of variation: it does not presuppose any trust whatsoever among the firms as a prerequisite for learning. It does not require any close contact or even an arm-length interaction between the firms. While suppliers and customers in a vertically organized production chain simply *need* to interact with each other in order to do business; competitors don't. Most *manifest* relations in the cluster will therefore be along the vertical dimension,[7] and this creates inherent difficulties for any empirical investigation of the horizontal benefits of clustering when relying on what is easily observable.

That firms in the horizontal dimension of the cluster do not necessarily *have* to interact is not the same as implying that they never in reality co-operate by helping each other in overcoming technical problems, by lending materials and swapping surplus capacity or by exchanging information. In fact, they may interact regularly, even intimately so, in order to forward some particular scheme (Allen, 1983). On the other hand, they might just as well hate each other intensely, newer exchanging a single useful comment. This might be one explanation to the disappointing results of many empirical studies trying to document the abundance of inter-firm collaboration in localized clusters.

The proposition put forward here simply suggests that the cluster exist because of localization economies *that are independent of the degree of internal interaction at least in principle*. The only requirement is that *many firms undertaking similar activities* are placed in circumstances where they can monitor each other constantly, closely, and almost without effort or costs. Variation emanates naturally when firms possessing somewhat similar bodies of knowledge must act on incomplete and uncertain information. Co-location helps firms identify and imitate superiour solutions while combining them with ideas of their own.

The cluster exists, we thus propose, primarily because of the benefits of enhanced knowledge creation that takes place when many co-located firms undertake similar activities. These benefits are enhanced by the ease of interaction across different bodies of knowledge when sharing the same local circumstances. We shall now turn our attention to the important, but secondary, types of knowledge creation stemming from the vertical dimension of the cluster, once its existence has been secured.

---

[6] White's proposition can be found in several later works and his idea is at the core of Porter's (1990) concept of rivalry.

[7] This theoretical point has been supported by empirical findings (Håkansson, 1987).

The internal differentiation of the cluster

Specialized suppliers and sophisticated customers become attracted to the cluster, once established, by the particular opportunities available. The vertical dimension of the cluster might, however, also be developed by task-partitioning as firms become more specialized. Very specialized firms often find solutions and notice peculiarities otherwise overlooked, even when specializing in performing some particularly trivial tasks. The perception of minute anomalies, previously unnoticed, leads in turn to new insights and ways of improvement and, as a result, to a general acceleration of the growth of knowledge.

By creating an appropriate vertical differentiation, the cluster can therefore develop knowledge far beyond the reach of any of its members. With the internal growth of knowledge, new economic activities become possible, the economy of the cluster progresses, and the resulting extension of the internal market makes the process self-reinforcing (Young, 1928).

It follows from the concept of variation dealt with above, that if all firms specialize so much that they all become part of the vertical dimension then all learning through variation and monitoring must necessarily cease. A continued division of labour among firms in the cluster might thus only be expedient for the overall knowledge creation up to a certain point. Beyond that point, the benefits might be offset by the corresponding reduction in knowledge creation as variation is diminished and fewer possible avenues of progress are tried out in parallel. Only by a steady increase in the number of firms in the cluster would it be possible to create knowledge simultaneously by variation and by the division of labour, and this restriction should be kept in mind when studying empirically the internal structure and the growth of the clusters.

The benefits and returns of the division of labour must also be balanced against the obstacles and costs of reassembling and coordinating separate bodies of knowledge (Loasby, 1999). Dispersed knowledge must be reassembled in order to be useful. Inter-firm learning[8] is subject to both thresholds, before the knowledge-bases of firms divided have grown sufficiently apart for interaction to imply learning, and ceilings, after which the cognitive distance becomes too great for firms to bridge, and where learning will cease. Even when situated between the two extremes it can be very difficult, and sometimes even impossible, to transfer and re-use knowledge even if it is openly available, but often less so when the transfer takes place within a community that share the same language, beliefs, judgements and values. When firms co-locate, a spatially defined community is usually formed that makes it easier for them to bridge communication gaps resulting from heterogeneous knowledge endowments. The innovative capabilities of firms are enhanced because co-location can provide them with an arsenal of instruments to obtain and understand even the most subtle, elusive and complex information of possible relevance developed because they were *separate firms* pursuing their individual agenda. Hence, the process of clustering tilts the balance between advantages of specialization and costs of coordination so that a higher level of knowledge creation can be obtained. The ability to de-code and utilize knowledge residing elsewhere is not a phenomenon to be captured by input/output analyses of trade flows or accounts of business contact patterns.

Clusters and the wider local milieu: the institutional fit

The learning processes identified so far are rooted in the day-to-day operations of the firms in the cluster, but influenced by institutions such as social capital. It is reasonable to assume that the cluster's particular set of institutions has emerged as a response to the special requirements of the activities performed by the firms making up the cluster.[9] Once a dominating institutional pattern has been established it will attract those firms most compatible with it. A favourable institutional pattern will also attract entrepreneurs with ambitions to start firms in the particular industry. This is also why many of the most talented wannabes within the film industry tend to end up in Hollywood and many of worlds best specialists in information and communication technology find themselves in Silicon Valley.

[8] The product innovation literature has firmly established that firms learn from each other when inter-acting. See, for instance, Rosenberg 1972, Freeman 1982, 1991, Håkansson 1987, Kline and Rosenberg 1986, Hagedoorn and Schakenraad 1992, OECD 1992, DeBresson 1998.

[9] On the national level recent research has proved the existence of such a correlation between patterns of specialization in production and trade, on the one hand, and the knowledge-base on the other (Archibugi and Pianta, 1992).

There is thus a fundamental *interdependence* between the economic structure and the institutions of the cluster. Just as the set of firms undertaking similar and complementary activities differ among clusters, so do institutions. It is the particular set of activities in the localized cluster that affects what is done within and among the firms and therefore *what is learnt*, but the institutions in the cluster define how things are done and consequently *how learning takes place* (Lundvall and Maskell, 2000). Different modes of learning create different outcomes which might be more or less suited to the challenges and opportunities presented by the world outside the cluster.

Precisely because the institutions developed within the cluster become very specific, they will differ from one cluster to the next (Maskell and Törnqvist, 1999). The mechanisms that reduce cognitive distance *within* the cluster tend to increase the cognitive distance *between* clusters. This is not to be avoided nor regretted, because the cluster-specific institutions are often a major prerequisite for making the firms of the cluster attractive for outsiders to interact with.

However, there is a downside that becomes apparent when a successful fit creates routines of extraordinary durability: they are retained and sometimes even aggressively defended long after changes within or outside the cluster have made them redundant (Demsetz, 1988). It is difficult to *unlearn* successful habits of the past, also in cases where it is obvious to everyone concerned that they hinder future knowledge creation and success (Imai *et al.*, 1986, Hedberg, 1981).

In the cluster, as elsewhere in the economy, wanting unlearning might go hand in hand with an increasing resistance towards new ideas, a growing bureaucratic inertia and a general organizational degeneration, especially when the firms are operating in generous markets (Eliasson, 1996). Experienced success results in a flatter forgetting curve, and accepted best practices assume a life of their own (Hamel and Prahalad, 1994). It is an established fact of life that it is a lot easier to challenge the orthodoxy of others compared with one's own, and firms — sometimes the whole cluster — are sometimes led by their former success into trajectory-specific lock-ins (David, 1985, Arthur, 1989). Or in the words of Boisot:

> ... experiences work their way into the collective memory and expectations of a culture and remain embodied in institutional arrangements long after they have ceased to serve. They may then obstruct rather than assist the process of social adoption much as early childhood traumas become the source of phobias and pathologies in later adult life (Boisot, 1983: 160).

Institutions are linked to social history and they might be slow to adapt to the change in the related industries of the cluster. Therefore no one-to-one correlation between industrial structure and institutional endowment should automatically be expected, and empirical studies continuously add to our understanding of why a complete match is seldom found and how different degrees of mismatches affect the performance of the cluster. As we see it, the "institutional fit" is not in itself part of the explanation of the existence of the cluster. Rather, it contributes to the explanation of the successful path-dependent development trajectories of clusters – and to the lock-in situations in which clusters sometimes end up.

The three elements of the theory presented here – variation, vertical differentiation, and institutional fit – may show up in different ways in empirical assessments, and their relative impact may vary over time. One would expect variation would be a key feature in the early development of cluster, while differentiation and institutional adjustment would play increasingly important roles as time goes by.

5. Conclusion: where do we go from here?

The research problem attached to the agglomeration phenomenon may seem to be trivial. We have a phenomenon that can be observed in reality. Firms in the same or related industries often—although far from always—tend to locate at the same place, and this should reasonably indicate that there should be some advantages connected to such a location pattern: localized clusters do exist and this may legitimately make us assume that such a spatial structure is in some sense efficient or rational. This point of departure is common to many scholars, not only in economic geography but also in economics and business studies: "A cluster of independent and informally linked companies and institutions represents a robust organizational form that offers advantages in efficiency, effectiveness, and flexibility" (Porter 1998: 80). At the same time, it has turned out to be extremely difficult to identify empirically the mechanisms that are supposed to account for its existence.

One problem relates to the issue of spatial scale. The notions *local* and *regional*, which are often central in analyses of spatial clustering, are extremely elastic. First, the two notions are often used more or less as synonyms in the literature. Furthermore, they may denote a number of geographical scales, extending from the local neighbourhood (a street or block in a city, or a small town) through to entire nations or even groups of nations. Similar mechanisms or forces are held to explain both why advertising agencies flock together at a particular street (Maddison Avenue) in New York, and why the 'European banana' developed as a core area of heavy industrialization during the 19th century, an area which is extended across several countries in the heartland of what is now the European Union. It does not, however, seem possible to define, once and for all, a specific geographical scale at which one could argue that agglomeration economies exert a particularly strong influence. Rather, it seems reasonable to allow the scale to vary according to which type of phenomenon that is emphasized in the analysis.

Thus, if we focus primarily on the role formal institutions (such as the legal system) or cultural and linguistic aspects, the nation state may often work as a good proxy for 'the local milieu'. To the degree that the agglomeration economies work in a dichotomous way, such that they render advantages in the exchange between insiders in a milieu ('us') and in a corresponding way complicate the exchange (of goods, services and/or information) with outsiders ('them'), nation states are often the relevant scale. If, however, we primarily focus on business transactions between related firms, the geographical scale will be more of a continuum. In principle, one may assume than the costs of interaction will generally increase with increasing distance such that the agglomeration economies are maximally forceful in the case of the immediate juxtaposition of two trading partners, while it will gradually decrease as distance grows. When, on the other hand, the primary focus of the analysis is the mundane, everyday, often informal exchange of information and ideas, associated with the frequent face-to-face contacts and more or less unplanned meetings which are seen to forward the creation of new ideas and innovation, the relevant spatial resolution will turn out to be fine. Normally, we will then focus on small places, or limited parts of larger functional regions (urban neighbourhoods, city blocks).

It does at present not seem reasonable to attach the use of the concept localization economies to one particular spatial scale, even though the theoretical approach proposed in this paper would tilt our focus towards the smaller scale. Nevertheless, the conceptual elasticity indicated above, and the fact that individual theoretical accounts often lack explicit discussion of the issue of spatial scale and the interaction between various levels, is somewhat troublesome.

Another, and perhaps more significant issue is the problems of measurability. While it is relatively uncomplicated to asses empirically to what degree firms in a particular industry are agglomerated, it is much more difficult to investigate the degree of agglomeration across groups of firms which are related along some other dimensions (through business transaction, shared technology etc). As soon as one leaves the industry classification of official industrial statistics, the possibility of using existing data diminishes rapidly: "Clusters rarely conform to standard industrial classification systems, which fail to capture many important actors and relationships in competition. Thus significant clusters may be obscured or even go unrecognized" (Porter 1998: 79).[10]

It is also comparatively easier to map patterns of inter-firm linkages in terms of flows of commodities or money between firms, than it is to capture flows of information and knowledge. The studies of contact patterns carried out by some economic geographers in the late 1960s and early 1970s[11] meant that some steps were taken along such a path but they have not been followed up since. Finally, when it comes to the institutional, social and cultural milieu, some characteristics are indeed possible to map (see Putnam 1993, for an interesting although controversial attempt to measure the degree of 'civicness' in Italian regions), but it is evidently extremely difficult to study in a way that allows for systematic comparison across regions, phenomena which are stated to be 'in the air', which was the phrase Marshall (1890) used.

The problem is to come up with a model that can account for agglomeration phenomena without ending up in circular reasoning. Models of agglomeration have always run such a risk: the existence of agglomerations is explained by

---

[10] It should be noted that in reality, even to assess what is an agglomerated location pattern in a properly defined group of forms (e.g. belonging to an industry according to standard classifications like ISIC) is not that simple. Ellison and Glaeser (1994) show that even pure chance can produce agglomeration: They illustrate this by showing that one only needs to through six darts at a map of the U.S. before it is most likely that two will hit the same state. Thus, when concluding that a certain spatial distribution, e.g. a certain value of a Gini or localization index, is indicating that we are faced with agglomeration, we should be sure that we are actually observing 'more than random' agglomeration.

[11] See Törnqvist (1970) for a summary and overview.

agglomeration economies. These agglomeration economies can not really be specified or observed but their existence is proved by the existence of agglomerations (cf. Krugman 1995: 52). Such an approach is, of course, less than totally satisfactory.

Furthermore, there is an urgent need to increase out knowledge of some other possible trade-offs that are often, explicitly or implicitly, advanced in existing accounts. Thus, it seems obvious that firms need a supporting structure of trustful and collaborative actors in order to perform well. At the same time, however, there is a need for competition and rivalry to drive the innovation process, defined in a broad sense. In most existing accounts of the agglomeration phenomenon, the collaborative element is emphasized at the expense of competitive. In the present article a different approach is suggested emphasizing the possible learning effects along the horizontal dimension of the cluster.

We suspect that the reason for the relatively meager results coming out of studies attempting to study empirically the magnitude and intensity of local inter-firm collaboration may be that it is on the horizontal dimension of inter-firm observation, comparison and competition that the effects of agglomeration are most important. Thus, a 'nice' and collaborative atmosphere might not at all characterize most firms in a spatial agglomeration. Firms may dislike each other and refuse to talk to each other but can still, indirectly, contribute to each other's competitive success in the global market. To the degree that this is the case, there are reasons to be cautious against some of the policy initiatives which, with theoretical inspiration from some of the literature discussed here, are exclusively geared at promoting the establishment of local networking and inter-firm collaboration (cf. Maskell 2001a).

At the start of the paper we argued that research on spatial clustering is at the core of economic geography, highlighting as it does the impact of proximity and distance, institutions and local milieus on economic processes. We have subsequently argued that a theory of spatial clustering should play down the role of cost efficiencies in favour of focussing on the way clustering enhance knowledge creation. In that context we have argued that spatial clusters does induce variation, observability and comparability, while at the same time allowing increased differentiation without discouraging knowledge exchange by imposing too large cognitive distance. We suggest that rather than using the costs induced by 'friction of distance' as the starting point, the spatial attributes of interactive learning and innovation process would perhaps be a fruitful point of departure not only when analyzing spatial agglomerations, but also when it comes to re-invigorating research in economic geography generally.

Acknowledgements
Anders Malmberg made some of the work for this article while being a Fellow at The Swedish Collegium for Advanced Study in the Social Sciences in 1999. The participants at The World Conference on Economic Geography in Singapore, December 4-7, 2000, have provided valuable comments on an earlier draft. We thank The Nordic Center for Spatial Development (NORDREGIO) for supporting financially the research on which this article is based. The usual disclaimers apply.

References

Allen, R.C. (1983): Collective invention. *Journal of Economic Behaviour and Organization* **4**: 1-24.

Althin, T. (in collaboration with H. Nelson) (1948): Industrilokalisering. In: *Sveriges Industri*, tredje upplagan. Stockholm: Industriförbundet.

Alvstam, C. (1998): Ekonomisk geografi – ett ämne borttappat och återfunnet. In: Gren, M. & Hallin, P.O. (red) *Svensk kulturgeografi. En exkursion inför 2000-talet.* Lund: Studentlitteratur.

Amin, A, and Thrift (1992) Neo-Marshallian nodes in global networks, *International Journal of Urban and Regional Research,* **16**: 571-87.

Amin, A, and Thrift (1994) 'Living in the global', in: Amin, A. & N. Thrift (eds) *Globalization, institutions and regional development in Europe*, Oxford: Oxford University Press.

Angel, D.P. and Engstrom, J. (1995): Manufacturing systems and technological change: The U.S. personal computer industry. *Economic Geography,* **71**: 79-102.

Appold, S.J. (1995): Agglomeration, interorganizational networks, and competitive performance in the U.S. metalworking sector. *Economic Geography,* **71**: 27-54.

Archibugi, D. and Pianta, M. (1992): *The Technological Specialization of Advanced Countries*, Dordrecht: Kluwer.

Arthur, W.B. (1989): Competing Technologies, Increasing Returns, and Lock-ins by Historical Events. *Economic Journal* **99**: 116-131.

Asheim, B.T. and Isaksen, A. (1996): Location, agglomeration and innovation: towards regional innovation systems in Norway? Paper presented at the 28th International geographical Congress, The Hague, The Netherlands, 5-10 August.

Baptista, R. (1998): 'Clusters, Innovation and growth: A survey of the Literature'. In: Swann, P.G.M., Prevezer, M. and Stout, D., (Eds.) *The dynamics of industrial clustering. International comparisons in computing and biotechnology* pp. 13-51. Oxford: Oxford University Press.

Becattini, G. (1990): 'The Marshallian industrial districts as a socio--economic notion'. In: Pyke, F., (Ed.) *Industrial Districts and Inter-Firm Co-Operation in Italy* Geneva: International Institute for Labour Studies.

Bellandi, M. (1989): 'The industrial district in Marshall'. In: Goodman, E. and Bamford, J., (Eds.) *Small firms and industrial districts in Italy* pp. 136-152. London.

Bellandi, M. (1996): Innovation and Change in the Marshallian Industrial District. *European Planning Studies* **4**: 357-368.

Bianchi, G. (1998): Requiem for the Third Italy? Rise and fall of a too successful concept. *Entrepreneurship and Regional Development* **10**: 93-116.

Blanc, H. and Sierra, C. (1999): The internationalisation of RandD by multinationals: a trade-off between external and internal proximity. *Cambridge Journal of Economics* **23**: 187-206.

Boeri, T. (1989): Does firm size matter? *Giornale degli Economisti e Annali di Economia* **48**: 477-495.

Boisot, M., Boldrin, M. and Scheinkman, J.A. (1983): Convergence revisited: the knowledge of diffusion of knowledge in a British and a Japanese firm. *Journal of Management Studies* **1**: 159-190.

Brusco, S. (1982): The Emilian model: Productive decentralisation and social integration. *Cambridge Journal of Economics* **6**: 167-184.

Camagni, R.P. (1995): The concept of "innovative milieu" and its relevance for public policies in European lagging regions. *Papers in Regional Science* **74**: 317-340.

Carter, A.P. (1994): *Measuring the Performance of a Knowledge-based Economy.* Working Paper no.:337, Brandais Brandais University, Departments of Economics

Casson, M. (1982): *The entrepreneur. An economic theory*, Oxford: Martin Robertson.

Chinitz, B. (1961): Contrasts in Agglomeration: New York and Pittsburgh. *American Economic Review* **LI**: 279-289.

Cohen, S.S. and Fields, G. (1999) *Social capital and capital gains, or virtual bowling in Silicon Valley. An examination of Social Capital in Silicon Valley.* BRIE Working Paper, Berkeley University.

Cooke, P. (1996): Regional innovation system: an evolutionary approach, in: H. Braczyk, P. Cooke and R. Heidenreich (Eds.) *Regional innovation systems. London: UCL Press.*

David, P.A. (1985): Clio and the economics of QWERTY. *American Economic Review* **75**: 332-337.

DeBresson, C. (1989): Breeding Innovation Clusters: A Source of Dynamic Development. *World Development* **17** (1): 1-16.

Dei Ottati, G. (1994): Co-operation and competition in the industrial district as an organisational model. *European Planning Studies* **2**: 463-483.

Demsetz, H. (1988): *Ownership, Control and the Firm*, Oxford: Blackwell.

Dicken, P. and Lloyd, P.E. (1990): *Location in space: Theoretical perspectives in economic geography. Third edition.* New York: Harper and Row.

Eliasson, G. (1996): *Firm Objectives, Controls and Organization. The Use of Information and the Transfer of Knowledge within the Firm*, Dordrecht: Kluwer Academic Publishers.

Ellison, G. and Glaeser, E.L. (1994): *Geographical Concentration in the US Manufacturing Industries. A Dartboard Approach.* Working Paper no.:4840, Cambridge, Mass. National Bureau of Economic Research (NBER)

Enright, M.J. (1998): Regional clusters and firm strategy. In: Chandler, A.D., Hagström, P. and Sölvell, Ö. (eds) *The dynamic firm. The role of technology, strategy, organization and regions.* Oxford: Oxford University Press.

Estall, R.C. and Buchanan, R.O. (1961): *Industrial Activity and Economic Geography*, London: Hutchinson.

Freeman, C. (1982): *The economics of industrial innovation*, London: Pinter Publishers.

Freeman, C. (1991): Networks of Innovators: a Synthesis of Research Issues. *Research Policy* **20** (5).

Fuellhart, K (1999) Lozalization and the use of information sources: the case of the carpet industry. *European Urban and Regional Studies* **6**: 39-58.

Gottardi, G. (1996): Technology strategies, Innovation without RandD and the creation of knowledge within industrial districts. *Journal of Industry Studies* **3**: 119-134.

Greenhut, M.L. (1970): *A theory of the firm in economic space*, New York: Appleton-Century-Crofts.

Hagedoorn, J. and Schakenraad, J. (1992): Leading companies and networks of strategic alliances in information technologies. *Research Policy* (21).

Håkansson, H., (Ed.) (1987): *Industrial Technology Development - a Network Approach.* London: Croom Helm.

Hamel, G. and Prahalad, C.K. (1994): *Competing for the Future*, Boston: Harvard Business School Press.

Hanson, G.H. (2000): 'Firms, Workers, and the Geographic Concentration of Economic Activity' Chapter 24 in Clark, L.G., Feldman, M.P. and Gertler, M.S.(eds): *The Oxford Handbook of Economic Geography* Oxford: Oxford University Press:

Harrison, B. (1992): Industrial Districts. Old wine in New Bottles? *Regional Studies* **26**: 469-483.

Harrison, B., Kelley, M.R. and Gant, J. (1996): Innovative firm behavior and local milieu: Exploring the intersection of agglomeration, firm effects, and technological change. *Economic Geography* **72**: 233-258.

Head, K., Ries, J. and Swenson, D. (1995): Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. *Journal of International economics* **38**: 223-247.

Hedberg, B. (1981): 'How organizations learn and unlearn'. In: Nystrøm, P.C. and Starbuck, W.H., (Eds.) *Handbook on Organizational Design - Adapting Organizations to their Environment* pp. 3-27. Oxford: Oxford University Press.

Herrigel, G. (1996): Crisis in German decentralized production: unexpected rigidity and the challenge of an alternative form of flexible organization in Baden-Württemberg, *European Urban and Regional Studies* **3**: 33-52.

Hirschman, A.O. (1958): *The strategy of economic development*, Clinton Mass.: Yale University Press.

Hoover, E.M. (1937): *Location theory and the shoe and leather industries*, Cambridge, Mass.: Harvard University Press.

Hoover, E.M. (1948): *The location of Economic activity*, New York: McGraw-Hill Book Company.

Hudson, R. (1999) "The learning economy, the learning firm and the learning region": a sympathetic Critique of the limits to learning, *European Urban and Regional Studies* **6**: 59-72.

Imai, K., Ikujiro, N. and Hirotaka, T. (1986): 'Managing the new product development process: Companies learn and unlearn.'. In: Clark, K.B., Hayes, R.H. and Lorenz, C., (Eds.) *The uneasy alliance. Managing the productivity-technology dilemma* Boston.: Harvard Business School Press.

Isaksen, A. (1996): Towards increased regional specialisation? The quantitative importance of new industrial spaces in Norway, 1970-1990, *Norsk Geografisk Tidsskrift* **50**: 113-123.

Jacobs, J. (1961): *The death and life of great American cities*, New York: Random House.

Johnson, B. (1992): 'Institutional learning'. In: Lundvall, B.-Å., (Ed.) *National systems of innovation* London: Pinter.

Kahlin, T. (1961): Industrilokalisering. I: *Sveriges Industri*, fjärde upplagan. Stockholm: Industriförbundet.

Kirat, T. and Lung, Y. (1999): Innovation and Proximity: Territories as Loci of Collective Learning Processes. *European Urban and Regional Studies* **6**: 27-38.

Kline, S.J. and Rosenberg, N. (1986): 'An overview of innovation'. In: Landau, R. and Rosenberg, N., (Eds.) *The positive sum game* Washington D.C.: National Academy Press.

Krugman, P. (1991a): *Geography and trade.* Cambridge, Mass: MIT Press.

Krugman, P. (1991b): Increasing returns and economic geography. *Journal of Political Economy* **99**: 483-499.

Krugman, P. (1991c): History and Industry Location: The Case of the Manufacturing Belt. *The American Economic Review* **81**: 80-83.

Krugman, P.R. (1995): *Development, Geography, and Economic Theory*, Cambridge, Mass.: The MIT Press.

Larsson, S. (1998): Lokal förankring och global räckvidd. En studie av teknikutveckling i svensk maskinindustri, *Geografiska regionstudier nr. 35,* Uppsala: Kulturgeografiska institutionen, Uppsala universitet.

Larsson, S. and Lundmark, M. (1991): Kista – företag i nätverk eller statusadress? En studie av Kistaföretagens länkningar. *Forskningsrapport nr. 100,* Uppsala: Kulturgeografiska institutionen, Uppsala universitet.

Larsson, S. and Malmberg, A. (1999) Innovations, competitiveness and local embeddedness. a study of machinery producers in Sweden, *Geografiska Annaler Series B Human Geography* **81B**: 1-18.
List, F. (1841): *The National System of Political Economy*, London: Longmans, Green and Co.
Loasby, B.J. (1999): 'Industrial districts as knowledge communities'. In: Bellet, M. and L'Harmet, C., (Eds.) *Industry, space and competition. The contribution of economistc of the past* pp. 70-85. Cheltenham: Edward Elgar.
Loasby, B.J. (2000) *Organisations as interpretative systems*. Paper presented at the DRUID summer conference, Rebild, Denmark, June.
Lundquist, K.-J. (1996): Företag, regioner och internationell konkurrens. Om regionala resursers betydelse, *Meddelanden från Lunds universitets geografiska institutioner, Avhandlingar nr. 129.* Lund: Lund University Press.
Lundquist, Karl-Johan and Lars-Olof Olander (1998), *Regional economies: a threat to the nation-state?* (CFE Working Paper, 3. Lund: Centre for European Studies, Lund University).
Lundvall, B.-Å., (Ed.) (1992): *National systems of innovation: Towards a Theory of Innovation and Interactive Learning.* London: Pinter.
Lundvall, B.-Å. and Maskell, P. (2000): 'Nation states and economic development - From national systems of production to national systems of knowledge creation and learning'. In: Clark, G.L., Feldmann, M.P. and Gertler, M.S., (Eds.) *The Oxford Handbook of Economic Geography* Oxford: Oxford University Press.
Maillat, D. (1991): 'Local dynamism, milieu and innovative enterprises'. In: Brotchie, J., Batty, M., Hall, P. and Newton, P., (Eds.) *Cities of the 21st century* London: Longman.
Maillat, D. (1998): Innovative milieux and new generations of regional policies. *Entrepreneurship and Regional Development* **10**: 1-16.
Malerba, F. (1989) *Learning by firms and incremental technical change* Working paper, Università degli Studi, Brescia.
Malmberg, A. (1996): Industrial geography. Agglomerations and local milieu. *Progress in Human Geography* **20**: 392-403.
Malmberg, A. (1997): Industrial Geography: location and learning. *Progress in Human Geography* **21**: 573-582.
Malmberg, A., Malmberg, B. and Lundequist, P. (2000): Agglomeration and firm performance: Economies of scale, localisation and urbanisation among Swedish export firms. *Environment and Planning A* **32**: 305-321.
Malmberg, A., Sölvell, Ö and Zander, I. (1996): Spatial clustering, local accumulation of knowledge and firm competitiveness, *Geografiska Annaler* **78B**: 85-97.
Malmberg, A. and Maskell, P. (1997): Towards and explanation of regional specialization and industry agglomeration. *European Planning Studies* **5**: 25-41.
Markusen, A. 1999. Fuzzy concepts, scanty evidence, policy distance: the case for rigour and policy relevance in critical regional studies. *Regional Studies* **9**: 869-884.
Markusen, A., Hall, P. and Glasmeier, A.K. (1986): *High Tech America*, Boston: Allen and Unwin.
Marshall, A. (1890): *Principles of Economics*, London: MacMillan. (8th edition 1920).
Maskell, P. (1998): Successful low-tech industries in high-cost environments: The case of the Danish furniture industry. *European Urban and Regional Studies* **5**: 99-118.
Maskell, P. and Malmberg, A. (1999): The Competitiveness of Firms and Regions: 'Ubiquitification' and the importance of Localised Learning. *European Urban and Regional Studies* **6** (1): 9-26.Maskell, P. and Malmberg, A. (1999): Localised learning and industrial competitiveness. *Cambridge Journal of Economics* **23** (2): 167-186.
Maskell, P. (2000): 'Social capital and competitiveness'. Chapter 6 in: Baron, S., Field, J. and Schuller, T., (Eds.) *Social Capital. Critical Perspectives.* Oxford: Oxford University Press.
Maskell, P. (2001a): 'Regional Policies: Promoting Competitiveness in the Wake of Globalisation'. In: Felsenstein, D. and Taylor, M., (Eds.) *Promoting Local Growth: Process, Practice and Policy*: Ashgate.
Maskell, P. (2001b): 'Towards a learning-based theory of the cluster', *Industrial and Corporate Change* (forthcoming).
Maskell, P. , Eskelinen, H., Hannibalsson, I. Malmberg, A. and Vatne. E. (1998): *Competitiveness, localised learning and regional development. Specialisation and prosperity in small open economies.* London: Routledge.
Maskell, P. and Törnqvist, G. (1999): *Building a Cross-Border Learning Region. The Emergence of the Northern European Øresund Region*, Copenhagen: Copenhagen Business School Press.
McCann, P. (1995): Rethinking the economics of location and agglomeration. *Urban Studies* **32**: 563-577.
Morgan, K. (1997): The Learning Region: Institutions, Innovation and Regional Renewal, *Regional Studies* **31**:491–504.
Myrdal, G. (1957): *Economic Theory and the Underdeveloped Regions*, London: Ducksworth.
Norton, R.D. (1992): Agglomeration and Competitiveness: From Marshall to Chinitz. *Urban Studies* **29**: 155-170.

OECD (1992): *Industrial Policy in the OECD countries.* Annual Review Paris.: The Organisation for Economic Co-Operation and Development
Porter, M.E. (1990): *The competitive advantage of nations.* London and Basingstoke: Macmillan.
Porter, M. E. (1994): The role of location in competition. *Journal of the Economics of Business* **1**: 35-39.
Porter, M. E. (1998) Clusters and the new economics of competition. *Harvard Business Review,* November/December: 77-90.
Pred, A. (1966): *The spatial dynamics of U.S. urban-industrial growth, 1800-1914. Interpretive and theoretical essays.* Cambridge, Mass. and London: The Harvard University Press.
Pred, A. (1977): *City Systems in Advanced Economies. Past Growth, Present Processes and Future Development Options*, London: Hutchinson.
Putnam, R. D. (with Leonardi, R. and Nanetti, R. Y.) (1993): *Making Democracy Work. Civic Traditions in Modern Italy*, Princeton NJ: Princeton University Press.
Ratti, R., Bramanti, A. and Gordon, R., (Eds.) (1998): *The Dynamics of Innovative Regions. The GREMI Approach.* Aldershot: Ashgate.
Rosenberg, N. (1972): *Technology and American Economic Growth*, White plains, N.Y.: Sharpe.
Ross, E.A. (1896): The location of industries, *Quarterly Journal of Economics* **10**: 247-268.
Sabel, C.F. (1989): Flexible specialisation and the re-emergence of regional economies, in: P. Hirst and J. Zeitlin (eds) *Reversing industrial decline? Industrial structure and policy in Britain and her competitors.* Oxford: Berg Publishers Limited.
Saxenian, A. (1994): *Regional Advantage. Culture and competition in Silicon Valley and Route 128*, Cambridge, Massachusetts.: Harvard University Press.
Scott, A.J. (1983): Industrial organisation and the logic of intra-metropolitan location – 1. Theoretical considerations. *Economic Geography* **59**: 233-250.
Scott, A.J. (1988): New industrial spaces: Flexible production organisation and regional development in North America and Western Europe. London: Pion.
Scott, A.J. (1998): 'From Silicon Valley to Hollywood. Growth and development of the multimedia industry in California'. In: Braczyk, H.J., Cooke, P. and Heidenreich, M., (Eds.) *Regional innovation Systems.The role of governances in a globalized world* pp. 136-162. New York: UCL-Press.
Schumpeter, J. 1934 (1959). The Theory of Economic development: An inquiry into Profits, Capital, Credit, Interest, and the Business Cycle. Cambridge MA: Harvard University Press.
Spender, J.-C. (1994): 'The Geographies of Strategic Competence: Borrowing from Social and Educational Psychology to Sketch an Activity and Knowledge-Based Theory of the Firm'. In: Chandler, A.D., Hagström, P. and Sölvell, Ö., (Eds.) *The Dynamic Firm. The Role of Technology, Strategy, Organizations, and Regions.* pp. 417-439. Oxford: Oxford University Press.
Staber, U. (1996): Accounting for variations in the performance of industrial districts: The case of Baden-Württemberg, *International Journal of Urban and Regional Research* **20**: 299-316.
Staber, U. (1997): An ecological perspective on entrepreneurship in industrial districts, *Entrepreneurship and Regional Development* **9**: 45-64.
Storper, M. (1993): Regional worlds of production: Learning and innovation in the technology districts of France, Italy and the USA. *Regional Studies* **27**: 433-455.
Storper, M. (1995): The resurgence of regional economies, ten years later: the region as a nexus of untraded interdependencies. *European Urban and Regional Studies* **3**: 191-221.
Storper, M. 1997. *The Regional World*. New York: Guilford.
Swann, P.G.M., Prevezer, M. and Stout, D. (1998): *The dynamics of industrial clustering. International comparisons in computing and biotechnology*, Oxford: Oxford University Press.
Teece, D.J. (1980): Economies of scope and the scope of the enterprise. *Journal of Economic Behavior and Organization* **3** (1): 223-247.
Törnqvist, G. (1970): Contact systems and regional development. *Lund Studies in Geography, Ser. B. Human Geography,* No. 35.
Ullman, E.L. (1958): Regional development and the geography of concentration. *Papers of the Regional Science Association* **IV**: 179-198.
von Hayek, F.A. (1937): Economics and knowledge. *Economica* **IV** (New series): 33-54.
Weber, A. (1909/1929): *Theory of the location of industries.* Chicago: The University of Chicago Press.
White, H.C. (1981): Where do markets come from? *American Journal of Sociology* **87**: 517-547.
Wiig, H. and Wood, M. (1995): What comprises a regional innovation system? An empirical survey. STEP Report 1-95, Oslo: STEP Group.
Winter, L. (1997): Regional teknisk forandring og stiafhængig udvikling. dansk industris regionale udvikling 1980-1994. PdD.-afhandling, Geografisk Institut, Københavns Universitet.

Yeung, H.W. (2000): Organising "the Firm" in Industrial Geography I: Networks, institutions and Regional Development. *Progress in Human Geography* **24**: 301-315.
Young, A. (1928): Increasing returns and economic progress. *Economic Journal* **38**: 527-542.