Mysteries of the Region:
Knowledge Dynamics in Silicon Valley

John Seely Brown

Paul Duguid

Abstract

For more than a century, commentators have predicted that industrial "clusters" would disappear with improvements in communications technology. Yet still clusters form, most noticeably in Silicon Valley, at the heart of the latest revolution in information technology, the Internet. How can such clusters be explained? We argue that collaborative practices and their contribution to the dynamics of knowledge are more important than communications technologies in explaining clusters. We advance a unitary, practice-based, account of both internal and the external flows of organizational knowledge. We then use this account to explain the importance of proximity to the movement of knowledge and hence the advantages of clustering, despite the disadvantages that come from losing knowledge to competitors. We sketch a matrix view of clusters and suggest that the details of the matrix explain why networks of knowledge have variable density and topography. In conclusion, we replace this simple matrix view with a richer ecological view of knowledge in regions. This helps explain both why regions are so strong and why they are so difficult to imitate.

* * *

The mysteries of the trade become no mysteries; but are as it were in the air.

Alfred Marshall
Principles of Economics, 1890

Been there, done that?

Despite all the recent insightful writings on "clusters" (Porter, 1998), "technopoles" and "innovative milieux" (Castells and Hall, 1994), and "regional advantage" (Saxenian, 1996), it can feel as though researchers are only adding footnotes to Alfred Marshall's magisterial economic exploration of "localization," written more than a century ago.

For example, Saxenian's study importantly shifts attention from economic issues to social and cultural ones. That shift recalls Marshall's insight that in localization "social forces co-operate with economic ones." For their part, Castells and Hall, reflecting Sturgeon's analysis (1988, 2000), illuminate the historical determinants of regions. Silicon Valley, they note, goes back much further than the silicon chip or even the fabled garage of Hewlett and Packard. Its roots extend at least to the development of radio technologies before the First World War. Marshall would no doubt have approved such historicizing. He traced the roots of the south Lancashire steel industry in the nineteenth century, for example, to the settlement of smiths after the Norman conquest in the eleventh century.

Again, recent analyses have noted the importance of the lawyers, venture capitalists, marketing firms, and the like that have grown to support the core firms of Silicon Valley. Kenney and von Burg (2000) insightfully refers to this as the "second economy"; Lynn et al. (1996) as the "superstructure." Both these approaches call to mind Marshall's account of the "subsidiary" trades that develop around the primary firms of an industrial localization.

Finally, analysts increasingly argue that the conventional dichotomy of firm and market inherent in the "transaction cost" view of business organization is too stark to explain something like Silicon Valley. With its complex array of subcontracts, cross-licenses, joint ventures, and so on, much of Silicon Valley falls organizationally into what Hennart (1993) calls the "swollen middle" between firm and market. Indeed, Gilson (1996) provocatively asks, "Is Silicon Valley--a complex network of shifting alliances, relationships, intermediaries, firms and investors--in fact a firm?" The heart of Marshall's analysis, of course, concerns the way that industrial localization has collective, systemic properties that fall somewhere between market and firm.[1]

Limits to prophecy

In all, Marshall was remarkably insightful and "foresightful," presenting and predicting at the end of the nineteenth century many of the characteristics now associated with Silicon Valley at the end of the twentieth. Given his prescience, and given how hard prediction of any sort is, it will seem churlish to concentrate on one point where he seems less clear and accurate about future trends. Great thinkers, however, are often as interesting when they get things wrong as when they get them right. Moreover, Marshall treads with uncharacteristic timidity in exactly the spot where many people since (most of whom have probably never heard of Marshall) rush in with confident predictions. Marshall's hesitancy and the limitations of these predictions, we suggest, should prompt us all to rethink many of the explanations of, to put it briefly, why clusters cluster.

Standing on his robust analysis of localization and looking to the future, a future in which we now live, Marshall notes, "Every cheapening of the means of communication alters the action of the forces that tend to localize industries" (Marshall, 1916, 273) The phrase suggests that clusters are held together by the limits of communications technologies. As these technologies develop, Marshall seems to believe, the need for clusters should fall away. Yet contemplating better communications (and lower tariffs), Marshall seems unable to discern, path dependency aside, any future logic to the distribution of industry. On the one hand, he suggests, some industries may continue to cluster. On the other, they may spread apart. He points to "two opposing tendencies," but offers no clear idea of what drives either. What Marshall seems to lack, and what we want to offer here, is a diagnostic tool for understanding how these opposing forces, one centripetal, the other centrifugal, will resolve themselves.

More recently, theorists have been far less hesitant than Marshall. They speak clearly and confidently of, for instance, the "death of distance" (Cairncross, 1997) or what Castells (1996) calls the "space of flows." A society "organized around flows of information," Castells argues, will create a "new spatial logic." Arguments like these tend to focus on Marshall's centrifugal forces and to assume, as he did, that with the improvement of communications, the centripetal forces of localization will wither away.

But before we accept the compelling logic of such theoretical arguments, it's important not to ignore the stubbornness of history. For since Marshall wrote the technological means of communication have cheapened and changed beyond wildest expectations. In 1890 (when the first edition of *The Principles of Economics* appeared), telephone was a fledgling technology. Radio was barely a dream. Television unthought of. Computers and the Internet almost completely unthinkable. Since Marshall's prediction all these and more means of communication (from the fax to Federal Express) have fallen into place and fallen in price. Yet clusters continue to exist as robust and important social and economic entities. So while "disagglomeration" is undoubtedly an industrial reality, in some industries and some regions dramatic cheapening does not appear to have significantly altered "the action of the forces" that keep competitors geographically local. Hence "regional advantage" and "clustering" are still salient issues, politically and economically, today.

Of course, it may seem reasonable to argue that what clustering exists today is no more than a path-dependent legacy of old industrial organization. Yet that argument must confront the evidence that the cluster that is probably the best known and economically most significant, the one societies around the world most want to imitate, actually lies not in the remnants of the industrial age, but on the cutting edge of the information age and communication technologies. Not only did Silicon Valley initially develop around early work in radio technology, but in the last decade it has reinvented itself around the technologies of networking and the Internet. In so doing, the region seems only to have reinforced its centripetal attraction. And while it inevitably spawns imitators around the world, it makes the formation of rival regions difficult by continuing to draw the best scientists and engineers from across the globe, with what Marshall called its "constant market for skill" (Marshall, 1916, 271). Even organizations at the heart of other clusters succumb to its magnetism. Thus in 1998, Microsoft, though itself a pole for localization in Washington state and committed to the distal power of the Internet, moved part of its research arm to the Valley to take advantage of Marshall's mysteries in the air. It would seem that as Silicon Valley has cheapened the "means of communication" and improved its capacity to disaggregate, the "forces that tend to localize industries" have, in its case, paradoxically only intensified.

The critical questions, then, are *why* and *where*. Why, despite a century of predictions to the contrary, and the development of technologies for overcoming distance, does localization remain and even thrive? And where, given the significant forces of disagglomeration, might we expect localization to continue? Exploring these questions throws light on a number of critical issues, as we try to show in this chapter. Overall, the tenacity of Marshall's localization suggests how little is still understood about the character of the local and the importance of direct human interaction. Particularly where knowledge is a critical factor, these need to be taken seriously. Usually, they are taken for granted.

So, taking Marshall's quotation about mysteries in the air as our starting point, we argue first for the importance of proximity and, in particular, of collaborative practices to the flows and dynamics of knowledge. Next, we suggest that the challenge of moving knowledge helps explain the function and internal structure of the firm. Economics and sociology often turn the inside of the firm into an almost magical "black box" where communication is "costless" or otherwise unproblematic. Only by understanding the difficulty of moving knowledge can we really understand the challenge of innovation and the contribution of the firm to the innovative process. We then go on to claim that the same explanation that

accounts for the internal relations of the firm in terms of knowledge dynamics also explains many of its external relations to other firms.

Explaining internal and external relations in terms of knowledge leads us to lay out the region as a matrix structure made up of knowledge flows, some that run within firms, some that run between them. Conceptually, this matrix appears to extend indefinitely into a national or even global network. It soon becomes clear, however, that the overall network has a topography that can vary significantly. Here, the simple lines of the matrix give way to multi-directional flows, complex feedback loops, structural forms of amplification, and intricate relations of interdependence. At this point in our argument, despite the ease and elegance of matrix or network models, it seems more useful to think in terms of ecologies. And the ecological model of regions then answers the question that seems always to be in the air when people talk about regions: Why are regions so hard to replicate?

Mysteries in the air

In the air

The idea that critical organizational knowledge that would usually be secret and hard to find is widely known in clusters is familiar to almost anyone who has lived or worked in Silicon Valley. (Indeed, Marshall's phrase about mysteries in the air turns up repeatedly in articles about the Valley.) As Saxenian (1996) has shown, there is an extraordinary level of knowledge available to all who work in the Valley.

To understand exactly what this means, let us make a loose distinction. Undoubtedly there is a high level of knowledge *in* firms in the Valley. These firms thrive on their proprietary knowledge. But so do all hi-tech firms, whether in clusters or in isolation. In the Valley, however, there's also a very high level of knowledge *about* the firms. People are remarkably well informed about what their competitors are up to. It's not hard to find out who's good at a particular task (and who's not), who's reliable (and who's not), and so forth. This ambient knowledge helps people understand almost implicitly what's old and so not worth pursuing, and what's new and worth a second look. Inevitably, much of this is also evident from outside the Valley. What seems to be less evident from outside is any idea of what's missing or what's coming: where the new opportunities, the "next new thing," is likely to come from.

As an example, consider 3Com, which through its work on Ethernet technologies developed as a networking firm. Initially, its obvious market was minicomputers. These, several firms realized, badly needed good networking capabilities. But, as Kenney and von Burg (1999) argue, 3Com's position in the Valley allowed the firm to see a little further. Its managers understood that however lucrative they might appear in the short term, minicomputers formed a soon-to-be-dying market. Thus the company bypassed this obvious market, while its competitors continued to battle for it. Instead, 3Com positioned itself for a barely visible but ultimately far more lucrative and enduring PC market that was developing in the Valley. Such foresight, we believe, is principally available to those who very literally are *in* the know—in the heart of a cluster, breathing the air in which its mysteries spread.

More generally, the level of knowledge about firms leads to a great deal of collective benchmarking. Ambient awareness of what its close competitors are up

to drives a firm to innovate to stay ahead. And the "tipping" effect in high technology, which quickly turns small leads into insurmountable ones, is such that no one can let a rival increase market share uncontested. If one firm seems to be getting ahead, its competitors in the Valley will be among the first to know and thus the first to respond. Consequently, the almost-public knowledge that comes from constantly looking over each other's shoulder leads to collective bootstrapping in all the different specialties of the Valley.

This sort of benchmarking develops around both formal and informal components. Formal components include such things as announcements of future products (vaporware aside) or the registration of patents. Of course, such public declarations are open to all, whether inside or outside the Valley. But close competitors working on similar issues within the zeitgeist of the Valley understand more quickly and better than others what the implications of a product or a patent might be and can, as a result, make well-informed anticipatory moves. This process of action and reaction drives close competitors through rapid cycles in a spiral of innovation.[2] Hence the Valley is currently watching a struggle between digital and optical network switches as proponents of each try to prove that they have the key advantages for next-generation networking.

But the Valley is rich in informal sources of benchmarking, too. Within its narrow confines, competitors shop in the same stores, eat in the same restaurants, send their kids to the same schools, travel in the same car pools, work with the same suppliers, even smoke in the same groups outside the smoke-free office buildings. And, of course, many shuffle between firms, changing jobs with relative ease. All this intermingling makes it almost impossible for people not to know what others are up to. And it gives participants the extra insight to interpret a product announcement, read a patent, understand the significance of a product, or use a new tool. Critically, then, the level of shared informal knowledge "in the air" within a locality provides an unrivaled key for interpreting the formal knowledge produced there.

Mysteries

This interplay between the informal and the formal suggests to us why Marshall doesn't simply say that in a cluster there's a lot of information in the air. Instead, he uses this curious word "mysteries." By it, he no doubt means that what would be secret elsewhere is public knowledge within these clusters. But to anyone who was a student of English manufacturing as Marshall was, the word means a good deal more. Historically, "mystery" was a term for the old Guilds, associations or networks of craftsmen. Marshall seems to suggest that in clusters, these networks are no longer formal organizations, but come with the territory. This is an important idea that we shall return to a little later. For the moment, we want to turn to yet another meaning of the word. "Mystery" also denoted skills, crafts, and the sort of embodied, implicit knowledge that they represent. Thus Marshall puts in the air the sort of knowledge that comes from learning *in situ*, from being where the knowledge is used and getting the opportunity to use it.

This sense of "mystery" takes us beyond ideas of information. And it starts to put some limits on where knowledge will flow and where it won't. Knowledge of the sort that is valuable in the Valley doesn't spread easily unless people are engaged in the skill, craft, or practices to which the knowledge applies (Brown and Duguid, 2000). Without a shared practice, it's hard to get knowledge to move. This sort of "stickiness" helps to explain the confidence of Gordon Forward, CEO of the innovative steel manufacturer Chaparral Steel. He told Dorothy Leonard-Barton

(1995) that his firm has no problem with giving plant tours. "We will be giving away nothing," he argued, "because they can't take it home with them." They can't because they don't share the underlying practice that makes what they see usable.

Knowledge, then, is hard to acquire in a usable form unless the people who would acquire it engage in the actual activity or practices of which the knowledge is a part. Consequently, it doesn't travel indifferently over digital networks, local or wide-area, as information does. As Jeff Papows, president of Lotus, whose Notes is a widely used tool for knowledge management, acknowledges, for all the power to communicate that Notes and similar groupware provide, "spreading the practice has not been easy." And spreading practices is the key to spreading actionable knowledge.

Actionable knowledge and collective practice

The difference between knowledge that is actionable and knowledge that, by contrast, is sterile, builds on a well-known distinction made by the philosopher Gilbert Ryle (1949). Ryle argued that actionable knowledge has two components, know *how* and know *that*. Know *how* is akin to practical experience, know *that* to abstract information. Without the requisite know *how*, know *that* has limited usefulness. With only know *that*, you might talk a good game, but you would never be able to play one. To play you need know *how*. And we learn know *how*, Ryle argues, through engaging in the relevant practices.

This argument suggests that knowledge you can put into practice curiously comes out of practice. New ideas and inventions are spun out of a practice and circulate most readily among people who share that practice. These claims may appear to violate several almost sacred tenets about human knowledge. For example, on their face they seem to deny the importance of theory. In fact, they do not. Theorizing is important. But it is also, as Ryle argues, just another form of practice—the practice of theorizing. Here, as anywhere else, practical experience counts for a good deal.

The importance of a *shared* practice to the creation and circulation of knowledge also seems to deny the primacy of the individual in the creation of knowledge. The individual genius and lone inventor is a cherished figure. Yet it does not detract from any geniuses to note the ways in which they drew on (and inspired) collaborators and competitors. Shakespeare, for example, worked at the center of a group of quite exceptional London dramatists. At different times, members of this group collaborated. They also all stole willfully from one another. (Consequently, there is a great deal of work that cannot easily be attributed to an individual author.) Simultaneously, they all competed, driving each other to greater and greater heights. None of this limits Shakespeare's towering genius. Similarly, the early impressionists as a group experimented with radical innovations in the world of painting, working together, driving each other on, and providing a sympathetic audience for each other while the public at large could make no sense of what each was up to. None of this lessens Monet's achievements.

Nor will such arguments offend those familiar with modern scientific or hi-tech innovation. This has long associated major breakthroughs with clusters of names: Watson and Crick; Shockley, Bardeen, and Brattain; Gates and Allen; or Jobs and Wozniak, for example. But in almost all of these cases, the known names in fact stand for much larger groups of unacknowledged collaborators.

Communities of practice

The way small, tight-knit groups of people working together develop and spread knowledge helps explain the power of what are known as "communities of practice." Lave and Wenger (1991) showed how these groups of interdependent members circulate and reproduce a corpus of actionable, community-based knowledge. Such groups are also critical to the creation and development of new knowledge (Constant, 1987; Brown and Duguid, 1991). Furthermore, as we might expect from Ryle's refusal to separate theory and practice, the community of practice explains knowledge dynamics for groups regardless of whether they are predominantly practical or theoretical. Studies have used the concept to examine data entry processors (Wenger, 1998), service technicians (Brown and Duguid, 1991; Orr, 1996), engineers (Constant, 1989), research scientists (Brown and Duguid, 2000), or top management (Spender and Kessler, 1995). At whatever level or task, small groups, working closely together, sharing insights and judgement, both develop and circulate knowledge inevitably as part of their practice.

Practice, however, is not a knowledge panacea. The strength of these groups is simultaneously their weakness. Shared practice makes it easy to circulate new ideas within such groups. But the absence of shared practice beyond a group's boundaries can make it difficult to get these ideas out of the community. Consequently, nascent artistic movements, for example, can seem absurd to their contemporaries who don't share the practice (as was the case for the impressionists). And new scientific discoveries can seem unintelligible to the scientist's contemporaries. As in the arts and the sciences, so in all regions of life and work, knowledge tends to stick where shared practice ends. The practice has to spread before the ideas can. "Practice," as Ryle notes, "precedes the theory." When they come in this order, once unfathomable ideas can quickly become familiar. Now that people are used to viewing and assessing impressionist art, it no longer seems exceptional at all. Now that working with the ideas of relativity or quantum mechanics has become commonplace, the underlying theory is no longer unfathomable. Thus anyone interested in spreading new knowledge—in the workplace, in the marketplace, through the organization, or wherever, must attend not simply to the relevant know *that* but to the practice of which it is a part. Failure to attend to the practice involved in knowledge dissemination may not only lead to failed dissemination. It may also lead to the inability to understand (and so remedy) the causes of failed dissemination.

High-tech history is littered with accounts of now-familiar ideas that initially "stuck" at the boundary of the local community and its practices. The transistor stuck among a group of core researchers in Bell Labs. The Graphical User Interface (the GUI, which lies behind—or rather, in front of—both the Macintosh and Windows operating systems) stuck at Xerox PARC. It's easy to tell these stories as simple cases of inept management. But that doesn't do justice to the difficulty of spreading new knowledge. Though the value of new technologies may look obvious with hindsight, their implications are very hard to see as they emerge and before the relevant practice are established. Alexander Graham Bell misunderstood the potential of his own telephone, Thomas Edison, his gramophone. Tim Berners-Lee underestimated the potential of the World Wide Web. Gordon Moore, co-founder of Intel and writer of Moore's Law, which reveals a great deal of foresight, candidly admitted that when he first contemplated the PC, he "didn't see anything useful in it, so [Intel] never gave it another thought" (Lester, 1998, 120). The difference between hindsight and foresight is that hindsight can take advantage of the diffusion of practice. Now that everyone depends on transistors, uses personal computers with graphical user interfaces,

and spends time on the web, the application of the theories behind them is much easier to understand.

Inside out

This view of the relationship between knowledge and practice throws new light on the firm. Economists tend to describe the firm as a region of costless information. And sociological views, too, tend to render firms as, internally, unitary "interpretive systems." But firms built around an internal division of labor are inevitably made up of diverse communities, with different practices and as a result different interpretive systems. Moving knowledge in these conditions is a profound challenge. Indeed, from the perspective of knowledge dissemination (the critical perspective for the knowledge economy), the firm's strength lies not in its ability to render knowledge flows costless or to amortize transaction costs, but rather in its willingness to take on the high initial costs of getting knowledge to flow between groups with minimal shared practice. From this perspective, the costless view of the firm obscures the firm's core function of generating and disseminating knowledge.

Once the internal barriers to knowledge-flow become visible, it is possible to understand the essential role firms play in innovation. For, as is now well recognized (see Rosenberg, 1994; Teece, 1986), innovation is usually a systemic process that involves linking the inventive knowledge of diverse communities into something robust and rounded enough to enter the marketplace. Many of the systemic components may be technological. So, for example, a new printing process may require new kinds of ink, new ways of drying, new paper paths and the like. But, as often as not, innovation requires other types of organizational coordination. Thus xerography required not only new technologies, but also new kinds of marketing and new kinds of leasing to succeed. Single communities, then, may be less well equipped than multi-community firms to make the journey from invention to innovation. Innovative firms succeed by bringing together different communities—scientists and engineers, engineers and designers, designers and marketing, and so forth—and coordinating their different practices and belief systems.

So innovative firms might be said to exist to drive knowledge across the gulfs between the different practices necessary to build a path from invention to innovation. It is a difficult process that must always do battle with the diversity it seeks to unify and the risk involved in making everyone adapt to a new idea that may or may not succeed. Given that people's livelihoods are at stake when such risks are taken, resistance can be intense. Corporate antibodies will regularly swarm to keep out a new idea. Inevitably, the challenges of innovation grow dramatically when the invention is disruptive and the attendant risk involves completely letting go of the old—old technologies, old ways of doing things, old organizational forms, and old income streams, despite their proven potential—before the new can prove its worth.

While managing the difficulties and resistance inherent in coordinating new knowledge this way, firms also face the problem that what sticks hard within them may, paradoxically, leak out easily. For example, the knowledge of the transistor left Bell Labs and ended up on the West Coast in a series of firms that began with Shockley, then went from there to Fairchild and the "Fairchildren," which are dominated at the moment by Intel. And the knowledge of the GUI left Xerox and ended up first at Apple and later at Microsoft. Now that venture capitalists stand at the door offering knowledge a ready route out, firms must

apply their coordination skills much more quickly and effectively or risk losing the knowledge altogether. The conventional decision of "make or buy" may be becoming less significant for innovative firms than the decision to "make or sell"—take productive advantage of the knowledge they have developed, that is, before they simply lose it.

Back in the air

But if knowledge sticks within firms, why should it not be just as sticky on the outside? To explain this leakiness, let us first turn back to Marshall's comments about mysteries in the air. We noted that the term "mysteries" could encompass secrets and skills, both of which pervade a cluster. But we also noted that it could refer to guilds or other types of professional or craft associations that extended across particular employers, linking people who do similar jobs for different firms. These sorts of networks are very much in the air in Silicon Valley, where researchers, programmers, engineers, and managers from all the different firms regularly rub shoulders with their counterparts in rival firms.

We think of these networks as "networks of practice" to suggest that they are related to, but distinct from communities of practice. Networks of practice are made up of people that engage in the same or very similar practice, but unlike in a community of practice, these people don't necessarily work together. So, for example, there is a network of hematologists that runs across hospitals, research labs, and medical schools. All members of the network have a lot in common by virtue of the work they do.[3] Though the people in such networks do not all work directly with one another, such a network shares a great deal of common practice. Consequently, its members share a great deal of insight and implicit understanding. And in these conditions, new ideas can circulate. These do not circulate as in a community of practice, through collaborative, coordinated practice and direct communication. Instead, they circulate on the back of similar practice (people doing similar things but independently) and indirect communications (professional newsletters, listservs, journals, and conferences, for example).

Because such a network is not completely uniform, however, knowledge within it does not necessarily circulate uniformly. Hematologists working on leukemia will be able to share some specialized knowledge with other such specialists outside their immediate community, but what they share may not be intelligible to everyone else in the broader hematology network. Furthermore, those doing specialized leukemia research in experimental centers will form a yet tighter circle, again larger than a community of practice but significantly smaller than the network of practicing hematologists. Finally, cutting-edge researchers working directly together on unprecedented work will form the tightest subset of the network, which is the community of practice. Much of their work may be unintelligible beyond that immediate circle.

Take as another example of network and community members a local group of photocopier repair technicians. They generate a great deal of local knowledge, which they share, often through shared practice, with one another (Brown and Duguid, 1991; Orr, 1996,). But the group of ten or so who work closely with one another, forming as they do a community of practice, are only part of a much larger network of Xerox technicians, amounting to some 25,000 technicians around the world. In such a large network, only certain very general ideas will make it across the network as a whole. Some will only make sense to the subset that works on similar machines or for similar customers, or in similar

geographical regions. So where the community is more or less uniform within, the network is not. Its density (and ability to circulate knowledge) will depend on the extent of shared practice.

These sorts of networks help explain how ideas sticking within a firm may leak outside it. The GUI, for example, wouldn't pass readily to the Xerox engineers, with their distinct practice and lack of any experience with personal computers. It leaked, however, to like-minded researchers engaged in similar practices—and hence part of an implicit network—within Apple. Thwarted within, knowledge leaked along a path of lesser resistance to other groups who stood a better chance of overcoming stickiness. Again the mystery, this time in the sense of a virtual guild, was in the air, determining where this knowledge could and would flow.

Mysteries of the matrix

We have thus far described formal and informal links that connect communities and allow them to share knowledge. The firm provides formal links, joining diverse communities into a coupled system for getting work done and, in particular, for promoting new ideas into marketable products or services. Highly schematically, we could represent it as in figure 1.

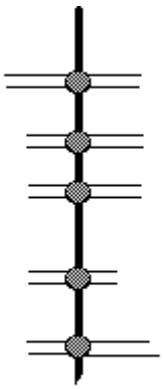

Figure 1

Each of the small blobs represents a different community of (a different) practice linked together in the firm. (We need to emphasize that this is highly schematic. It is not meant to preclude "U" shaped, "M" shaped, matrix, or other organizational forms. Equally, there usually is some overlap between communities within an organization, though we have not dealt with this important issue here.) The solid vertical line represent the connections established by business processes within the firm. The open lateral lines, in contrast, represent the network-of-practice connections that link each of those communities to other communities of similar practices in different firms. Accountants in one firm, for example, have implicit, practice-based links—through their professional newsletters and journals, though informal contacts, and so forth—to accountants in another.

Consequently, a similar but orthogonal diagram can represent a network, as in figure 2.

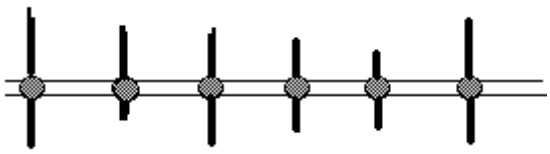

Figure 2

In this figure, the broken solid lines running vertically represent the links within the different firms to which these communities belong, while the horizontal link represents the networks of practice running between them and along which knowledge about that practice will more easily flow. Professional associations, which lie along these lines, have often been important routes for the spread of new knowledge (Constant, 1987).[4]

The links across such a network may be fairly distant. Most people in a large professional association, for example, are unknown to one another. Even those from different firms that do know one another may meet only sporadically at national conferences. This sort of distance puts significant limits on the amount of knowledge that can be shared. But such links can also be fairly close—as they are, of course, in clusters. Here, people in similar jobs who see each other every day in a car-pool, who once worked together and have kept in touch because they are still neighbors, or whose kids are on the same team will tend to intensify the relations within a network of practice. (Indeed, connections can become so dense that intra-firm communities of practice may form de facto. This can happen in joint ventures when engineers from different partners, for example, work closely together across firm boundaries.) Clusters provide the sort of density that allows for proximity and interaction like this. Consequently, in clusters, knowledge, supported on the rails of practice but accelerated by interpersonal relations, can travel particularly easily between different organizations.

The relations of the region, then, comprising both network and organizational links, can be represented as in figure 3.

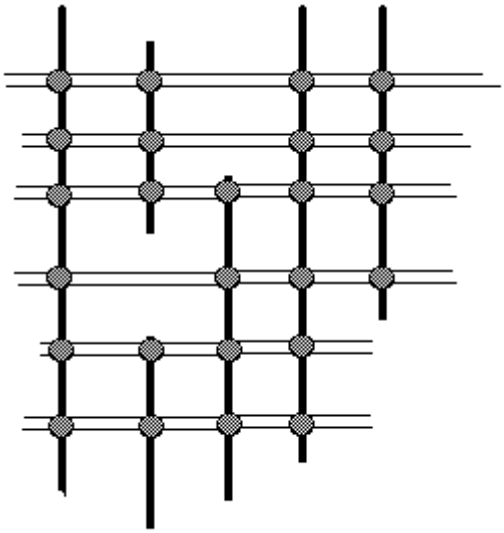

Figure 3

Again, the vertical lines represent organizations vertically integrating diverse practices, while, conversely, the horizontal lines are networks linking similar practitioners in diverse organizations. The organizations, of course, are not all the same. Some may be universities, some government agencies (the military, after all, was long one of the most important organizations in Silicon Valley), and many will be private firms. Depending on their constitution, these different organizations will integrate different practices in different ways. In the Valley, for example, some firms internalize almost all functions in a sweep from finance to fulfillment. Some are only Marshall's "subsidiary" trades—venture capitalists, marketing firms, fabricators, and the like. And others are stripped down to a few basic functions—research and engineering but not manufacturing, for example. Hence, in figure 3 some of the firms represented intersect many of the networks of practice that run throughout the region, others only a few. The latter may then integrate complementary functions not under the hierarchical order of a firm, but through market relations of subcontracting or through some of the other hybrid arrangements that, as we mentioned, fall somewhere between market and firm. These hybrid links are most easily formed where interfirm relations are close, the links between them dense. This sort of density is particularly important in fast-changing areas of the economy, where all partners to a venture need to be able to change in a coordinated fashion.

Such density doesn't only allow partners to coordinate closely. It also allows people to differentiate finely between different firms, finding the most apt for a particular task or idea. And again this discrimination tends to promote leakiness. Because of the density a regional cluster creates, the horizontal links of a network will be more like those found in a community of practice than like the distant sort found across a large professional association. Here, as we have argued, knowledge can leak more readily and successfully. Moreover, as Arrow (1984) has argued, classic knowledge workers—scientists, academic researchers, and the like—are often as loyal to the knowledge they work with as to the organizations they work for. Consequently, new ideas that are going nowhere across the gulfs within an organization may not just leak but get a significant push towards fellow practitioners who, those pushing know, can give the idea a better home.

In all, as a consequence of the density and differentiation within a cluster, localization can make efficient use of people and ideas. In an account of labor movements in Silicon Valley, Angel (2000) notes that "the decisions of individual workers, whether to stay within an existing employer or to change jobs, would appear to be a highly efficient means through which to deploy labor skills and experience within the local labor market." These fluid movements are most likely, Angel suggests, when likely employers are clustered together, when there's a great deal of knowledge "in the air" about who can do what, when the costs of moving are low (no need to sell your house or move the kids to a new school), and when there is, in Marshall's words, a "constant market for skill."

Similarly, localization promotes fluidity of ideas (what we are calling leakiness), by lowering the cost (in terms of knowledge dissemination) of moving them. If a firm fails to use them effectively, ideas, like people, are unlikely to stick around for long. And where informal connections are dense and the mysteries of practice are in the air, the inefficiencies that keep ideas within isolated firms, hedged around by intellectual property strategies and ignorance, are less of a constraint on mobility. People in closely situated and closely related organizations will not be ignorant of what the ideas might signify and how to use them. So ideas will travel (or be pushed or pulled) along networks of practice until they are used. One of the great motivating forces here, as we have noted, is venture capital (see Kenney and Florida, 2000), which tempts both people and ideas to move out of existing firms into start-ups. (That people and ideas move in similar directions is not surprising, given our argument that ideas cling very closely to people and travel along rails built by practice.)

Structure and spontaneity

As ideas and people move rapidly, driven by venture capital, it's easy to believe that a cluster like Silicon Valley conforms to the "Law of the Microcosm" (Gilder, 1989) or the "Law of Diminishing Firms" (Downes and Mui, 1998). Such laws predict that ultimately there is no place for the vertical integration of communities, which we have suggested is the essence of the firm. All such relations will dissolve in the marketplace. Arguments of this sort, we believe, misunderstand Marshall and misread Silicon Valley. Marshall's localizations, though populated with small, specialist firms, do not represent a shift from firm to market. Rather they lie adeptly between firm and market, drawing on both.

Similarly, Silicon Valley does not represent a relentless progress from large firm to small so much as a symbiotic relationship between the two. Small firms, after all, are often the product of large ones. As Kogut, Walker, and Kim (1995) show, large firms sometimes spin off the small to test markets, explore niches, or develop networks of users committed to a particular standard. Certainly, this is not always the case. Often start-ups are the result of a competitive departure, when disappointed employees leave to form a rival, much as the founders of Intel left Fairchild. Nonetheless, whether the spin-out was friendly or hostile, it is important not to overlook "spin-ins," which move in the opposite directions as large firms take small firms in. Some firms in the Valley, like Cisco Systems, seem masterful at spinning in and remaining coherent.[5] Indeed, just as many spin-outs are deliberate, so many apparently independent start-ups are really designed to be spun in. And of course, some that are not taken in, grow instead to be even bigger than their forbears, as Intel did.

The reciprocating movements of the Valley need to be read as an elaborate balance between stabilizing structure, on the one hand, and dynamic spontaneity,

on the other. Venture capitalists, as we have suggested, play an important role in pulling ideas and people out of large firms and into start-ups. Playing this role, they usefully destabilize the "core rigidities" (Leonard-Barton, 1995) of established firms, the "tradition" of long-term networks (Constant, 1987), or the simple reproduction (Lave and Wenger, 1991) of communities of practice. So doing, they are playing a vital part in Schumpeter's idea of creative destruction.

But venture capital thrives on capital gains, which in turn rely on the explosive growth of new firms. By its very nature, venture capital doesn't hang around for profits that come from steady income streams. For these, a different process takes over, one that is less spectacular and gets less attention. Now the firms to which the venture capitalists played destructive midwife must develop a structure capable of continuous innovation. They must establish practices, build traditions, and establish core competencies. At this stage, the creative abrasion that comes from holding things together, rather than the creative destruction that comes from pulling them apart, becomes important. The destruction and the construction are obviously two quite different processes calling for quite different skills. (Jim Clark of Silicon Graphics and Netscape, as the *Economist* pointed out in 1999, has proved himself very good at starting business but poor at running them.) But each depends on the other. For ultimately if there is no promise of long-term profitability, there will be no short term, explosive growth. Without a stable future for the firms at issue, IPOs are no more than Ponzi schemes.[6] So undoubtedly, small firms will continue to emerge. And certainly large ones will come under repeated attack. But large ones will continue to grow as well and small ones will continue to turn into large ones..

Reach and reciprocity

As we try to understand the complexities of the regional matrix in terms of structure and spontaneity, it's important to see that such a matrix also balances reach and reciprocity. Market structures, as Hayek (1945) famously showed, achieve remarkable reach through the information available in the price system. In working markets, price signals the state of supply and demand over global distances. And, in a sense, it is this sort of reach that Marshall seems to believe will come with cheaper communications technology. Firms will be free to locate "near to the consumers who will purchase their wares" without losing what, with less efficient communication, localization provided (Marshall, 1916, 274).

But Hayek's market works by obliterating the details of local knowledge and the minutiae of practice. Reducing all these to price, it makes such knowledge and practice (which are hard to fathom from a distance) superfluous for the purpose of exchange. But as we have been arguing, such things are not superfluous for the purposes of learning, innovating, sharing practice, circulating new, actionable knowledge, and the like. (Markets, as Arrow (1984) convincingly argued, are not well adapted for knowledge.) For these, extensive reach, whether offered by markets or by technology, can be highly problematic, as the survival of Marshall's localizations, particularly in knowledge-dense industries, and despite the cheapening of the "means of communication," argues. More than reach, learning, innovating, sharing practices, and circulating inchoate knowledge all require reciprocity—close interaction and mutual exchanges among the people involved. The links between communities of practice within our matrix (both horizontal and vertical) are two-way links that reflect dense, interpersonal, and often face-to-face interactions. Hence, they are not indifferent to distance.

The necessary reciprocity might be seen, for example, in the workings of the universities within the region. While canonically these are assumed to generate knowledge that passes out into the region, in fact the flow goes both ways. All the schools and departments closely connected to firms in the region—computer science, engineering, and business, for example—live by two-way traffic. Faculty and students carry ideas to the firms, helping the firms to develop their knowledge. But the schools reciprocally develop their knowledge through visits—for talks, seminars, tutorials, and so on—from people who work in the region's firms. Similarly, faculty consulting in the region and students engaged in internships carry knowledge both ways.

The reciprocity is also evident in much of the subcontracting and joint venturing of the region. As Saxenian (2000) shows, firms like Sun and Apple insist on having their most important suppliers close by. Distant connections, she suggests, are inadequate for the constant cycle of ideas back and forth needed to make subcontracting in dynamic knowledge industries work. No contracts can be written to deal with the contingencies involved when specifications must change at the pace of change in the Valley. And even if they could, they could never be enforced. Instead, these agreements must rely on personal connections and trust.

Again, reciprocity seems to be needed to develop the grounds for trust. Cohen and Fields (1999) argue persuasively that the Valley is not a site of the sort of long-term social capital and trust that Putnam (1993) describes based in part on his experiences in the "Third Italy." The networks of family relations Putnam relies on do not exist in Santa Clara County, where almost everyone seems to have come from somewhere else. Instead, the Valley seems to give rise to what Meyerson, Weick, and Kramer (1966) call "swift trust," a trust that can develop over short, intense periods of interaction. Again, for this sort of trust, close interdependent interaction and reciprocity, and not distal communications, seem particularly important, as indeed, do durable networks of practice. Arguing against Putnam, Cohen and Fields conclude by arguing "the sequence runs from performance to trust, not from community" (Cohen and Fields, 1999, 126). We suspect that they oversimplify the question a little. As we have been arguing, reliable performance (or practice) builds communities and networks, and out of this can come trust. But these are not the familial communities of Northern Italy. They are the workplace communities developed as people work together.

From matrix to ecology

We have been trying to explain these complexities of the region in terms of a simple two-dimensional matrix. With each line we write it becomes more difficult. It's hard to see structure and spontaneity in a matrix. You might see reach and reciprocity, but not the complex feedback loops that develop as a result of reciprocity; nor the symbiosis that leads to swift trust; nor the way communities seem to spill over the borders marked by firms and networks nor the constantly changing relations (firm, market, hybrid) that ripple across the region; nor the spaces that new firms occupy.[7]

In all, it seems more useful to switch, as others have, to an ecological metaphor for the region. Seeing the region as an ecology that is home to multiple species but whose growth is ultimately a collective process does much more justice both to Marshall's insight and the richness of regions like Silicon Valley. For the ecological view provides a systemic perspective. What is good for the ecosystem as a whole is not necessarily good for individual species or firms. Indeed, some of

these may have to die for the region as a whole to survive. (Deaths in Silicon Valley get much less attention than births—though see Freeman (1990)). Conversely, protecting individual species, as many working to build a region try to do, may be counterproductive.

The leaking of proprietary knowledge may represent a significant loss to the firm that loses it. If it flows to where it will be more effectively used, the region as a whole, by contrast, gains. A firm faced with such a loss may try to seal itself off from the system. But such isolation, as Saxenian (1996) and Mounier-Khun (1994) indicate, can be quite damaging. Firms that feed into the ecology will, by the same routes (in particular, networks of practice), feed off it. Closing off these routes isolates a firm, and isolation easily means death.

The ecological view helps explain why Marshall's focus on the means of communication limited his otherwise expansive view of localization. From the ecological perspective, the means of communication are only a small part of the overall complexity of the knowledge dynamics of the region. Ecological robustness is built—mysteries are put in the air—through shared practice, face-to-face contacts, reciprocity, and swift trust, all generated within networks of practice and communities of practice. New communications technologies can certainly reinforce these. It is more doubtful that they can readily replace them.

An ecological perspective also addresses the burning question of replication. Throughout the world, people are trying to imitate Silicon Valley by creating what Castells and Hall (1994) call "technopoles." Politicians and business groups are seeking to "bootstrap" new hi-tech clusters in their regions so that they can compete with the established ones and propel themselves from the periphery to the center of the knowledge economy. Such ventures, Castells and Hall argue, would "have to be launched on a huge scale," setting in place a large number of interdependent organizations, institutions, and networks at simultaneously. But the problem is probably not simply one of scale and cost. Even if these costs, which could clearly be enormous (replicating even part of such institutions as Stanford University, the University of California at Berkeley, San Jose State University, and Santa Clara College along with the infrastructure of the Valley would be prohibitive), that is surely not enough. For the Valley and its residents embody a situated, shared experience that has been developed over time and in practice. Both Marshall's history and Silicon Valley's present suggest that knowledge ecosystems develop over time, building connections between participants until they reach a critical mass and take on a collective dynamic all their own. By which time, the units in the whole are not only heavily interdependent, but also path dependent, reflecting the cumulative, participatory practices that have brought them to this point. In all, when regions have reached a point where the mysteries are in the air, if these are mysteries of any great worth, it has probably taken a great deal of time, effort, experience, and trial and error to get them up there.

Our argument is not intended to dismiss attempts to develop strong regional competencies. Quite the reverse. Unlike those who assume the death of distance, we believe "regional advantage" will play a significant economic role well into the future and that those who do not want to be marginalized must try to develop their own regional strengths. Our point, rather, is that these collective competencies are grown organically, not simply implemented mechanically. Aspiring regions can clearly learn from established ones, but because local factors (of culture, institutional forms, and so forth) contribute critically to localization, it would be a mistake (and probably an expensive one) for aspirants simply to imitate the established.

Conclusion

At the outset, we suggested that, given the persistence of clusters despite the cheapening of the means of communication, the two critical questions to address were "*why* and *where*": Why do clusters persist? and, given that the modern economy has seen plenty of disagglomeration, Where might we expect to find them? Our answer to *why* is that clusters persist despite advances in communications technologies because, marvelous though many of these technologies are and extensive though their reach is, they do not have the necessary reciprocity to spread fast-breaking knowledge. In championing the distal powers of new communications technologies to spread information, prognostications easily underestimate the richness of face-to-face interactions and local communications to spreading knowledge.

The answer to *why* then provides the answer to *where*. Though he hints that clusters might continue to exist, Marshall, as we argued, offers no clear idea of where they will continue to exist. We, in contrast, have tried to show that localization meets the demands of knowledge. At least with the current generation of communications technologies, clusters will continue to exist in exactly those industries where fast-breaking knowledge is at a premium. Consequently, as we have suggested, those wanting to develop a robust knowledge economy need to learn how to develop (and not simply imitate) a robust knowledge ecology.

Bibliography

Angel, David. 2000. High-Technology Agglomeration and the Labor Market: The Case of Silicon Valley. In M. Kenney (ed.), *Understanding Silicon Valley: The Anatomy of an Innovative Region*, Palo Alto, CA: Stanford University Press.

Arrow, Kenneth J. 1984. *Information and Economic Behavior.* In K. Arrow*, Collected Papers*. Cambridge, MA: Harvard University Press, 136-152.

Brown, John Seely and Paul Duguid. 1991. Organizational Learning and Communities of Practice: Towards a Unified View of Working, Learning, and Innovation. *Organization Science* 2(1): 40-58.

Brown, John Seely and Paul Duguid. 2000. *The Social Life of Information*. Harvard, MA: Harvard Business School Press.

Cairncross, Frances. 1997. *The Death of Distance: How the Communications Revolution Will Change Our Lives*. Cambridge, MA: Harvard Business School Press.

Castells, Manuel. 1996. *The Rise of the Network Society*. Oxford, U.K.: Blackwell.

Castells, Manuel and Peter Hall. 1994. *Technopoles of the World: The Making of Twenty-First-Century Industrial Complexes*. London: Routledge.

Cohen, Stephen S. and Gary Fields. 1999. Social Capital and Capital Gains in Silicon Valley. *California Management Review* 41(2): 108-130.

Constant, Edward W. 1987. The Social Locus of Technological Practice: Community, System, or Organization. In W. Bijker, T. Hughes, and T. Pinch (eds.), *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge, MA: MIT Press, 223-242.

Constant, Edward W. 1989. Science in Society: Petroleum Engineers and the Oil Fraternity in Texas, 1925-1965. *Social Studies of Science* 19: 439-472.

Downes, Larry and Chunka Mui. 1998. *Unleashing the Killer Ap: Digital Strategies for Market Dominance*. Cambridge, MA: Harvard Business School Press.

Duguid, Paul and Teresa da Silva Lopes. 1999. Ambiguous Company: Institutions and Organizations in the Port Wine Trade, 1814-1834. *Scandinavian Economic History Review*, (special issue on Institutional Theory and Business History, Mary Rose and Sverre Knutson, eds.) 47 (1): 83-102.

Economist. 1999. Know Thyself. October 30th-November 5th, p 76.

Freeman, J. 1990. Ecological Analysis of Semiconductor Firm Mortality. In J.V. Singh (ed.), *Organizational Evolution: New Directions*. Thousand Oaks, CA: Sage Publications.

Gilder, George. 1989. *Microcosm*. New York: Basic Books.

Gilson, Ronald J. 1996. The Fading Boundaries of the Firm: Comment. *Journal of Institutional and Theoretical Economics* 152: 80-84.

Hayek, Freidrich. 1945. The Use of Knowledge in Society. *American Economic Review* 35(September): 519-30.

Hennart, Jean-Francois. 1993. Explaining the Swollen Middle: Why Most Transactions are a Mix of "Market" and "Hierarchy". *Organization Science* 4(4): 529-547.

Kenney, Martin and Urs von Burg. 1999. Technology and Path Dependence: The Divergence between Silicon Valley and Route 128. *Industrial and Corporate Change* 8(1): 67-103.

Kenney, Martin and Richard Florida. 2000. Venture Capital in Silicon Valley: Fueling New Firm Formation. In M. Kenney (ed.), *Understanding Silicon Valley: The Anatomy of an Innovative Region*, Palo Alto, CA: Stanford University Press.

Kenney, Martin and Urs von Burg. 2000. Institutions and Economies: Creating Silicon Valley. In M. Kenney (ed.), *Understanding Silicon Valley: The Anatomy of an Innovative Region*, Palo Alto, CA: Stanford University Press.

Kogut, Bruce, Gordon Walker and Dong-Jae Kim. 1995. Cooperation and Entry Induction as an Extension of Technology Rivalry. *Research Policy* 24: 77-95.

Lave, Jean and Etienne Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.

Leonard-Barton, Dorothy. 1995. *Wellsprings of Knowledge: Building and Sustaining the Sources of Innovation*. Cambridge, MA: Harvard Business School Press.

Lester, Richard. 1998. *The Productive Edge: How U.S. Industries Are Pointing the Way to a New Era of Economic Growth*. New York: Norton.

Lynn, Leonard H., N. Mohan Reddy, and John D. Aram. 1996. Linking Technology and Institutions: The Innovation Community Framework. *Research Policy* 25: 91-106.

Marshall, Alfred. 1916. *Principles of Economics: An Introductory Volume*. Seventh Edition. London: MacMillan and Co.

Meyerson, Debra, Karl E. Weick, and Roderick M. Kramer. 1996. Swift Trust and Temporary Groups. In R. Kramer and T. Tyler, *Trust in Organizations: Frontiers of Theory and Research*. Thousand Oaks, CA: Sage Publications, 166-195.

Mounier-Kuhn, Pierre E. 1994. Product Policies in Two French Computer Firms: SEA and Bull (1948-64). In Bud-Frierman, Lisa, ed. 1994. *Information Acumen: The Understanding and Use of Knowledge in Modern Business*. London: Routledge, 113-135.

Orr, Julian. 1996. *Talking about Machines: An Ethnography of a Modern Job*. Ithaca, NY: IRL Press

Porter, Michael E. 1998. Clusters and the New Economics of Competition. *Harvard Business Review* Nov-Dec: 77-90.

Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton: Princeton University Press.

Rosenberg, Nathan. 1994. *Exploring the Black Box: Technology, Economics, and History*. New York: Cambridge University Press.

Ryle, Gilbert. 1949. *The Concept of Mind*. London: Hutchinson.

Saxenian, AnnaLee. 1996. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Second Edition. Cambridge, MA: Harvard University Press.

Saxenian, AnnaLee. 2000. The Origins and Dynamics of Production Networks in Silicon Valley. In M. Kenney (ed.), *Understanding Silicon Valley: The Anatomy of an Innovative Region*, Palo Alto, CA: Stanford University Press.

Spender, J-C and Eric H. Kessler. 1995. Managing the Uncertainties of Innovation: Extending Thompson. *Human Relations* 48(1): 35-57.

Sturgeon, Timothy J. 2000. "How Silicon Valley Came to Be." In M. Kenney (ed.), *Understanding Silicon*

Valley: The Anatomy of an Innovative Region, Palo Alto, CA: Stanford University Press.

Sturgeon, Timothy J. 1992. "The Origins of Silicon Valley: The Development of the Electronics Industry in the San Francisco Bay Area." Master Thesis, Department of Geography, University of California, Berkeley.

Teece, David J. 1986. Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing, and Public Policy. *Research Policy* 15: 285-305.

van Maanen, John and Stephen R. Barley. 1984. "Occupational Communities: Culture and Control in Organizations. In *Research in Organizational Behavior* vol. 6, edited by Barry M. Staw and L.L. Cummings, 287-365. Greenwich, CN: JAI Press.

Wenger, Etienne. 1998. *Communities of Practice*. New York: Cambridge University Press.

Footnotes

[1] The hybrid character of localization may be true of more than industrial organization. See Duguid and Silva Lopes (1999).

[2] Formal components of benchmarking can also involve selling your competitors the very tools you use to stay ahead. Hewlett-Packard, for instance, created an important market by selling its high-quality testing tools although these inevitably gave its competitors the means to improve their products.

[3] Networks of practice are similar to the "occupational communities" discussed by van Maanen and Barley (1984). Part of our point in changing the terminology is to direct attention from the "community" aspect of such groups, to the centrality of practice in these networks.

[4] Constant also shows how such associations can stop the flow of knowledge. Professional antibodies, like corporate antibodies, may emerge when new ideas threaten.

[5] We are grateful to Martin Kenney for this point.

[6] In the rapidly changing world of modern, high-tech innovation, building a structure for creative abrasion may be much more of a challenge than unleashing the spontaneity of creative destruction, yet the former tends to get far less attention.

[7] "The whole idea of a firm with definite boundaries cannot be maintained intact," Kenneth Arrow (1984, 147) noted while contemplating flows of knowledge and information.

With thanks to Martin Kenney, William Miller, Henry Rowen, Marguerite Hancock, and Sharron Wood.