

# Localization of Knowledge and the Mobility of Engineers in Regional Networks

Paul Almeida • Bruce Kogut

McDonough School of Business, Georgetown University, Washington, D.C. 20057

Department of Management, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

almeidap@gunet.georgetown.edu

---

**K**nowledge, once generated, spills only imperfectly among firms and nations. We posit that since institutions and labor networks vary by region, there should be regional variations in the localization of spillovers. We investigate the relationship between the mobility of major patent holders and the localization of technological knowledge through the analysis of patent citations of important semiconductor innovations. We find that knowledge localization is specific to only certain regions (particularly Silicon Valley) and that the degree of localization varies across regions. By analyzing data on the interfirm mobility of patent holders, we empirically show that the interfirm mobility of engineers influences the local transfer of knowledge. The flow of knowledge is embedded in regional labor networks. (*Knowledge; Spillovers; Mobility; Regions; Networks*)

---

Ideas, because they have no material content, should be the least spatially-bounded of all economic activities. Being weightless, their transport is limited only by the quality and availability of communication. Since ideas serve both as the inputs and outputs in their own production, their location need be constrained neither by the happenstance of the spatial distribution of raw materials, energy, and labor, nor by that of demand and markets.

Yet, there is good reason to believe that the production of ideas may be, contrary to its economics, prescribed within spatial boundaries. In his comparative analysis of nations published in *Industry and Trade* (1920), Alfred Marshall noted that economic activity was drawn to regions rich in the “atmosphere” of ideas. Vibrant regions are those that produce knowledge externalities that denote the spillover of ideas from innovating firms to other firms.<sup>1</sup> The existence of

these stable regions implies that these externalities are also localized; that is, they do not spill perfectly over spatial borders.

Yet, the economic treatment of externalities largely assumes them to be “there”—such as embodied in capital goods—rather than a property that itself deserves to be explained. Knowledge externalities, however, are not simply generated by a given technology. The relationship among firms, universities, star scientists, and engineers strongly conditions the extent by which knowledge spills over.<sup>2</sup>

The importance of regions in economic development has been a persistent, though often lost, theme in economic sociology. Jane Jacobs (1969) put forth an argument that the growth of cities is based on a positive cycle of linkages among industries; the social and economic linkages among diverse activities gen-

<sup>1</sup> We use externalities and knowledge spillovers interchangeably to denote the benefit of knowledge to people, or to firms, not responsible for the original investment in the creation of this knowledge.

<sup>2</sup> See Allison and Long (1987) for evidence that institutional affiliation provides a significant spur to productivity; also Crane (1965) on invisible colleges, and Brown and Deguid (1991) on communities of practice.

erate and sustain growth. In a seminal study, Annalee Saxenian (1994) carried out an ethnography of engineers in Silicon Valley (south of San Francisco) and Route 128 (which rings Boston), and attributed the success of the former to a more robust exchange of ideas among firms.

The relationship between social networks and the spatial localization of knowledge is usually neglected in economic studies on externalities. In an important exception, Jaffe et al. (1993) analyzed patent citation data pertaining to domestic university and corporate patents to test the extent of localization of knowledge spillovers. At three different geographic levels (country, state, and SMSA), they found evidence that patents citations tend to belong to the same geographic area as the originating patent (the patent they cite), even after controlling for the existing concentration of patenting activity. Their findings indicated that knowledge localization exists in the aggregate. Because they did not analyze the variation of localization by region or technology, they left open the issue of whether the properties of technology and institutions determine knowledge externalities.

We hypothesize that variations in regions influence the spatial character of knowledge externalities. This study applies the methodology of Jaffe et al. (1993) to investigate if and why a particular kind of knowledge (i.e., the design of semiconductor devices) is localized to particular geographic communities and *not* to others. A consideration of the sociology of localization addresses two questions previously left on the table, namely: Is there variation across economic regions? Does this variation occur because of differences in the sociology of the local labor networks and relations among firms?

Our analysis consists of two stages. In the first stage, we find that the localization of patentable knowledge varies across regions. Semiconductor knowledge in the Silicon Valley and, less so, New York triangle and Southern California tends to be localized but this is not true for other regions. We show through several diagnostic tests that these results are robust.

The second stage of research seeks to show how ideas are transferred through labor markets. By an examination of the mobility paths of patent-holders,

we trace the effect of inter-firm mobility on the pattern of patent citations. We show that ideas are spread, in part, by the mobility of patent-holders. This pattern suggests, along the lines of Jacob's (1969) argument, that localized knowledge builds upon cumulative ideas within regional boundaries. We offer the speculation that a driving force for local externalities in semiconductor design is the mobility of people.

## 1. Research Setting

The interdependence of technological accumulation and regions has marked the development of the semiconductor industry from its origins. The industry originated from the invention of the first solid state transistor at the laboratories of AT&T (Bell Labs) in New Jersey in 1947. Over the next five decades process and product technology in the semiconductor industry has advanced at a rapid pace while the industry has grown increasingly international.

Within the United States, interfirm linkages between domestic companies are common. Most firms, including Intel, Advanced Micro Devices, National Semiconductors, and Texas Instruments, have a history of alliances of various types with other semiconductor firms. Of the over 1800 recorded alliances in the industry between 1961 and 1989, nearly 1200 involved U.S. firms (Kogut and Kim 1992). Formal technology transfer arrangements have also helped to diffuse technology internationally. Japanese and European firms have both benefited from extensive strategic alliances with U.S. firms.

In addition to the formal transfer of technology, there is also impressive ethnographic evidence of the spread of knowledge through more informal channels that may differ by region. Since its inception, the American semiconductor industry has been characterized by interfirm mobility of scientists and engineers. Rival firms actively courted key engineers leading to extensive interfirm mobility of personnel (Rogers and Larsen 1984). Entrepreneurship has been another significant characteristic of the American industry. Ever since William Shockley left Bell Labs to start Shockley Semiconductors in Palo Alto, California, start-ups have played an important role in the diffusion of knowledge and the evolution of the industry (Moore

1986). Several of Shockley's assistants left his firm and formed Fairchild Semiconductors in 1957. The origins of almost every firm in Silicon Valley can be traced back to Fairchild. In addition to the role of the pioneering firms, universities played an important role. The area boasts two frontier universities in electrical engineering, University of California at Berkeley and Stanford University, which proactively pursued the diffusion of knowledge to the region (Leslie and Kargan 1996). The significance of university research for local diffusion is confirmed in several studies, notably Jaffe et al. (1993) and Zucker et al. (1994).

### Regional Differences

Saxenian's (1994) study presents a compelling and yet puzzling comparison of two of America's best-known regions of innovation in the electronics industry. These regions had similar histories, yet they face different futures—Silicon Valley is flourishing while Route 128 has stagnated in recent years. Saxenian argues that while Silicon Valley developed a system of collaboration and learning among small specialist companies, Route 128, dominated by a few large corporations, was slow to adjust to changes in markets and technologies. In effect, Saxenian puts forth the argument that the higher externalities among relatively smaller networked firms in the Silicon Valley leads dynamically to higher rates of innovation and productivity in the region. Angel (1991) conducted a survey of personnel at 67 semiconductor firms and found that those located in the Silicon Valley tended to hire more labor with substantial experience, suggesting a dynamic by which experience accumulates broadly in the region.

An important difference between the Silicon Valley and other regions, and the US and other countries, is the role played by start-ups. Of the 176 start-ups founded in the world semiconductor industry between 1977 and 1989, 88% were located in the United States, and 55% percent were located in Silicon Valley (Dataquest 1990). The study by Eisenhardt and Schoonhoven (1990) on regions and start-ups in semiconductors showed that survival rates of new firms in the Silicon Valley, despite its high density of activity, was not significantly different from other regions. In investigating why, they found that entrepreneurs in

Silicon Valley were more closely networked with venture capitalists and, to a lesser but still significant extent, with other firms.

## 2. Knowledge and Regional Networks

The ethnography of Saxenian poses the important question: Why should there be variation in the dynamic trajectory of these two regions? The two regions, Route 128 and Silicon Valley have the same industry (i.e., semiconductors), the same "cluster" (i.e., computers), and yet the dynamics by which innovations are created differ dramatically. In fact, by any standard economic logic, Silicon Valley appears disadvantaged for two reasons. First, the inability of a firm to establish property rights over knowledge should lead to a decrease in its willingness to invest in R&D. The possibility that free riders will benefit from the investment of other firms should lead to a vicious cycle that erodes the innovative investment of the region. Second, if inventive knowledge leaks across firms, it should leak across the boundaries of a region. In short, why should the spatial borders of a region be less permeable than the proprietary borders of a firm?

One obvious reason why knowledge should be regional is that it is held tacitly by skilled engineers who remain within the region. Studies on innovation point clearly to the importance of their comprehensibility, as Rogers (1983) and Winter (1987) argued, as a factor in their diffusion. The degree to which knowledge is not codifiable and is tacitly held by individuals have been found to be important determinants in the speed by which major innovations are transferred within and among firms (Kogut and Zander 1994, Zander and Kogut 1995).

An important aspect of diffusion is not only the inherent qualities of knowledge (i.e., whether it is tacit or easily imitated), but also whether there is a regional labor market for the engineers, scientists, and workers. Some regions appear as remarkable in this regard. In a memorable quotation, an engineer from the Silicon Valley observes that "people change jobs out here without changing car pools." Saxenian (1994) cites an engineer who claims that:

Here in Silicon Valley there's far greater loyalty to one's craft than to one's company. A company is just a vehicle which allows you to work. If you're a circuit designer, it's most important for you to do excellent work. If you can't succeed in one firm, you'll move on to another one.

The observations that innovative knowledge is held by individuals and that there are active labor market networks in some locations are critical components to explaining why the localization of ideas may vary by region. The central engine of this argument is the necessary conditions that the knowledge held by design engineers has a tacit quality and that these same engineers are mobile among firms *within* the spatial boundaries of a region. Parenthetically, differences in intellectual property law do not appear to explain differential mobility of engineers, and thereby the creation of externalities. Neither Silicon Valley nor other American regions differ in the common law treatment that the firm owns an employee's knowledge (Hyde 1997).

We test the assertion that mobility influences the creation of localized spillovers through two steps. First, we show that there are regional variations in the localization of knowledge. Second, we test whether the mobility of engineers holding major patents leaves a trace in the patent citation records and whether mobility varies by region. By tracking individual engineers, we link the stronger presence of externalities in the Silicon Valley to the movement of individual patent holders who remain within the region.

### 3. Data and Methods

For the following statistical analysis, we use patent citation analysis of important semiconductor innovations and apply a case-control methodology and regression analysis to test for the localization of knowledge.<sup>3</sup>

#### Patent Data and Citations

Patent documents provide data on the inventor and his location at the time of the invention, the owner

(assignee) of the patent (usually a firm), the time of the invention and also the technology of the invention. In addition, through patent citations, we are able to infer the technological influences on a particular invention.

The patent citations contained in a patent document have two possible sources: (a) the inventor and the patent lawyer and (b) the patent examiner. The patent applicant is obliged by law to specify in the application any and all of "the prior art" of which he or she is aware. Interviews with patent reviewers reveal that the examiner undertakes a thorough search of files to determine the patent's relationship to existing patents. In the final list, some citations represent direct technological influences on a particular innovation, while other citations may only represent indirect technological influences (since the patent examiner added them). Several studies (Albert et al. 1991, Carpenter et al. 1981, Narin et al. 1987) have shown that patent citation counts are a good indicator of the technological importance of an invention. Further, Trajtenberg (1990) in his study of CT scanners, showed that the number of citations to a patent serves as an indicator of social and economic value of the innovation as well. We therefore analyze only highly cited patents that tend to be of both technological and economic importance.<sup>4</sup>

#### Establishing the Regions

Major regions of semiconductor activity in the United States were identified by plotting actual plant locations of over 750 facilities and demarcating the corresponding regional clusters. (See Appendix for data sources.) The location of semiconductor plant clusters was confirmed through the use of county level establishment and employee data from County Business Patterns (corresponding to SIC 3674). Regions were demarcated around contiguous counties having two or more semiconductor establishments. The analysis reveals 18 regional clusters. We analyzed patents belonging to the top 12 regions. These 12 regions accounted for more than 95% of the highly cited patents.

<sup>3</sup> A patent is the grant of a property right to an inventor for an invention conferred by the government. A US patent is granted for an invention which is "useful," "novel," and "nonobvious to a person of ordinary skill in the art" (U.S. Department of Commerce 1992).

<sup>4</sup> We later show that the use of highly cited patents (as opposed to less cited patents) does not bias our study towards finding regional localization.

**Table 1** Descriptive Statistics

1a: Full Panels						
Panel	Number of Major Patents	Number of Citations		Mean Citations	% Self Citations	
1980	131	2371		18.1	10.5	
1985	172	2722		15.8	17.9	

  

1b: Distribution of Major Patents by Region						
Region	1980			1985		
	Number of Major Patents	Citations		Number of Major Patents	Citations	
		Number	Mean		Number	Mean
NY-NJ-PA	20	601	30.5	20	510	25.5
AZ	4	40	10	11	107	9.7
CO	7	126	18	5	47	9.4
FL	3	32	10.7	5	59	11.8
MA-CT	10	99	9.9	10	156	15.6
VT	11	145	13.2	9	109	12.1
OR-WA	4	31	7.8	12	126	10.5
NORTH CA	20	478	23.9	20	468	23.4
SOUTH CA	20	323	16.2	20	233	11.7
TX-DALLAS	20	324	16.2	20	334	16.7
TX-AUSTIN	6	64	10.7	20	280	14
TX-HOUSTON	6	108	18	20	293	14.7

### Identification of Major Patents

We first isolated patents related to semiconductor design with the help of experts in the Patent and Trademark Office. These patents belonged to two time periods—those patents filed in 1980 and those filed in 1985. For every patent in these two panels, we counted the total number of subsequent patents (up to 1995) that cited them. For each region and for each time period, we selected the top 25% most highly cited patents. Since this method generated an extraordinarily large number of patents for large regions, we took the 20 most highly cited patents as our sample. For some regions, the top 25% consisted of less than 20 patents. As a result, the samples for some regions had less than 20 major patents by which to generate the citations. To the degree that more highly cited patents are more or less geographically localized than less highly cited patents, this procedure potentially opens

the possibility of a sample selection bias. We show below that our results are robust to this concern.

### Descriptive Statistics of Major Patents

Table 1a displays descriptive statistics of the panels of major patents analyzed. Since not every region had 20 highly cited patents, the total number of major patents were 131 for the 1980 panel and 172 for the 1985 panel. Table 1b gives the distribution of major patents and the number of citations to these patents by U.S. region. While the larger regions had at least 20 highly cited patents, some regions such as Arizona and Florida had fewer.

### Citations and Controls

Localization is the use of knowledge created by others in the same region. Operationally, it is defined by the joint condition that the citing patent and the major innovation belong to the same geographic region. To

measure the frequency of localization, we geographically matched each major patent with the citing patents.

Clearly, the observed frequency of geographic coincidence of the major patent and the citing patents may also reflect the distribution of patenting activity (rather than the localization of spillovers). Silicon Valley has a lot of semiconductor firms, they patent a lot, and hence they cite each other a lot because the region dominates the overall patent count. To adjust for any bias due to this existing distribution of technological activity, we followed Jaffe et al. (1993) in the construction of a control sample. For each citing patent, we identified a corresponding control patent. This patent was identified such that the patent (technology) class was identical to that of the citing patent and the application date was as near as possible to the citing patent. Since the control patent does not cite the major patent, the frequency of a geographic match between the two reflects the existing concentration of patenting activity for a particular region. This frequency of geographic matches between the major patent and the control patent sets the baseline against which we compare the frequency of major patent-citing patent matches. This research design is very conservative. By use of controls, we isolate spillovers "above and beyond" agglomeration effects. Certain regions are unusually rich in their innovative activities, and this richness is itself suggestive of an externality. The controls capture this baseline agglomeration effect.

#### Statistical Test for Localization of Spillovers

Let  $P_{\text{cit}}$  be the probability that the major patent and citing patent are geographically matched, and  $P_{\text{con}}$  be the corresponding probability for the major patent-control patent match. Assuming binomial distributions, the null hypothesis is

$$H_0: P_{\text{cit}} = P_{\text{con}},$$

and the alternate hypothesis is

$$H_a: P_{\text{cit}} > P_{\text{con}}.$$

The  $t$  statistic is calculated as:

$$t = (P_{\text{cit}} - P_{\text{con}}) / [(P_{\text{cit}}(1 - P_{\text{cit}}) + P_{\text{con}}(1 - P_{\text{con}})) / n]^{0.5}.$$

The "t" statistic tests the difference between two independently drawn binomial proportions. We calculated the statistic for each of the 12 regions in each panel.

The case comparison controls for the effect of technology. By matching patents by their technological relatedness, we control for the differential degree of spillovers across a broad technological space. Podolny and Shepherd (1996) show that the evidence for spillovers is stronger among technologically dissimilar patents. By matching a control to the citing patent by technology class, our method conservatively eliminates the effects of technological distance and isolates the geographic dimension to diffusion.

#### Differences in Regional Localization

One of the major goals of this paper is to establish that externalities simply do not exist, but vary systematically by region. To demonstrate that these variations are statistically significant, we also test for differences in the extent of localization between regions.<sup>5</sup> For Region A,  $P_A$  represents the degree of localization of knowledge and is given by  $P_A = P_{\text{Acit}} - P_{\text{Acon}}$ .

For Region B,  $P_B$  represents the degree of localization of knowledge. We call  $P_B$  a baseline region for comparison. We test the hypothesis  $H_0: P_A = P_B$  against the alternate hypothesis  $H_a: P_A > P_B$ . The  $t$  statistic is calculated as:

$$t = (P_A - P_B) / [(P_A(1 - P_A) / n_A + P_B(1 - P_B) / n_B)]^{0.5}.$$

## 4. Testing for Regional Variations in Localization

The main results of the case-control tests for both the samples are given in Table 2a. The "Number of Citations" corresponds to the total number of cites for the major patents. "A" and "B" are the percentage of citations and controls, respectively, that belong to the

<sup>5</sup>We would like to thank a referee and Rebecca Henderson for encouraging us to augment our original sampling methodology by collecting sufficient original patents from each region in order to conduct these tests.

**Table 2** Test of Localization of Knowledge in US Regions (Significant Results at 0.05% in Bold)

2a: Results with All Cites								
Region	Number of Citations		A = Citation Matching %		B = Control Matching %		T-statistic	
	1980	1985	1980	1985	1980	1985	1980	1985
NY-NJ-PA	601	510	22%	29%	13%	8%	<b>4.51</b>	<b>8.81</b>
AZ	40	107	5%	7%	0%	1%	1.45	<b>2.19</b>
COLORADO	126	47	6%	11%	2%	2%	1.55	1.71
FLORIDA	32	59	3%	2%	0%	0%	1.02	1.01
MA-CT	99	156	7%	12%	5%	4%	0.60	<b>2.74</b>
VT	145	109	13%	14%	1%	1%	<b>3.95</b>	<b>3.75</b>
OR-WA	31	126	10%	5%	0%	2%	1.82	1.02
NORTH CA	478	468	27%	45%	9%	17%	<b>7.56</b>	<b>9.89</b>
SOUTH CA	323	233	16%	16%	4%	4%	<b>4.93</b>	<b>4.35</b>
TX-DALLAS	324	334	10%	19%	5%	8%	<b>2.84</b>	<b>4.19</b>
TX-AUSTIN	64	280	11%	10%	0%	4%	<b>2.80</b>	<b>2.91</b>
TX-HOUSTON	108	293	13%	20%	3%	2%	<b>2.83</b>	<b>7.08</b>
<b>TOTAL</b>	2371	2722	17%	22%	7%	7%	<b>11.31</b>	<b>16.32</b>

  

2b: Results Without Self-cites								
Region	Number of Citations		A = Citation Matching %		B = Control Matching %		T-statistic	
	1980	1985	1980	1985	1980	1985	1980	1985
NY-NJ-PA	502	385	12%	12%	11%	5%	0.30	<b>3.11</b>
AZ	37	96	0%	3%	0%	0%	0.00	1.76
COLORADO	115	46	1%	9%	1%	2%	0.00	1.39
FLORIDA	32	54	3%	0%	0%	0%	1.02	0.00
MA-CT	92	127	1%	2%	5%	2%	-1.67	-0.45
VT	103	74	0%	0%	0%	0%	0.00	0.00
OR-WA	26	112	0%	1%	0%	0%	0.00	1.00
NORTH CA	448	431	24%	43%	8%	16%	<b>6.71</b>	<b>9.09</b>
SOUTH CA	307	209	11%	8%	4%	4%	<b>3.34</b>	1.45
TX-DALLAS	303	235	7%	3%	3%	3%	1.88	0.00
TX-AUSTIN	57	242	4%	4%	0%	2%	1.44	0.79
TX-HOUSTON	78	166	0%	1%	0%	1%	0.00	0.00
<b>TOTAL</b>	2100	2177	11%	12%	6%	5%	<b>5.96</b>	<b>8.36</b>

same geographic region as the major patent. The *t*-statistic tests the equality of the control and citing proportions as described previously.

For the overall samples, there are significantly higher proportions of citation matches than control matches indicating localization effects. (Results significant at the 0.05 level or better are given in bold.) These results confirm the principal findings of Jaffe et al. (1993). The results indicate quite strongly that

knowledge is localized at the regional level. Silicon Valley shows the strongest localization effects, while the results for the Southern California, New York-New Jersey-Pennsylvania, Vermont, and all three Texas regions are also significant.

An examination of the underlying data reveals that localization may often be driven by self-citations (when the major patent and the citing patent have the same owner). Often the citations belonged to the same

plant (not just firm) and thus these citations did not represent cross-border knowledge flows. Further, several firms (especially Texas Instruments, IBM, and AT&T) have shown a strong propensity to cite their own patents, contributing to a strong (but perhaps misleading) localization effect.

We, therefore, run the analysis without including self-cites (Table 2b).<sup>6</sup> We find that only Silicon Valley exhibits strong localization effects and indeed contributes significantly to the overall localization findings. NY-NJ-PA and Southern California are each significant in one panel. The evidence qualifies the interpretation of the results of Jaffe et al. (1993). As with the earlier study, the above analysis indicates that, in the aggregate, patent citations tend to be localized. However, localization of knowledge is not a universal phenomenon. Geographic regions reveal different patterns in the local diffusion of knowledge externalities.

In Table 3 we test whether the degree of knowledge

<sup>6</sup> In order to treat the citing and control samples evenly, when the assignee of the control patent was the same as that of the major patent, the record was also considered a self-cite and excluded from the final sample.

**Table 3** Test of Localization of Knowledge Regional Differences

Region B	Results of <i>t</i> -Tests					
	Region A					
	NY-NJ-PA		North CA		South CA	
	1980	1985	1980	1985	1980	1985
NY-NJ-PA			<b>8.75</b>	<b>8.39</b>	<b>4.35</b>	-1.65
AZ	1.74	1.44	<b>9.26</b>	<b>8.56</b>	<b>4.87</b>	0.10
COLORADO	1.74	-0.07	<b>9.26</b>	<b>4.83</b>	<b>4.87</b>	-0.82
FLORIDA	-0.82	<b>5.06</b>	<b>3.67</b>	<b>12.60</b>	1.19	<b>2.69</b>
MA-CT	—	<b>7.43</b>	—	<b>13.96</b>	—	<b>4.30</b>
VT	1.74	<b>5.06</b>	<b>9.26</b>	<b>12.60</b>	<b>4.87</b>	<b>2.69</b>
OR-WA	1.74	<b>3.52</b>	<b>9.26</b>	<b>11.25</b>	<b>4.87</b>	1.61
NORTH CA	-8.75	-8.39			-3.91	-9.53
SOUTH CA	-4.35	1.65	<b>3.91</b>	<b>9.53</b>		
TX-DALLAS	-2.50	<b>5.06</b>	<b>6.33</b>	<b>12.60</b>	<b>2.15</b>	<b>2.69</b>
TX-AUSTIN	-1.18	<b>3.51</b>	<b>4.20</b>	<b>11.40</b>	1.28	1.47
TX-HOUSTON	1.74	<b>5.06</b>	<b>9.26</b>	<b>12.60</b>	<b>4.87</b>	<b>2.69</b>

*Note:* All figures are *t*-statistics for differences in regional localization; bold indicates significance at 0.05% or lower.

localization is significantly different across regions. Because three regions (i.e., Northern California, Southern California, and the New York-New Jersey-Pennsylvania) evidence significant degrees of localization (see Table 2), we use them to compare differences in localization across regions. The results show that knowledge is significantly more localized in Silicon Valley than in any other region (though the other two regions also evidence considerable localization).

### Tests for Sample Bias

It is interesting to note that the three regions showing significant localization are ones for which we have 20 patents. These three regions could be exhibiting localization because highly cited patents are more localized than less cited patents. We therefore test whether the localization findings are an artifact of the sampling procedure by comparing two samples of patents from Northern California (Silicon Valley) for the 1980 time period.

Our first sample consisted of the 20 highly cited patents for Northern California for the 1980 panel considered previously. We matched every patent from our "Highly Cited" sample with another randomly selected patent controlling for the region (Northern California), time period (1980 panel) and technology (same technology class). We thus had a second sample of "Other" patents that represent less highly cited patents. We then compared the frequency of regional citation matches between the two samples and tested to see whether the "Highly cited" sample was more localized than the "Other" sample.

The results of the test, available on request, indicate that the "Other" sample has, as expected, fewer citations but the frequency of local citation matches does not differ significantly across the samples. The result holds whether or not we include self-cites. These findings indicate that highly cited patents are not significantly more localized than less cited patents and therefore the sampling scheme used here does not introduce any bias.

### Tracing Diffusion by Tracing Patent-Holders in the Network

We hypothesized earlier that regions that are marked by spatially defined labor markets should evidence higher rates of localization. For semiconductor design



technologies, skilled engineers, some of whom hold major patents, hold significant knowledge. By compiling deeper data on patent-holders and through interviews, we collected information on the importance of individuals in the localization of knowledge. To measure the interfirm mobility patterns of semiconductor engineers, we developed a database of the career paths of semiconductor engineers with patenting records. For all the 438 individuals who hold major patents being analyzed, the career paths from 1974 to 1994 were traced through the records of their patenting activity. There were a total of 174 intraregional moves observed and 181 interregional moves. While some information on job changes is undoubtedly missed, the data are surprisingly revealing.

Table 4 shows the mobility patterns of these engineers at the regional level. The first column, "Total Years," is the summation of the number of years for which we counted moves for all the patent holders in our sample for the particular region. (For instance, if we observed 20 patent holders on average for 10 years each within Region A, the "Total Years" would be 200.) Moves are defined as the number of times that a major patent holder changes firms, as revealed in an analysis of all semiconductor patents. Thus for NY-NJ-PA we observed 48 moves by the patent holders. Of these 48 moves, 17 were within the region and 31 were from NY-NJ-PA to other regions. Next, to make comparisons across regions we standardized the

moves by calculating the "Moves per 100 Years." The last column provides the most important data. It represents the net intraregional moves of all major patent holders, standardized across regions.

The Silicon Valley is clearly unique in terms of inter-firm mobility. The level of intraregional mobility is very high, while extent of interregional moves is much smaller. Only MA-CT, Southern California and Texas-Austin show more intraregional movement than interregional movement, though to a much lesser extent than Silicon Valley. The table demonstrates considerably that regions differ to a considerable degree in the extent to which they facilitate interfirm job transfer through mobility.

### Regression Analysis

The simplicity of the means tests runs the objection that unobserved factors influence these results. In fact, the slight differences between the results for the two time periods suggest that changes in structural parameters (e.g. regional and institutional variables) over time influence the degree of localization. An alternative strategy is to test directly for the institutional effects on the generation of regional externalities. We investigate the factors influencing the localization of knowledge through a logistic regression analysis. (Variable definitions and data sources are given in Appendix 1; descriptive statistics and correlation tables are available on request.)

Agglomeration and size effects are measured through the number of establishments in the region and the density of design establishments per square mile in the county with the largest number of establishments.<sup>7</sup> We also included two mobility measures, defined to capture the intraregional turnover and the interregional turnover among the patent holders. In addition, a variable for the number of startups over the period in the region was added, since startups recruit, by definition, new engineers and managers. By controlling for start-ups, we can observe directly the effect of mobility independent of the opportunities for new employment.

<sup>7</sup> We took the density of the county with the most establishments due to the problem of comparing regions that differ widely in size, populated areas, etc.

**Table 4** Regional Mobility of Major Patent Holders

Region	Total Years	Moves		Moves per 100 Years		
		Intra	Inter	Intra	Inter	Net
NY-NJ-PA	920	17	31	1.85	3.37	-1.52
AZ	181	1	5	0.55	2.76	-2.21
CO	101	5	7	4.95	6.93	-1.98
FL	82	1	4	1.22	4.88	-3.66
MA-CT	240	9	7	3.75	2.92	0.83
VT	304	2	5	0.66	1.64	-0.99
OR-WA	156	4	10	2.56	6.41	-3.85
NORTH CA	750	76	23	10.13	3.07	7.07
SOUTH CA	568	21	16	3.70	2.82	0.88
TX-DALLAS	475	13	19	2.74	4.00	-1.26
TX-AUSTIN	297	9	7	3.03	2.36	0.67
TX-HOUSTON	243	1	15	0.41	6.17	-5.76

**Table 5** Logistic Analysis: Localization of Knowledge

Variables	Model 1 without self-cites	Model 2 without self-cites	Model 3 without self-cites	Model 4 all cites
Intercept	-4.2233** <i>0.3983</i>	-4.1801*** <i>0.3991</i>	-4.1817*** <i>0.4717</i>	-2.1679*** <i>0.2576</i>
Control	0.8374*** <i>0.1562</i>	0.8398*** <i>0.1561</i>	0.8372*** <i>0.1563</i>	0.7692*** <i>0.1221</i>
Regional Variables				
Density of Establishments	0.6022*** <i>0.0982</i>	0.5862*** <i>0.0984</i>	0.5969*** <i>0.1032</i>	0.3818*** <i>0.0543</i>
Intraregional Mobility	0.1535*** <i>0.0269</i>	0.1463*** <i>0.0273</i>	0.1414 <i>0.0791</i>	-0.1979*** <i>0.0427</i>
Interregional Mobility	-0.2204** <i>0.0805</i>	-0.2298** <i>0.0802</i>	-0.2193** <i>0.0806</i>	-0.0044 <i>0.0409</i>
Start-Ups	-0.2160** <i>0.0666</i>	-0.2110** <i>0.0667</i>	-0.2151** <i>0.0667</i>	-0.0582 <i>0.0477</i>
North California			0.1012 <i>0.6218</i>	1.9525*** <i>0.324</i>
Temporal Variables				
Lag	0.0619*** <i>0.0181</i>	0.0639*** <i>0.0181</i>	0.0619*** <i>0.0181</i>	0.0196 <i>0.0132</i>
Period 1980	-0.0831 <i>0.1511</i>	-0.1377 <i>0.1539</i>	-0.0938 <i>0.1515</i>	-0.3876*** <i>0.0878</i>
Universities				
Major Patent	0.3442 <i>0.2617</i>		0.3469 <i>0.2624</i>	0.4642 <i>0.2457</i>
Citing Patent	-0.3369 <i>0.5441</i>		-0.3357 <i>0.5441</i>	-1.0561 <i>0.5299</i>
Individual Variable				
Inventor Intraregional Mobility		0.2240* <i>0.1143</i>		
Maximum Likelihood	1294.91	1293.99	1294.90	2374.96
Number of Observations	4357	4357	4357	5093

Note: \* = Significance level of 0.05, \*\* = 0.01, \*\*\* = 0.001. Standard errors in italics.

To sort out possible period effects, we created dummies for the panels, plus added a time term for years lapsed since the original major patent. We control also for whether a university held the major or citing patent. The logit odds are coded 1 if the citing and major patent are from the same region; 0 if otherwise.

In Table 5, we report the results from this logit specification by estimating four models. The first

model gives the results for regional variables alone, plus the control (patent) variable, plus the temporal and university variables. (There is no change in the results when the model is reestimated absent the temporal and university effects.) It is important to keep in mind that the control patent variable already incorporates many of the unobserved variables to the regression. In all the regressions, the control variable is significant, as to be expected; it covaries with the

localization of the dependent variable. The interesting issue is what can be explained above and beyond the baseline expectation for localization.

The results of the specified models in Table 5 point quite clearly to the significant role played by agglomeration economies, mobility, and start-up activity. (To check for potential multicollinearity, we also estimated the model without start-ups; there is no change in the significance levels of the other variables.) The most interesting result is that the coefficient to intraregional mobility is positive but that to interregional mobility is negative. The institutional variables of universities were not significant. Clearly, the degree of mobility at the regional level is associated with the degree of localization.

To pinpoint more concretely the relationship of mobility to localization, we created an additional variable called "inventor intraregional mobility." This variable is a dummy that indicates if the inventor of the original patent subsequently moved within the region. As shown in the second regression, this variable is positive and significant at the 0.05 level. The parameters to the other mobility variables remain significant, indicating that the localization of knowledge is generally related to the mobility of top inventors within and between regions.

#### **Diagnostic Tests for the Regressions**

In the last two models given in Table 5, we report two diagnostic regressions to test the robustness of the results by including a dummy for Northern California (principally the Silicon Valley) and using self cites. Given the high correlation of 0.94 between the dummy for Silicon Valley and intraregional mobility, it is not surprising that the variables are not significant. A common test for whether an additional collinear variable adds explanatory power is to compare the likelihood scores with and without this new variable. Comparing the likelihoods of Models 1 and 2 shows no significant improvement. This test implies that intraregional mobility is equivalent to the Silicon Valley effect. Interesting enough, the coefficient to interregional mobility remains negative and significant, indicating that this variable captures a source for the loss of localization.

The last regression includes self-cites. Self-citing is

especially prominent among the larger firms, such as Texas Instruments and IBM, that populate the Texas and New York area regions. The results are very intuitive. Interregional mobility, though still negative, is no longer significant. We infer from this result that the larger firms, such as Texas Instruments, Intel, and IBM that have plants in several regions and account for most of the self-citations, build upon their own knowledge across regions, leading to the weaker negative coefficient result for interregional mobility. In this expanded data set, both the Silicon Valley and intraregional mobility variables are positive and significant. A particularly interesting change is the loss of significance for the time lag. In effect, large firms appear to build more rapidly upon their knowledge, as suggested in a shorter time to citation of their major patents. (The mean time to citation for the sample with self cites is 5.4 years; for a cite by another firm, the mean is 5.9 years.) Intraregional mobility generates local diffusion, but the process is less rapid than intrafirm diffusion.

In results not reported here, final diagnostics split the sample between cases in which the controls reveal localization and those that do not. By estimating the regressions on the split samples, we check for important interaction effects between the specified variables and the unobserved sources of localization embedded in the control variable. The results do not reveal any important interactions.

These results indicate clearly what is meant by the "above and beyond" effect of mobility on the localization of knowledge. Even after controlling for agglomeration and unobserved localization captured through the control variable, mobility still has a significant and positive effect on the probability that a patent will build upon a major patent from the same region. The implications suggest that the ability to build upon semiconductor design knowledge is tied significantly to the career paths of innovative individuals. The lower intraregional mobility of engineers appears to be related to the presence of large firms that build upon their knowledge in-house. However, these regions also show higher rates of departure of innovative engineers to other locations. These observations suggest that engineers in larger firms face a choice

between building careers within the internal labor markets or entering the external labor market, often by departing for other regions and carrying their innovative knowledge to new firms in new sites. A corollary to this speculation is that the entrepreneurial and high intraregional mobility through local labor markets is a factor behind the localized nature of the diffusion of innovation among firms confined to the Silicon Valley region.

## 5. Conclusions

One of the most important trends in the economics of research has been the diminishing role of the individual in patenting. This trend, long noted since Schmookler (1966), disguises the persisting importance of individuals in research. Across the landscape populated by laboratories and organized research, individuals appear as active agents in the creation and spatial diffusion of knowledge.

Ultimately, some appeal to institutions and the structure of relations is required to explain why certain regions show a higher degree of localization and in the ability to absorb and build upon previous knowledge. The many studies on technology transfer indicate the importance of social capability, prior experience, and access to the absorption of knowledge and its creation.<sup>8</sup> But these observations are fairly sterile unless understood in the context of the economic sociology of these regions or nations.

Externalities play a central role in economic theory, and yet are rarely studied. On examination, externalities are reflections of the nature of the knowledge held by individuals or groups in the context of specific social networks. To the mobile engineers of the Silicon Valley, the transfer of their knowledge and abilities is made through a partially visible network. Our results offer the speculation that externalities are the outcome of actions of skilled labor in spatially defined markets.<sup>9</sup>

<sup>8</sup> See Pack and Westphal (1986) and Rosenberg (1987).

<sup>9</sup> However, for some kinds of technologies, mobility does not seem to be important; see Argote et al. (1990) who found that learning by doing among shipyard workers did not transfer by rotation to other yards.

Because these markets differ geographically, regions also evince large differences in their social structures and stimulation for innovation. The ethnographies of Saxenian (1994) and Rogers and Larsen (1984), and the statistical results derived from patent citations, indicate that externalities are not created uniformly across all regions, nor are they natural by-products of particular technologies. Rather, externalities are created through the existence of broader social institutions that support a viable flow of ideas within the spatial confines of regional economies.<sup>10</sup>

<sup>10</sup> The authors are grateful to Peter Farkas, Patrick Abouchalache, Berlin Lai, and Jason Shrednick for their research assistance. They gratefully acknowledge the contribution of Mike Albert of CHI Research and thank Paul Allison, Tony Frost, Rebecca Henderson, Adam Jaffe, Paul Rosenbaum, Naren Udayagiri, and Sid Winter for comments and suggestions. Financial support for this project has been provided by the Huntsman Center for Global Competition and Innovation and the Reginald H. Jones Center at the Wharton School, and the Capital Markets Research Center, at the McDonough School of Business at Georgetown University.

## Appendix 1. Variable Definitions and Data Sources

The definition and sources of data used in the logit regression are given below:

- 1) Dependent variable: Regional match/no match between major patent and citing patent.
- 2) Control variable: Regional match/no match between control patent and major patent.
- 3) Density of establishments: Number of semiconductor plants per square mile in the largest county (in terms of establishments) in each region.
- 4) Intraregional Mobility: Number of intra-regional inter-firm moves per 100 years by patent holders. Interregional Mobility: Number of interregional, interfirm moves per 100 years by patent holders. Inventor Intraregional Mobility: Dummy variable coded 1, if major patent holder moved across firms within the region.
- 5) Start-ups: Log of total number of regional start-ups between 1975 and 1990.
- 6) Lag: Citation lag (Filed date of citing patent—Filed date of major patent).
- 7) Period 1980 and Period 1985: Dummy variables for 1980 and 1985 samples.
- 8) Universities: Major patent and Citing patent: Dummy for university patents.
- 9) Northern California: Dummy variable, coded 1 if major patent is from Northern California.

**Data Sources**

- 1) Patent data were obtained through the on-line patent database available on LEXUS-NEXUS and through CHI Research, a private research firm.
- 2) Plant locations were obtained from company reports and Dataquest (1990).
- 3) Establishment and employee data were obtained from County Business Patterns (1975–1990), a US Department of Commerce publication.
- 4) Data on start-up firms were obtained from Dataquest (1990).

**References**

- Albert, M. B., D. Avery, F. Narin, P. McAllister. 1991. Direct validation of citation counts as indicators of industrially important patents. *Res. Policy* **20** 251–259.
- Allison, Paul, Scott Long. 1987. Interuniversity mobility of academic scientists. *Amer. Sociological Rev.* **52** 643–652.
- Angel, D. P. 1991. High-technology agglomeration and the labor market: The case of Silicon Valley. *Environment and Policy* **23** 1501–1516.
- Argote, Linda, Sara Beckman, Dennis Epple. 1990. The persistence and transfer of learning in industrial settings. *Management Sci.* **36** 140–154.
- Brown, John Seely, Paul Deguid. 1991. Organizational learning and communities of practice. *Organ. Sci.* **2** 40–57.
- Carpenter, Mark, Francis Narin, Patricia Woolf. 1981. Citation rates to technologically important patents. *World Patent Inform.* **3**(4) 160–163.
- Crane, Diana. 1965. Scientists in major and minor universities: a study of productivity and recognition. *Amer. Sociological Rev.* **30** 699–714.
- Dataquest. 1990. *A Decade of Semiconductor Start-ups*. San Jose, CA.
- Eisenhardt, K. M., C. B. Schoonhoven. 1990. Organizational growth: founding teams strategy and environment and growth among US semiconductor ventures 1978–88. *Admin. Sci. Quart.* **35** 504–529.
- Hyde, Alan. 1997. How Silicon Valley has eliminated trade secrets (and why this is efficient). Mimeo. Rutgers University, New Brunswick, NJ.
- Jacobs, Jane. 1969. *The Economies of Cities*. Random House, New York.
- Jaffe, Adam B., Manuel Trajtenberg, Rebecca Henderson. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quart. J. Econom.* **108** 577–598.
- Kogut, Bruce, Dong-Jae Kim. 1992. Strategic alliances in the semiconductor industry. Unpublished report to Dataquest, The Wharton School, Univ. of Pennsylvania, Philadelphia, PA.
- , Udo Zander. 1994. Knowledge of the firm and the evolutionary theory of the multinational enterprise. *J. Internat. Bus. Studies* **24** 625–645.
- Leslie, Stuart, Robert Kargan. 1996. Selling Silicon Valley: Frederick Terman's model for regional advantage. *Bus. History Rev.* **70** 435–472.
- Marshall, Alfred. 1920. *Industry and Trade*. Macmillan, London, UK.
- Moore, Gordon E. 1986. Entrepreneurship and innovation: the electronic industry. R. Landau, N. Rosenberg, (eds.) *Positive Sum Strategy*. National Academy Press, Washington, DC.
- Narin, Francis, Elliot Noma, Ross Perry. 1987. Patents as indicators of technological strength. *Res. Policy* **16** 143–155.
- Pack, Howard, Larry E. Westphal. 1986. Industrial strategy and technological change: theory vs. reality. *J. Development Econom.* **22** 87–128.
- Podolny, Joel, Andrea Shepherd. 1996. When are technological spillovers local? Patent citation patterns in the semiconductor industry? Working Paper, Stanford University, Stanford, CA.
- Rogers, Everett. 1983. *The Diffusion of Innovations*, Free Press, New York.
- , Judith Larsen. 1984. *Silicon Valley Fever*. Basic Books, New York.
- Rosenberg, Nathan. 1987. Learning by using. *Inside the Black Box*. Cambridge University Press, Cambridge, UK.
- Saxenian, AnnaLee. 1994. *Regional Advantage*. Harvard University Press, Cambridge, MA.
- Trajtenberg, Manuel. 1990. A penny for your quotes. *Rand J. Econom.* **21** 172–187.
- US Department of Commerce. 1992. *General Information Concerning Patents*. Patent and Trademark Office, US Government Printing Office, Washington DC.
- Winter, Sid. 1987. Knowledge and competence as strategic assets. D. Teece, ed. *The Competitive Challenge—Strategies for Industrial Innovation and Renewal*. Ballinger, Cambridge, MA.
- Zander, Udo, Bruce Kogut. 1995. Knowledge and the speed of the transfer and imitation of organizational capabilities: an empirical test. *Organ. Sci.* **6** 76–92.
- Zucker, Lynne, Michael Darby, Marilynn Brewer. 1994. Intellectual capital and the birth of US biotechnology enterprises. Working Paper No. 4653, NBER, Cambridge, MA.

Accepted by Rebecca Henderson; received September 24, 1997. This paper has been with the authors 5½ months for 2 revisions.