This paper discusses the results of a two-year research project on the sources of success in regional clusters of entrepreneurship and innovation like Silicon Valley. Our project has studied a number of locations, most of which have shown spectacular rates of growth of information and communications technology-related activities during the 1990s. Our case studies comprise some emerging regions, notably in Ireland, India, Israel and Taiwan, along with more advanced areas like Northern Virginia in the US, Cambridge, UK, the Scandinavian countries and the Silicon Valley 40 years ago by way of the memory of one of its ‘father founders’, Gordon Moore. Through visits, interviews and other materials, we uncovered some regularities about the determinants of success of these entrepreneurial-led models of economic growth. We find that the economic factors that give rise to the start of a cluster can be very different from those that keep it going. Agglomeration economies, external effects and ‘social increasing returns’ of any sort arise almost naturally after a cluster has taken off. But the most difficult and risky part is to get the new clusters started. At that stage, ‘old economy’ factors like firm-building capabilities, managerial skills, a substantial supply of skilled labor and connection to markets were crucial for the take off of these ‘new economy’ clusters (including Silicon Valley 40 years ago).

1. Introduction

This paper discusses the sources of success in regional clusters of entrepreneurship and innovation like Silicon Valley. It draws from the results of a two-year research project carried out by a large international and
interdisciplinary team. The project has looked at how different forces, including public policy, business strategy, and institutions at regional and national level have combined to encourage the emergence, growth and maintenance of clusters, and how they might achieve the level of positive feedback and ongoing success of Silicon Valley itself. Our viewpoint has been comparative. We have used both in-depth case studies and statistical methods to analyze examples of early, preliminary and partially successful development of regional clusters in information and communications technology (ICT), inside and outside the United States.

The topic of this paper has been difficult to attack until now, both for practical and conceptual reasons. On the practical side, international comparison of Silicon Valley imitators has suffered the difficulties of comparing a roaring success to some bitter failures. Where Silicon Valley is entrepreneurial, decentralized, and only loosely and flexibly connected to broader national institutions, many efforts at imitation have been government sponsored, top down or tightly linked to established firms, perhaps ‘national champions’. This wide divergence makes an analytical approach difficult, for one cannot easily investigate success drivers for getting over the positive feedback hump by looking at places that have been a roaring success or at places where so little of the logic is right.1 As a result, opinion in the policy arena has fractured, with some believing that there is a magic formula for creating a new cluster of innovation and others thinking that the Silicon Valley model is unique and inimitable outside the United States. Much of the same ‘all good/all bad’ thinking characterizes analyses of Silicon Valley as well. At one stage, the region and its companies could do no wrong, at another, no right.

Against this background, the mission of our project has been to analyze a number of different attempts to gain national economic advantage from regional clusters of development in ICT. We have sought to avoid both the hagiographic ‘SV is great’ mode and the hypercritical ‘there is no new economy’ mode. Our main goal has been to assess and possibly identify the sources of long-term economic growth in clusters of industrial activity.

We define a regional cluster simply as a spatial and sectoral concentration of firms; and we measure success by the ability of the cluster as a whole to grow, typically through the expansion of entrepreneurial start-ups. In our project, we have then identified a number of nascent clusters of technology-based innovative activity around the world. They are all distinguished by entrepreneurship and growth, and they all have substantial focus on ICT. In

1 The study that overcomes this difficulty most successfully (Saxenian, 1994) looks only at one country and still finds itself comparing a region (Silicon Valley) wherein many structures support entrepreneurship to one (Route 128) where few structures do.
short, they are all 'success stories', and in this sense they are all 'young Silicon Valleys'. Indeed, as we shall see below, one nascent cluster in our study is the Silicon Valley of four decades ago, and we asked one of its 'father founders', Gordon Moore, to recount this story. With that one exception, however, we do not yet know if the other centers that we have looked at will achieve the level of commercial and technological success of the mature Silicon Valley. It is not even clear whether these nascent clusters will build sophisticated support and service industries (like venture capital) and experienced entrepreneur-mentors that would make founding a new firm as easy in them as it is in Silicon Valley today.

Our sampling scheme allows us to take up two kinds of questions, but it rules out others. It lets us examine the similarities and differences among emerging clusters, but does not let us contrast these nascent clusters with failed efforts to get clusters going. The reason for this is simple. Unsuccessful cluster attempts, such as Malaysia’s Multimedia Super Corridor, have typically failed to provide useful contrasts for the same reason as earlier literature failed. The similarities and differences among the nascent clusters we have selected, however, provide a rich field for investigating other analytically important questions. For instance, what factors are systematic and what are unique feature of the western United States in these emerging regional clusters? Answering this question is critical for breaking out of the 'recipe' mode of analysis that has served policy formation so badly in recent decades.

Specifically, our study examines the forces that create new clusters of entrepreneurship-led growth. More broadly it is an empirical investigation into the microeconomics of growth and trade. Our international and interdisciplinary team has compared nascent regional clusters both inside and outside the United States. We pose two interrelated questions. First, how did these ‘imitators’ of Silicon Valley initially become centers of ICT-related growth; and second, what accounts for the subsequent success (or lack thereof) of these clusters to ensure that success builds on success in a self-reinforcing fashion. The goal is to understand how factors including national and regional policy, business strategy, and national and regional institutions combine to encourage the growth of regional clusters.

Our team has included the Stanford SIEPR-affiliated economists Tim Bresnahan, Kevin Davis, Michael Horvath and Scott Wallsten; business people/entrepreneurs, notably Ralph Landau and Gordon Moore; and management specialists and economists from other universities or institutions who work on the development of technology clusters, including AnnaLee Saxenian from Berkeley, Suma Athreye from The Open University in the UK, Ashish Arora from Carnegie Mellon University, Erran Carmel from American
University, Catherine de Fontenay from the University of New South Wales, Alfonso Gambardella from the Sant’Anna School of Advanced Studies, Tamah Morant from the University of North Carolina, John Richards from McKinsey & Co., and Salvatore Torrisi from the University of Camerino.

We have looked at a number of different clusters: outside the US in Ireland, Cambridge, UK, Israel, Scandinavia, India and Taiwan, and, within the US, Northern Virginia in the present and Santa Clara County (Silicon Valley) in the 1960s. In spite of their significant differences, India, Ireland and Israel are the prototypical cases of nascent ICT clusters that we have in mind. They have all exhibited a significant acceleration in the production of ICT (especially software and services) during the 1990s. Their ICT growth has been exceptional according to practically all major indicators: annual double-digit growth in the number of new firms, in ICT revenues and employment, in exports; increasing share of ICT of total exports of the country (up to one-third of total exports in the case of Israel). Taiwan has shown a similar pattern in the manufacturing of PCs and related businesses. In the advanced world, Northern Virginia has shown similar features, including double-digit growth figures. Our two special cases are the two European regions. While Cambridge, UK has all the feature of a cluster in the sense we are defining it here, it does not match the patterns, and particularly the growth figures, of the other regions that we have looked at. This is not to say that Cambridge, UK has not been successful. But compared to the other regions (including Silicon Valley), the growth figures have simply not been the same. The Scandinavian story, as we shall see, is interesting because it is the only case in which growth was triggered by a whole different vertical ICT market, wireless hardware.

Our team of researchers has worked with the data on these clusters and we have also visited each of the regions to interview key actors. We have met a number of times over the course of a two-year period in order to review and compare preliminary results and analysis as well as to jointly develop conclusions. In what follows, we discuss the key findings of the entire research.

2. Agglomeration Economies and External Effects

Many of the existing theories of clusters of innovative activity focus on external effects and the resulting agglomeration economies. One central

---

2 Of course, when we mention India we do not mean the entire country but the few regions that have shown a significant growth of ICT during the 1990s.

3 Some of the classical references are Krugman (1991), and more recently Porter (1998). Saxenian
feature of clusters of innovative activity is external effects among the technology firms located there. A local external effect is anything that raises the return to particular firms located in a region as a result of the location of other firms in the same region. External effects can be direct, as when managers or technologists learn about market or technical developments from colleagues in neighboring firms, when firms in closely related industries serve as one another’s customers or suppliers, and so on. External effects can also be indirect, as when key inputs are in abundant supply or when the overall level of commercial technology activity is high. These indirect external effects arise from increasing returns to scale in the supply of key inputs such as venture capital, which may locate where entrepreneurship is dense but support the development of new entrepreneurial firms; a thick labor market in technical personnel; or commercially oriented activities in universities or national laboratories, to name just a few. Both direct and indirect external effects generate positive feedback loops that ensure that technology-related firms locate in regions where other technology firms are already located.4

These external effects have (at least) two distinct implications. One implication is for economic growth, both within a region and in the broader economy. External effects among innovative firms inventing general-purpose technologies are highly levered mechanisms for increasing the rate of growth of an overall economy. By raising the rate of return to invention, either direct or indirect external effects can, if the field of the invention is important enough, push the commercialization of valuable technologies faster and closer to markets. Similarly, in the case of indirect external effects the economic return to key inputs such as highly skilled labor, the knowledge to be found in universities or the market judgment of venture capitalists can be raised by clustering among firms. A second implication of external effects is that nothing succeeds like success under external effects, and the private return to participation in these clusters of innovation by entrepreneurs, venture capitalists, technologists and those in the key supplier industries can be enormous.5

(1994), which we quoted earlier, is another classical citation. Arthur (1990) discusses specifically increasing returns in the context of ‘Silicon Valley locational clusters’. There are also, of course, the classical, early references to agglomeration economies that date back to Marshall (1920), Perroux (1950), Myrdal (1957) and Hirschman (1958). Finally, Jaffe et al. (1993) and Audretsch and Feldman (1996) provide evidence of the extent of geographically localized knowledge spillovers.

4 The direct/indirect language follows the standard usage in external effects theories. This mirrors the classical Marshallian distinction, which pointed out three sources of external economies: a thick labor market; specialized input producers (and related increasing returns), which arise because close-by producers entail larger local markets; knowledge spillovers and ‘untraded interdependencies’.

5 This is how thoughtful analysis of agglomeration economies and external effects states what those concepts can explain. See, for example, Fujita et al. (1999), who argue against use of ‘ad hoc dynamics’ to take models too far past the question ‘When is a spatial concentration of economic activity sustainable?’
These two points—external economies of scale, and the resulting capture of the rents by producers and the regions in which they are located and headquartered—explain the great interest in agglomeration economies. But network analysis alone cannot explain how regional clusters emerge. The existence of external effects explains why a region would like to have entrepreneurial-led growth and why the world’s consumers would like to have more clusters founded. But examining the positive feedback in an existing successful cluster does not tell you how a cluster begins. Moreover, while clusters of entrepreneurial-led growth are valuable social institutions, we know little about why clusters begin where they do or how many clusters will emerge within a given industry. There is a strong element of social increasing returns to scale in the argument, which one would normally expect to lead to a limited number of clusters in any particular industry or technology. Understanding how nascent clusters overcome these limits has been an important part of our study.

The positive feedback logic of the theory and of established and successful clusters makes it difficult for analysts to identify a starting point. Positive feedback, when it is working, appears as a virtuous cycle, and, when it is not working, as a difficult chicken-and-egg problem. For a region that is not yet succeeding, ‘nothing succeeds like success’ is an empty remark. The related more analytical observation that all of the elements of success (entrepreneurship, venture capital, etc.) feedback to one another positively does not communicate what it is important to get started. Our analysis, by looking at the nascent clusters that exhibit significant signs of growth, lets us address the question of how to start. We return to this question of the factors leading to the start of a cluster in Section 3 below.

The second aspect of external effects, namely the possibility of social increasing returns and the resulting producer rents, also contains the analytical temptation to draw a false dichotomy which positions ‘old economy’ explanations of growth as alternatives to ‘new economy’ ones. Here we do not need to linger on a sophisticated distinction between what is old and what is new economy. For the purposes of our discussion the common sense with which these expressions are currently used is sufficient. ‘Old economy’ is a shorthand for a number of concepts: organizational and firm-building activities, investment in general and industry-specific human capital, larger
companies and related economies of scale at level of the firms, lengthy periods of investment in capability before their exploitation. ‘New economy’ means instead entrepreneurship, economies of scale at the level of regions or industries rather than firms, external effects, etc.⁷ Our point is that a strong opposition between old and new economy explanations of growth can lead to some analytical errors. Should we seek to explain the success of clusters only by external effects? Is this newly discovered feature of supply so important that it trumps all others? Much ‘new economy’ thinking goes down this false dichotomy path, drawing a distinction between the positive feedback forces that are at work in a cluster—guarantors of instant success in the most fervid ‘new economy’ thinking—and the supposedly ‘ordinary’ or ‘old economy’ work of firm-building, market-building, invention and commercialization.

The studies carried out in our project show that it is an error either to focus only on the external effects and conclude that new economy logic supersedes old, or to conclude that there is nothing to the positive-feedback external effects story at all. Instead, in the construction and maintenance of successful clusters, the new economy and old economy elements act as complements to one another; neither one can succeed without the other. Many governments have made the analytical error of focusing far too much on the second aspect of external effects, and have viewed clusters of innovative activity as no more than a ticket to producer rents. This has provided the intellectual foundation for largely failed policies that attempt to jump-start growth in clusters by directive policy. In our study, we have taken another look at that kind of policy, but do not dwell on it. Instead, we focus on places—and governments—that have got the mix of policies close to right.

3. Starting a Cluster

The first step in understanding the complementary relationship between external effects and ordinary investments is to step back from the already-built, successful Silicon Valley cluster of the present. We need to look, instead, at clusters in the making, whether in Silicon Valley in the 1950s and 1960s, or regions like those that we have looked in our SIEPR project, e.g. Ireland, Israel, India, Northern Virginia and Taiwan in the present. In each of these cases we found that there are some external effects at the nascent stage. Yet

⁷ This form of organization distinction, rather than the sectoral one, underlies our use of the old/new economy language.
very important forces for success arise in the ordinary business challenges of building firms and building markets.

Put simply, our argument is that the processes of starting and sustaining a cluster have different economics. Starting a cluster involves, first, building the economic fundamentals for an industry or technology, and, second, finding the spark of entrepreneurship to get it going. Both of these are supported by a number of common elements in the regions that we examined—and it is striking how similar all the nascent clusters are to one another in this regard, and how much more similar to one another they are than to the current established and thriving Silicon Valley cluster. It is particularly significant that the Silicon Valley of 40 years ago is also closer to today’s nascent clusters than either is to the Silicon Valley of today.

Many similarities between today’s successful clusters and the Silicon Valley of the 1960s arise because founding a new cluster, or the early firms in a new cluster, is a very different entrepreneurial and economic activity than founding a firm in an established cluster. New clusters, including Silicon Valley in the 1960s, offer substantially less support to entrepreneurship in the startup or pioneering phase than does a mature cluster like Silicon Valley today. External effects—namely benefits to particular technology firms that arise from the presence of other firms or of support structures like venture capital—play only a small role in the early phases. Such benefits typically come later in the development of a cluster.

Among others, one similarity across the new clusters that we have examined is that they have taken advantage of a new technological and market opportunity that had not already been exploited: the integrated circuit industry in the Silicon Valley of the 1960s; the internet- and network-security markets exploited in Ireland and Israel today; specific opportunities like software demand following the YK2 or the Euro problem in India; the hardware and equipment opportunities in new kinds of devices like cellphones and PDAs embraced in Scandinavia and Taiwan. The rationale for this is not hard to guess. Markets with substantial producer rents, like ICT today, are characterized by powerful forces that make a direct assault on an existing market position unpromising. In the first place, the advanced technological capabilities and market connections of incumbents may offer a commanding lead. These markets often have barriers to entry created by the deployment of commercialization assets in established markets. At the same time, there can be market forces associated, for example, with de facto standard setting, as in much of computing, or political forces that protect existing national champions, as in parts of telephony, which make entry difficult. Broadly speaking, the lesson from our case studies is that to blossom, the new-cluster
entrepreneurs have to turn away from established sources of rents to define new ones, and (at least initially) make their relationship with existing technologies and clusters complementary rather than competitive.

In this respect, practically all the technologies developed in the new clusters that we have looked at in our study are complementary to existing ICT technologies, mostly sold by US-based firms or US-linked multinationals. Since ICT is used in large complex systems, it is important for the new entrants to have their inventions and advances linked with other inventions and advances. At the same time, this means that the new clusters can take advantage of the sizable and growing demand for the leading technologies. Our clusters in Ireland, India, Israel and Taiwan come from regions that, for one reason or another, have easier potential interactions with the US market (language, cultural connections, diaspora, etc.). They have then been able to take great advantage of the significant US demand for ICT products, services and components, during the 1990s. Combined with other factors that we shall highlight below (particularly underemployed skilled labor), this has given rise to a powerful mechanism for sparking off growth opportunities in some ICT niches and segments.

At one level, this is no more than the obvious remark that ‘demand is important for growth’. In a debate over-emphasizing supply-side factors like agglomeration economies and external effects, however, this remark is often overlooked. That market and demand factors lead to linkages to the US does not reflect a ‘US-centric’ view. The linkages with the US are merely a powerful example for pointing out how important it can be to be connected to a sizable and growing source of demand. Finding the source of demand that may spark off the growth of the cluster can be critical for its rise, and in many respects it should be one of the policy focuses in this arena.

In brief, to obtain growth one cannot point only to the ‘surfing’ of agglomeration economies, but also to the underlying ‘wave’ of technological and market opportunity. For many of the regions that we have examined in our project, the wave was the great expansion of ICT following the commercialization of the internet.

Another similarity across our new clusters is the degree of investment, effort and building needed to set up the background for an innovation cluster’s take off. To take up our earlier remarks again, development of the ICT businesses is not the magic exploitation of some ‘new economy’ rules. Instead, it takes years of firm-building and market-building efforts. The long-term investment in education of a skilled labor force has been critical in a number of regions, notably in Taiwan, Ireland, India and Israel. The supply of highly skilled workers can come from any of a number of sources, but needs
to be plentiful: Stanford-trained scientists and engineers to Silicon Valley, military training and Russian émigrés to Israel, alumni of ‘beltway bandit’ federal contractors in Virginia, and an educated population earning far less than world-standard wages in Ireland and India. While it is critical to invest in the assets that will permit the emergence of a cluster, there is no magic recipe (take one great university, 47 venture capitalists, and . . .) that works. Instead, a number of different routes exist to building the backdrop—technology opportunity, educated labor, flow of entrepreneurial talent, etc. 8

Sometimes these long-term investments in national or regional capabilities can grow for a long time in what seems like a low-return mode before the take off into cluster growth.

The story of early entrepreneurship and the early integrated circuit industry seen from the perspective of Fairchild and Intel illustrates this point very well. This is the story of the developments that put the Silicon in Silicon Valley, though there was the beginning of a high-tech cluster in Santa Clara County before these times. 9 What was important in determining the success of the early Silicon Valley back in the 1960s? Moore and Davis (2001) bring forward several elements. First, there was a rich technological opportunity in the semiconductor business. Second, there were immediately available markets such as consumer electronics and defense. The creation of general-purpose semiconductors led quickly to sales in those markets, where the value of a miniaturized component was clear. There was also the prospect of whole new uses of semiconductors in information technology industries, a real prospect for long-term growth. Both the technological and the market opportunity were separate from the existing high-tech (tube-based) electronic industry, and provided an advantage to producers located far from existing sellers.

How was this technological and market opportunity exploited? The Moore–Davis story is as far as can be imagined from one in which young entrepreneurs instantly succeed in a supportive environment of external effects. None of the modern institutions of Silicon Valley existed, so none of the incoming benefits of external effects were there. No mentoring from experienced entrepreneurs now working as venture capitalists, no easy access to the required skills or to wise and experienced thoughts about business models, no networks of connections to supply partners and marketing partners—none are part of the early period story. Instead, the story is a one of investment in human capital, firm building and market building processes.

8 In our project, the story of Taiwan is told in Saxenian (2001); Arora et al. (2001) discuss India and Ireland; de Fontenay and Carmel (2001) studied Israel.

9 Hewlett Packard, for example, was established years earlier and it was growing at the time of the founding of the integrated circuit industry in Silicon Valley.
that took a long time and quite serious effort and risk. Allowing for changed circumstances, this same element is present in many of our regions. Building a new capability at the firm level which will lead to local increasing returns and positive feedback does not involve anticipating or exploiting those high-payoff features, it involves investing in the key assets which will permit later collective payoffs.

To summarize, the forces underlying the emergence of a cluster differ from those needed to ensure its continued growth. While increasing returns and external effects can keep a cluster going, the initial spark is more difficult to obtain and more risky to pursue. Our research suggests that these include the importance of being linked to a sizable and growing demand as well as the availability of a proper supply of key factors like, in the case of ICT, skilled labor. Other critical factors are firm- and market-building capabilities. These require significant and systematic efforts by the ‘pioneers’ of the cluster to promote organizational and technological capabilities of various sorts, create new firms and institutions, etc. Finally, another factor is plain ‘luck’. Founders such as Moore recognize that there was considerable uncertainty at the beginning about the potential size of rents and the appropriate firm and industry structure to pursue them. Similarly, the opportunities for many of the newer cluster involved a particular matching of regional supply capability to world demand.

In fact, there is a logical argument for suggesting that luck plays a role in this context. We noted that nascent clusters, and the entrepreneurs operating there, have to bet on new trajectories before they manifest their potential. But this also means that they have to bet on an opportunity before it is clear to everybody else that it is indeed an opportunity. Some degree of risk is therefore unavoidable. At the same time, this means that only some of these opportunities (and most likely few of them) will materialize. Many attempts at creating new clusters and successful new firms in certain industrial or technological trajectories will fail, and they will fail in spite of the fact that the key actors have done all the right things that are to be done in these contexts. In this area it appears that luck and skill are complements; those initiatives that embody a superior business model or technology are more likely to find the ‘luck’ they need.

This string of similarities across our cases corroborates our earlier remark that the ‘new economy’/old economy’ distinction is a false dichotomy. There are two errors. It is just as incorrect to say that clusters of innovative activity can take off as if by magic as it is to deny the huge national and regional advantages that accrue from the existence of an established Silicon Valley cluster. The truth is that old economy hard work, both in company building
and in regional investments such as education, and new economy external effects are complements—each is more valuable with the other than without it. There is a great deal of truth to the ‘external effects’ theory that entrepreneurship is easier in a cluster than outside of it. This is the realization of the theoretical idea of ‘social increasing returns to scale’. Yet that certainly does not mean that the world can support only one cluster, and the growing pressure on the limited stocks of skilled labor and land in the US and Silicon Valley are powerful forces favoring emergence of new clusters.

4. No ‘Recipes’ but Some Deep Regularities

4.1 Highly Skilled Technical Labor

All our regional stories point to the importance of highly skilled labor as a precondition for the growth of an ICT-based entrepreneurial cluster—Taiwan, Ireland, India and certainly Israel. In some cases, underemployed skilled labor is close to being the ‘only’ factor (India, but also Ireland or Israel), or at least the one that spurred the rest. Further, some of our other stories reveal that there is a role for universities both as a source of skilled labor and of technologies that are exploited for export and growth—e.g. Cambridge, UK and Silicon Valley in the 1960s.10

Taken as a group, however, these stories are not at all encouraging of the simple ‘recipe’ view of universities and higher education in starting a cluster. Our cases demonstrate that there are a number of different ways to achieve a supply of skilled labor and that it is the ultimate outcome (a highly skilled labor supply), not the particular mechanism (a university) that matters. Large firms can also play a critical role in growing the skill base. While the role of universities like Stanford or Berkeley has been widely emphasized in the Silicon Valley story, one should not neglect the potential training provided by established firms like Hewlett-Packard or Intel. Large firms often nurture technical competencies. For example, individual researchers can use equipment or they can be part of research teams that would hardly be available outside the leading companies. Similarly, many offer managerial training, and possibly even managerial connections. Moreover, this technical and managerial training can encourage spin-offs. This raises a classical problem of ‘private’ vs. ‘social’ interest. As pointed out by Moore in his study with Davis, established companies warn against spin-offs. Even though Moore himself notes that one can benefit from spin-off supplier industries (e.g. chip

10 Athreye (2001) discussed the story of Cambridge, UK in our project.
manufacturing equipment), established companies clearly look unfavorably to the spillovers that they create by training people who would then use these competencies outside the firm.

Our stories pointed out other sources of training, apart from universities and large firms. In Northern Virginia the development of sophisticated technical capabilities in ICT, and particularly in communication technologies, has stemmed from the existing bases of competencies provided by years of contract research for the government and the defense department in the area. Government contracts have for many years nurtured contractor firms, and the related technical skills in such firms. Once defense procurement slowed down, and new internet-related and other communication technology opportunities rose, skilled labor in these fields were available in the area, and this implied the supply of skills that were fruitfully used for launching the cluster. In Israel, the military proved to be the key supply of technical skills, along with some leading technical universities (and particularly the Israel Institute of Technology, or Technion). Finally, the supply of skills can come from outside the region. The clusters in Taiwan and India have drawn heavily on US-educated Chinese and Indian engineers. In Israel as well, immigration (largely from Russia) has been an important source of skill. Similarly, the nascent Silicon Valley attracted engineers from all over the United States (Bresnahan et al., 2001; de Fontenay and Carmel, 2001; Moore and Davis, 2001; Saxenian, 2001).

In sum, the conventional wisdom lumps education, skilled labor and universities in a single idea and often overlooks alternative mechanisms for achieving a skilled labor pool. It is therefore important to stress that a university per se is not essential to the emergence of a successful cluster. This qualification is useful for at least two reasons. The first one is that it brings further support to the rejection of a formulaic recipe for the rise of the clusters. Putting a university at the center of the cluster can help, but it is neither a necessary nor a sufficient condition. Second, there can be different sources of skills in different regions, and—given that it is the availability of the skills that matters—regions can look for the most appropriate way (or mix thereof) for acquiring skilled labor, from universities to larger firms and other local as well as distant institutions. The policy implication is also straightforward—forming or attracting skilled labor, rather than a particular means for doing it, is the crucial aim.

4.2 Managerial Labor

The contributions of returning expatriates also suggest that the key human
capital investments are managerial as well technical. Most of the nascent clusters we have analyzed are focused on ICT—as is much of the activity in Silicon Valley. They therefore offer difficult management and marketing challenges because these sectors present nontrivial problems of commercialization of scientific and technical opportunity.

Firms in new clusters and sectors develop sources of non-technical (or not narrowly technical) human capital in a variety of ways. Silicon Valley itself, being perhaps more ‘first’ than the other clusters, has produced an indigenous supply of scientist-managers and engineer-managers by having scientists and engineers learn a second skill set (management), often by experience, in the early stages. Of course, as Silicon Valley matured, it used a number of other models to create dual-knowledge-base managers, including more experiential training (think of Steve Jobs or Larry Ellison,) mentoring by experienced entrepreneurs or venture capitalists of younger managers-to-be, and career moves from established firms in the industry (in and out of Silicon Valley) into new entrepreneurial firms.11

Later-growing clusters have not needed such heroic means of creating managerial capabilities as did the early Silicon Valley. ICT markets are more established, and technical people native to a region can achieve training in a number of ways that introduce them both to the management problems of ICT enterprise and to world markets. Multinational enterprises (MNEs) have played an important initial role in this regard—e.g. Taiwan, and possibly Ireland, certainly Israel.12 Potential engineer-managers can work in MNEs in their home region and gain experience in a second skill set, while also learning about world markets. This mechanism for building managerial human capital can later be deployed in indigenous firms. This is an advantage of late development.

A second source of the same human capital building arises through the return of expatriate engineers and managers working in established clusters elsewhere, especially in Silicon Valley itself. When repatriated, these individuals bring with them experience and knowledge of management in ICT firms, in this case even in ICT startups. And, since their work was overseas (often in the largest market, the US) they return with connections to and knowledge of ICT markets. This mechanism is the long-distance analog of the one that has been going on within Silicon Valley for some time, namely

11 If anything, this problem was quite clear in the view of the early Silicon Valley, when Gordon Moore and the others, as emphasized in his story, had to hire a fully fledged manager, outside of their technical group, Ed Baldwin, who took up the management for the company and taught them stuff that is nowadays commonly taught in MBA courses (Moore and Davis, 2001).

12 The more traditional role of MNEs in providing technology transfer and a mechanism for connecting to world markets is not crucial in our cases (e.g. Arora et al., 2001; de Fontenay and Carmel, 2001).
development of human capital within established companies followed by a
spinout or a movement of experienced workers to a new firm. The difference
is that these spinouts and this labor mobility go to a place where the engineer
or manager would like to live—India, Israel or Taiwan—rather than to a site
just a few miles away.

4.3 New Firm Formation and Firm Building

We have already noted that firm-formation and firm-building capabilities are
important elements of successful clusters. In this respect, the studies in our
project show that there is a difference between the growth of firms and the
growth in the number of firms in a cluster. Growth in the number of firms
may not lead to the sustained creation of substantial economic rents. Similarly,
growth in the size of the high-tech sector, as in the early phases in Taiwan
and partly the current phase in India, does not necessarily ensure continued
growth. The question is whether some firms in the clusters emerge out of the
class of small entrepreneurial startups and stake out independent positions in
world markets.

To be sure, growth of firms could be a signal rather than a cause of success.
A firm that pioneers an important innovation or one that generates a new
market will most likely grow. In turn, the growth of local firms can be critical
for the success of a cluster. Growing firms create demand for other types
of employment and not just technology (e.g. manufacturing, marketing).
Moreover, as many technology markets have shown today, most of the rents
are downstream. Companies with market power have the resources to make
additional investments, with further implications for growth. More generally,
larger companies, like Intel or Hewlett-Packard in Silicon Valley, Ericsson and
Nokia in Scandinavia, produce backward and forward linkages which systems
of smaller firms may not be able to produce, at least to the same extent. These
are all important factors for sustaining the growth opportunities.

Our regional studies show that the growth of firms is not a natural outcome
of the rise of a cluster. For instance, the new companies could be consulting
or service firms established by individuals whose primary job is elsewhere
(e.g. university professors). In these cases the effort on firm building would
be more limited than those by individuals who would make primary bets
on their companies. Also, some entrepreneurs prefer not to have their firms
grow. Growth implies new management challenges, possible decentralization
of powers and greater tension, risk or responsibilities. It is not uncommon in
some locales, as in our case of Cambridge, UK, to find individuals who are not
interested in these challenges. If commitments to firm-building are important
this could limit the growth of a cluster, particularly compared with the risks and investments in firm building made by the early founders of Silicon Valley and by some of the entrepreneurs that populate the Indian, Israeli or Taiwanese scene today. In sum, we suggest that the growth of companies, and not just the growth in the number of firms, is a signal of the success of the cluster. These are the firms that will become one of the sources of increasing returns for the continuous growth of the cluster, in the form of training for potential spinouts, development of managerial and technical competencies, along with various forms of backward and forward linkages.

The growth (or lack thereof) of local firms may also stem from comparative advantage. Indian engineers and Cambridge boffins may be reacting to the incentives they face, and supplying only what is economically efficient to the world economy (in these cases, the services of highly skilled individuals.) Over time, this would lead to a classical gain from trade, i.e. wages for highly skilled labor far closer to the global level for labor of that type. Israeli high-tech companies, for example, are often bought by American firms and moved to the US as soon as they reach a certain size and stage in their evolution. This may not be bad for the Israeli economy if the companies are paid the right price, notably the present value of the company given its expected opportunities. If the market for companies showed few imperfections in this respect, the gains from trade could produce notable benefits. If they remained in Israel, these companies might be limited by the size of the domestic market. Since development and commercialization (unlike initial research) benefit from exploitation of larger markets, there may be an advantage from selling the companies once they reach the stage when a large market becomes an important asset. This then shifts the problem from a general preoccupation with the acquisition of the best domestic firms, toward the potential imperfections in the international market for companies.

Another regional path is specialization in the creation of new innovative companies, with acquisition and commercial development of those companies or the technologies they create by established firms outside the region. Success on this path requires a fertile field of technological opportunity. If the field matures or is exhausted there may be limits to creating rents from selling new technologies. This too can be seen as a problem of setting the right price for the technology, or for the acquired company. Research suggests that asymmetric information leads to payments for technology licenses that are below the actual present value of the firm and to the acquisition of technology firms at prices below their long-term contribution (Arora et al., 2001). Yet even under these conditions comparative advantage may still be critical. It may be that focusing on the early part of the innovation cycle is the right
specialization for Israeli companies. In turn, the advantages of specialization would increase if the imperfection of the global technology markets were reduced, and the companies sold for a price closer to the ‘right’ price.

4.4 Connection to Markets

We have already noted that our emerging ICT clusters in Israel, India, Ireland and Taiwan all had significant ties with the United States, which helped them to exploit the ICT-intensive US growth of the second half of the 1990s. In this respect, it was critical for these regions to position themselves in product spaces that were complementary to the main sources of demand (notably Silicon Valley and the US) rather than directly competing with them (Arora et al., 2001; de Fontenay and Carmel, 2001; Saxenian, 2001).

Our case studies suggest that there can be two patterns of connecting to markets and sources of demand. The first one is that of the emerging countries mentioned above. Here the linkage is given by the relationships with the main market (the US), and the complementarity of the products of the clusters with the existing leading technologies, particularly those produced in the US themselves. The second model is the one epitomized by our case of the Scandinavian countries. The key here is to position the companies in areas not covered by the existing leaders, once again the US firms. The problem is even more complex because, as the Scandinavian stories in our project point out, the issue is not simply to cover products not produced by the leaders. In ICT, the leaders—and particularly the US firms—have been able to occupy entire vertical markets, and in most cases they have occupied them on a global scale. This has left little room for other global players to occupy even parts of these vertical markets. The only opportunity left was to occupy other vertical markets, with potentially global, or at least continental demand, which had not been occupied by the leaders (see Richards, 2001).

Some leading European firms, and particularly the Scandinavian companies Ericsson and Nokia, successfully pursued this strategy. In this respect, the study by John Richards (2001) in our project shows that the success of the Scandinavian model depended in good part on the same factors that we have highlighted in other cases—highly educated workforce, supply of technically and managerially skilled people, and connection to demand (particularly continental demand). The peculiarity of the model, however, is that it focused on a vertical market that was not occupied by the leading US firms, notably the wireless hardware segment of the ICT business. Moreover, as we shall also see later, the opportunities in this area were further raised by one of the most successful institutional European accomplishments in the ICT business during
the last decade, i.e. the establishment of the GSM standard for communication throughout the continent. This was a key event for creating a continental demand, which enhanced the growth of this segment and benefited the producers that had occupied this area.

The connections to market and to the leading areas worldwide also open up the question of the potential advantage of being 'out of town' that may accrue to newcomers or to latecomers. The point is that being 'out of town' may imply lower stakes in existing technologies and activities, and this may help new clusters to seek new sources of rents or to tap new product spaces and technologies. Of course, with the advantages of being 'out of town' for novelty come disadvantages for connectedness to world markets. One particular disadvantage for not being connected to the leading clusters, or to the leading centers of technological growth, is that of not being able to take advantage of some of the results of external effects, such as the availability of the 'right' labor. But this also explains the rise of some of the new regions that we have discussed in this paper. As noted, some of them were indeed 'out of town', but they all had local sources, or characteristics—such as excess supply of skilled labor—that made up for the disadvantage they faced.

Thus, on the one hand, countries like Israel, India, Ireland and Taiwan compensated for the disadvantages of being out of town by virtue of focusing on activities that were complementary to those of the United States. They then combined the advantages of being out of town, which enabled them to focus on new areas, with the ability to link with the sources of growth (whether demand or technical linkages and the like). On the other hand, the Scandinavian model implied that the companies had the opportunity (and possibly the fortune) of exploiting a new vertical market, and in this process they also built up the new linkages (technical, with demand, etc.) that were critical for the rise of the new cluster.

In these last several sections, we have emphasized the deep regularities cutting across all our study regions in the domains of human capital of two kinds: market connection and new firm formation. In emphasizing the deep similarities, we do not want to gloss over the considerable differences across these countries in institutions and in market mechanisms. These matter considerably for the ways in which the various deep capability-building goals were accomplished in different places.

5. Cooperation vs. Competition Among Clusters in the World Economy

The new clusters that emerged in the late 1990s have had to deal with the
dominance of the most important existing technology areas by US firms, largely firms located in the western US. From the perspective of rich places like Northern Virginia or Israel, or poorer places like Taiwan or Ireland, or vastly poorer places like India, linkages to established centers of technology are very important in the growth phase. Maintaining these linkages can be a difficult management problem for companies that have to do something new, but at the same time have to find ways to stay linked to the old. The best solutions appear to be participation in a worldwide production network that is an extension of the Silicon Valley network itself.

We have already noted the importance of finding a new area of technology or product area in which to succeed. This involves cooperation with existing firms and sales of complementary rather than competing products—at least in the short term. This pattern is not, however, a policy recommendation so much as it is an observation of why and how certain more narrowly imitative policies are likely to be ineffective. The real policy implications arise from thinking carefully about the particular sources of advantage for a nascent cluster and why that source might yield short-term complements with the potential to become long-term substitutes.

All the nascent regions we have examined have escaped from the belief that cooperation with existing richer economies is ‘colonialist’. Linkages with the US have been critical to all, in one way or another. India and Taiwan are linked to the US via outsourcing of software services and manufacturing respectively. Israel and Taiwan are also linked to the US by a returning group of expatriates who have worked there, and who see the benefits of long-distance collaboration. There is, in these cases, a flow of people and ideas back and forth between the existing cluster and the nascent ones. The diaspora has been particularly valuable to the nascent clusters that we have looked at in emerging markets. Technology transfer (in the narrow sense) is not very important in these cases. Rather what is transferred is primarily organizational models, a valuable piece of understanding for a nascent cluster, and the opportunity to apply that knowledge in a new domain.

This pattern of connection-led growth varies across the areas we study. India began simply by arbitraging the differences in engineering labor costs between that country and the US in outsourcing. There are no guarantees that such an activity will lead to entrepreneurship-led growth, though many valuable assets (such as knowledge of and connections to the US market and

---

13 The role of social structures and community ties may also be strong. This is particularly true within bodies of entrepreneur-managers who participate in a brain drain and then a reverse brain drain (Arora et al., 2001, Saxenian, 2001, 2002) or who have common experiences, e.g. in the military (de Fontenay and Carmel, 2001).
up-to-date technical skills) were gained in the process. Taiwan similarly began as a source of low-cost labor for manufacturing PCs, which had become too expensive to make in Silicon Valley. This was strongly complementary to Silicon Valley, and, combined with economic and institutional arrangements inside Taiwan which permitted and encouraged entrepreneurship, led to considerable growth of technical and market capabilities and ultimately an indigenous industry. Entrepreneurs in Ireland, Israel and Northern Virginia have been content to position themselves as niche players, producing in areas where they have an advantage (either individually or regionally) and thinking of themselves as linked to and complementary to Silicon Valley.

This pattern has also meant a significant departure from earlier relationships between the US and other countries’ high-tech industries. For example, the pattern looks different from the more competitive, and even confrontational, relationship between the US and Japan in semiconductors in the 1980s. The increasing competitiveness of Japanese semiconductor memory producers directly threatened the market position of the US producers. The case of Cambridge, UK provides another contrast. The Cambridge cluster looks more similar to Silicon Valley in terms of product space than the other emerging countries that we have examined. At the same time, while Cambridge has certainly grown, its growth rates have been steady and ‘normal’ compared to those of some of the emerging countries, and of Silicon Valley itself. This is why we think of Cambridge as a ‘partial’ success story when compared to such high-growth clusters.\(^\text{14}\) The products of firms in Cambridge were similar to those of Silicon Valley, with its first-mover advantage, and so could only cover spaces in the worldwide market that were not already exploited by the leader. By contrast, companies in the emerging countries engaged in complementary activities. This allowed them to avoid competing directly with Silicon Valley, while taking advantage of the growth rates of the latter through their complementary relationships.\(^\text{15}\)

The worldwide economic forces that have enabled these complementary patterns between the US and the emerging countries in ICT are important. First, the ICT industries grew more rapidly and in more varied a way than could have been anticipated. A number of technologies turned out to have significantly larger markets than originally anticipated—the PC and associated complementary hardware and software, and the internet and associated close complements come immediately to mind. Further, many of these new

\(^{14}\) Another possible explanation of Cambridge’s only ‘partial’ success is the lack of market access. No mechanism like returning expatriates created strong links to major markets and the domestic market was small (Saxenian, 1988).

\(^{15}\) See Athreye (2001) for details on the Cambridge, UK case.
technologies raised rather than lowered the demand in existing markets—the internet has dramatically increased the demand for PCs and for mainframe computers along with the demand for telecommunications transport and switching services, for example. All this has provided tremendous market and technical opportunities, and put pressure on the existing innovation-supporting resources, notably engineering labor and conveniently located land, in the existing successful clusters in the US.

Immigration and the physical expansion of Silicon Valley have provided partial offsets to these powerful forces, but both face real limits. Not everyone wants to live in the United States, and what had once seemed a vast expanse of cheap land used in fruiticulture was discovered to be a valley—with walls. Out of people and out of conveniently located land, Silicon Valley was in a position to cooperate with nascent clusters.

The countries and regions that responded come from the ‘also rans’ of economic development in the 20th century. Apart from the Scandinavian countries and Northern Virginia, all were relatively poor and peripheral. Why have the less advanced economies responded so well to these opportunities? Even more interestingly, why have many richer countries (e.g. the continental European countries and Japan) been less responsive than these ‘also rans’? The first part of the answer was largely given in the previous pages. In particular, our regions offered underemployed skilled labor and connections with the US that helped match their excess supply of skills with the excess demand for (ICT) skills in the US. While our analysis does not extend to countries like Germany, France and Japan, it is clear that they also offered ample, if not greater, supplies of both technical and managerial skills. A strong possibility for explanation is that the human capital in these other countries had substantially higher opportunity costs. A German engineer has lucrative opportunities in existing industries such as automobile, chemicals and electronics. The incentives to set up a risky new venture in a new industry could not easily match the gains from working in such established industries.

To summarize, as new firms or regions emerge in a context where leading firms, regions and countries exist, the former have advantages and disadvantages. The ‘also rans’ can exploit the opportunities created by the markets and institutions in the ICT sector, if they are well connected to them, without bearing the costs of creating them. However, because the major product spaces have been filled, they can only serve niche markets or focus on complementary activities. But what about the longer (or medium) term future? The initial successes imply that competencies are gradually accumulated. The longer (medium) term implications are not straightforward. Will there eventually be competition with the leaders? In the specific context we
are analyzing here (India, Taiwan, etc.) the situation is still fluid, and it is difficult to give definitive answers.

One cannot rule out a priori that some leading companies in the emerging countries will eventually compete with companies in the leading countries. The leaders, however, will probably keep moving up the frontier. The situation may turn out to be like the classical one in product markets wherein mature products are taken up by the followers, and the leaders develop more advanced technologies. However, other trajectories are possible. For example, firms in the new countries may specialize according to comparative advantage. The Indian firms, for instance, may further their initial comparative advantage of being able to organize large-scale projects for developing relatively lower end software at low costs. Comparative advantage has certainly played a role in Taiwan, which has developed sophisticated chip-manufacturing capabilities complementary to chip design in the US. The improved functioning of global markets may affect these processes as well. For example, it may be easier today to operate an international market in firms. This would favor the acquisition of companies overseas, and reinforce the existing division of labor, as noted earlier in the case of Israel.

6. Policy Issues and Conclusions

The core of good advice to policy is concrete understanding of the forces driving the economy. Two very different bodies of understanding have been brought to bear on high-tech, entrepreneurial-led growth. Should we understand the foundation of new regional clusters of entrepreneurial activity using new economy theories of increasing returns, or should we focus on comparative advantage? Our answer is an element of each. Much of the opportunity for new regions arises because old regions find themselves running up a steeply rising supply curve of land and of highly skilled labor. Even the very substantial migration of highly skilled labor to the existing Silicon Valley cluster has crashed into this classical diminishing-returns phenomenon. As a result, many of the policy implications are simple and classical: invest in education, have open market institutions, tolerate and even encourage multinationals, tolerate and even encourage a brain-drain. For similar reasons, we find no support for the wisdom of protectionist, infant industry, national champion or directive industrial policy programs. Yet that is not the end of the story. The mechanism by which entrepreneurial-led growth takes off and becomes a contributor to regional and even national development is one with a strongly increasing-returns flavor. Once clusters are founded, they do indeed deliver the kinds of opportunities emphasized in
the increasing-returns framework. But the big issue is how to start a cluster, and what can policy do (and not do) in this respect.

For one, our case studies clearly show the foolishness of directive public-policy efforts to jump-start clusters or to make top-down or directive efforts to organize them. Clusters of innovative activity do not respond well to being directed, organized or jump-started, entrepreneurship being a quirky thing. For example, the study by Scott Wallsten in our project (Wallsten, 2001) could not empirically find any effect of a public-sponsored program as important as the Small Business Innovation Research (SBIR) program in the US on high-tech employment in US counties (see also Wallsten, 2000). In contrast, accommodative government policies can be an important part of cluster development. Apart from public investments in areas like education, governments played an important supporting (though not leading) role in making entrepreneurship easier in many of our regions, notably in Ireland, Taiwan, Virginia and Israel.

A powerful policy lesson from all of our cases is that none of our clusters engaged in protectionism or ‘strategic trade policy’—even while attempting to capture greater shares of rent-producing activities or sectors. In fact, it was the openness of many of these regions to the world economy that has allowed success in the market-connected model. Rather than seek to offer firms a protected domestic market, which in the cases we study would have led to suboptimal scale, these regions have sought to define new niches and maintained a strong export orientation.

One of the most effective policies for regional development that we have encountered in our analyses has been that which encouraged the creation of the GSM standard. The GSM standard was critical for the growth of a continental wireless hardware market in Europe that benefited the Scandinavian firms first, and firms from other European countries later on. Yet the purposes of the policy were not set to encourage regional growth in Scandinavia. The rise of the Scandinavian model, built around the two leading firms Ericsson and Nokia, relied on several of the factors that we have already discussed for the other countries or regions—a substantial supply of skills and a strong policy for education; the role of the larger firms in encouraging the supply of skills; the ability to cover a niche (wireless hardware, and possibly wireless internet software and services) that has eventually become quite important. The formation of a European standard was, however, key for creating a sizable demand. Public policy was not the only factor that gave rise to the standard. It was, however, an important one, as with many standard setting processes (whether technical, legal, or else) that require the coordination of different actors. In sum, rather than interventionist policies, often directed toward the
formation of the clusters themselves, and of their firms, the creation of the conditions for the clusters to arise, like a continental demand, can be a much more effective mechanism for promoting the opportunities of growth.

A natural end to our journey into the potentially new micro-determinants of economic growth suggested by the cases of regions like Ireland, India, Israel or Taiwan is to summarize the key findings of our research. Notably, novel factors at the turn of the century are adding relevant new dimensions and relevant new determinants of growth. These include greater opportunities to exploit external effects, as well as economies of scale at the level of regions or industries, rather than mainly at the level of the firms. Yet, the importance of old economic factors, and related determinants of growth, has not vanished. Even within the context of the new models, old economy ingredients, like firm-building capabilities, connection to markets and demand, along with the supply of skills (both technical and managerial), are critical for the new economy forces (external effects, agglomeration economies, etc.) to reach their full potential. In particular, these factors are critical forces for the initial push that is required to start any new entrepreneurial-led growth process. Old economy factors are crucial for new economy outcomes.

This also speaks to public policy. Direct, top-down policies are most likely to fail. Particularly worrisome are policies that would direct at a level of detail such as picking the specific industries or technologies to be sponsored. The right policies have elements of a ‘benign neglect’, and they allow for a significant decentralization in the choice of the initiatives. They would focus instead on the enabling conditions like the creation of suitable demand and markets (including formation of standards), openness, competition for encouraging the success of skilled people and people with entrepreneurial ambitions, along with policies focused on key supply-side factors and institutions, and on education in the first place.

The stories of the emerging regions that we have studied in our project have also shown that one driver of entrepreneurship in these areas has been the low opportunity costs of local human capital. When human capital has high opportunity costs because of alternative productive uses (employment in large established firms in leading industries like ICT itself, automobile, chemicals, etc.), this is hardly an issue for public policy. The question is far more serious when the opportunity costs for local human capital comes from artificially high wages in relatively unproductive jobs—e.g. excessive levels of employment in public administration, or in intermediaries of various sort, which is typical for instance of countries like Italy or Japan. The right policies for new cluster formation will fight such ‘unproductive’ opportunity costs, to allow for the full blossoming of true opportunities for their human capital.
This paper draws on the results of a two-year research project, ‘Silicon Valley and its Imitators’, organized by the Stanford Institute for Economic and Policy Research (SIEPR), and carried out by an international and interdisciplinary team. In drafting this paper we benefited from specific, and most helpful comments by Scott Wallsten, along with several fruitful discussions with other team-mates of the project, and particularly Ashish Arora, Suma Athreye, Kevin Davis, Michael Horvath, Gordon Moore, John Richards and Salvatore Torrisi. A special acknowledgment goes to Ralph Landau for his intellectual suggestions and for his encouragement in pursuing this research. All errors remain ours.

References


Cluster Formation in the New Silicon Valleys


