The Geography of Research and Development Activity in the U.S.*

BY KRISTY BUZARD AND GERALD A. CARLINO

I n the U.S., metropolitan areas contain the largest concentrations of people and jobs. Despite some drawbacks, these so-called agglomeration economies also have benefits, such as the cost savings that result from being close to suppliers and workers. Spatial concentration is even more pronounced among establishments that do basic research and development (R&D). In this article, Kristy Buzard and Jerry Carlino show that geographic concentration of R&D extends beyond locations such as Silicon Valley. In fact, many types of R&D establishments are highly concentrated geographically.

Although metropolitan areas account for less than 20 percent of the total land area in the United States, they contain almost 80 percent of the nation's population and nearly 85 percent of its jobs. Put differently, the United States has, on average, 24 jobs per square mile, but metropolitan areas average about 124 jobs per square mile.

This high degree of spatial concentration of people and jobs leads to congestion costs, such as increased traffic and pollution, and higher housing costs. Congestion has become so severe in London that in February 2003, the city imposed a fee, currently £8 a day, on all vehicles entering, leaving, driving, or parking on a public road inside the Charging Zone between 7:00 a.m. and 6 p.m., Monday through Friday. New York City recently considered a similar plan. To offset these congestion costs, workers must receive higher wages, and higher wages increase firms' costs.

If congestion costs were the only thing resulting from the spatial concentration of firms, firms could easily disperse to reduce these costs. Yet they do not. This is because the negative effects of concentration make up only one side of the urban ledger. The positive effects of agglomeration economies — efficiency gains and cost savings that result from being close to suppliers, workers, customers, and even competitors — make up the other. Other things equal, firms will have little incentive to move if congestion costs are balanced by the benefits of agglomeration economies.

While economic activity tends to be geographically concentrated, spatial concentration is even more pronounced among establishments doing basic research and development (R&D). For example, although the United States has more than 3100 counties, the 50 counties that contain the largest number of R&D labs account for almost 60 percent of all such labs, while the top 50 counties in terms of the overall number of plants across all industries account for only about one-third of all plants.

More than most economic activity, R&D depends on a particular byproduct of agglomeration economies called knowledge spillovers — the continuing exchange of ideas among individuals and firms. The high geographic concentration of R&D labs creates an environment in which ideas move quickly from person to person and from lab to lab. Locations that

Jerry Carlino is a senior economic advisor and economist in the Research Department of the Philadelphia Fed. This article is available free of charge at www.philadelphiafed.org/research-and-data/publications/.

*The views expressed here are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

Kristy Buzard is currently a graduate student in economics at the University of California-San Diego. Previously, she was an economic analyst in the Research Department of the Philadelphia Fed.

Business Review Q3 2008 1

are dense in R&D activity encourage knowledge spillovers, thus facilitating the exchange of ideas that underlies the creation of new goods and new ways of producing existing goods.

Policymakers view the success of areas such as Silicon Valley in California, the Route 128 corridor in Boston, and North Carolina's Research Triangle as a miraculous recipe for local economic development and growth. But are these examples exceptions rather than the rule? The answer appears to be no. Equally remarkable concentrations may be found in many other types of R&D activity, such as the concentration of R&D in the pharmaceutical industry in northern New Jersey and southeastern Pennsylvania. In this article, we show that many types of R&D establishments are highly concentrated geographically.

CLUSTERING OF R&D LABS

Some studies have looked at the geographic clustering of economic activity in a particular industry, such as manufacturing or advertising. A study by Glenn Ellison and Edward Glaeser and one by Stuart Rosenthal and William Strange find evidence of geographic concentration of employment in many U.S. manufacturing industries. The geographic concentration of manufacturing jobs is not simply an American phenomenon, as Gilles Duranton and Henry Overman demonstrate in their analysis of manufacturing plants in the UK.

A study by Mohammad Arzaghi and Vernon Henderson looks at the location pattern of firms in the advertising industry in Manhattan. They report that Manhattan accounts for 20 percent of total national employment in the ad industry, 24 percent of all advertising agency receipts, and 31 percent of media billings. They show that for an ad agency, knowledge spillovers and the benefits of networking with

other nearby agencies are large but the benefits dissipate very quickly with distance from other ad agencies and are gone after roughly one-half of a mile.

Thomas Holmes and John Stevens take a broader approach. They used employment data for *all* U.S. industries, not just manufacturing, and

not for just a single industry, such as advertising. Among the 15 most concentrated industries, they find that six are in mining and seven are in manufacturing; only two industries fall outside mining and manufacturing (casino hotels and motion picture and video distribution).

Our article differs from past studies in two ways. First, rather than looking at the geographic concentration of firms engaged in the production of goods (such as manufacturing) and services (such as advertising), we consider the spatial concentration of private R&D activity.[1] Second, rather than focusing on the concentration of employment in a given industry, we look at the clustering of individual R&D labs.[2] To do this, we used 1998 data from the *Directory of American Research and Technology* to electronically code the addresses and other

information about R&D labs. These data were not available in a machine-readable format. Since the directory lists the complete address for each establishment, we were able to assign a geographic identifier (using geocoding techniques) to 3,129 R&D labs in the U.S. in 1998.[3]

> Policymakers view the success of areas such as Silicon Valley in California, the Route 128 corridor in Boston, and North Carolina's Research Triangle as a miraculous recipe for local economic development and growth. But are these examples exceptions rather than the rule? The answer appears to be no.

A map of the spatial distribution of R&D labs reveals a striking clustering of this activity (Figure 1). In places that have little R&D activity, each dot on the map represents the location of a single R&D lab. For example, there is only one lab in Montana, represented by the dot in Flathead County. In counties with a dense clustering of labs, the dots tend to sit on top of one another, representing a concentration of labs. A prominent feature of the map is the high concentration of R&D activity in the Northeast corridor, stretching from northern Virginia to

[1] There are a number of other studies that look at innovative output across cities, such as the study by David Audretsch and Maryann Feldman. What is unique about our article is that we present information on local *private* R&D activity, which no one else has done.

[2] The study by Paulo Guimarães, Octávio Figueiredo, and Douglas Woodward is one of only a few other studies we are aware of that look at spatial clustering at the establishment level. Specifically, they look at the geographic concentration of over 45,000 plants in 1999 for concelhos (counties) in Portugal. Duranton and Overman use plant-level data to study the locational pattern of UK manufacturing industries.

[3] Our data on individual labs were limited to the top 1,000 U.S. public companies in terms of R&D expenditure in 1999. The 1,000 firms cover more than 95 percent of all R&D performed by public companies. Many of these firms have multiple labs. For example, the Lockheed Martin Corporation has 54 labs, and General Electric has 26.

FIGURE 1

Location of Total R&D Labs*
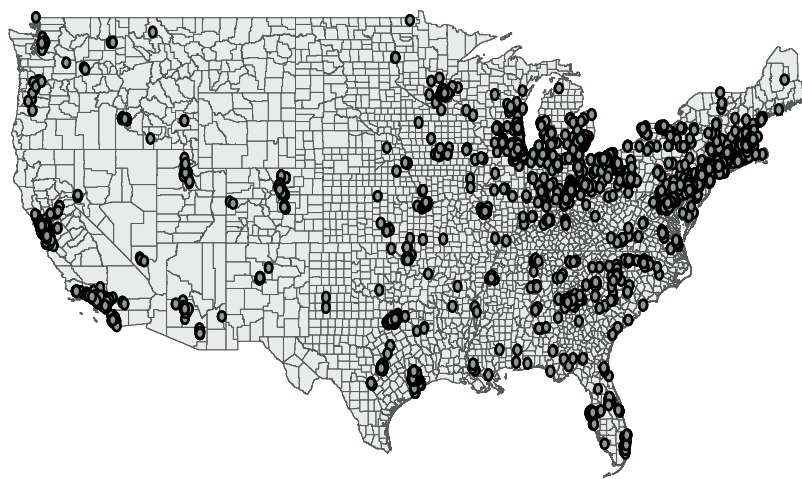
* In counties with relatively little R&D activity, the dots on the map represent the location of a single R&D lab. In counties having a dense concentration of labs, the dots represent a concentration of labs.

Massachusetts. There are other concentrations, such as the cluster around the Great Lakes and the concentration of labs in California's Bay Area and in southern California. But some states that account for a relatively large share of the nation's jobs account for a much smaller share of the nation's R&D labs. For example, Texas ranks second among states in terms of employment, but it ranks eighth in the number of R&D labs. Similarly, Florida ranks fourth in employment, but 13th in the number of labs.

However, as already noted, recent studies have shown that economic activity, especially manufacturing, also tends to be geographically clustered. We will show that R&D activity tends to be more spatially concentrated than total employment or manufacturing employment. There are 3,141 counties in the U.S., and all of them are engaged in some type of economic activity. All but 33 counties are en-gaged in some form of manufacturing activity. In contrast, only 519 of these counties have at least one R&D lab, and far fewer counties have a notable concentration of labs.

A simple way to quantify the concentrations of R&D relative to establishments in general or to manufacturing establishments in particular is to first compute each county's share of total R&D labs and rank counties by descending order of this share. Moving down this ranking, we compute a cumulative total for the share of R&D labs. Next, we construct a similar ranking for establishments in general and for manufacturing establishments in particular. The top 50 counties ranked by number of R&D labs account for 58 percent of all R&D labs, while the top 50 counties ranked by number of manufacturing establishments account for only 36 percent of all manufacturing establishments and only 32 percent of all establishments. It

appears that R&D labs are more highly concentrated than economic activity in general and overall manufacturing activity in particular. This is important because it means the concentration of R&D labs doesn't simply reflect the concentration of manufacturing activity. Since R&D is more concentrated than manufacturing activity, this suggests that some factors, such as knowledge spillovers, may be a more centralizing force for R&D than they are for manufacturing activity.

WHICH R&D LABS CLUSTER?

Paul Krugman and David Audretsch and Maryann Feldman developed a "locational Gini coefficient" to answer the question of which manufacturing industries cluster geographically. A locational Gini coefficient shows how similar (or dissimilar) the location pattern of employment in a particular manufacturing industry is from the location pattern of overall manufacturing employment. It does this by subtracting a county's share of national employment in manufacturing from the county's share of national employment in a given manufacturing industry, squaring the result, and summing over locations to arrive at a single number. The squaring of the difference in shares means that larger differences contribute more than proportionately to the overall value of the index. If the squared difference takes a value of zero, employment in a particular industry is allocated across counties in exactly the same way as employment in manufacturing. That is, this would indicate that employment in a given manufacturing industry is no more or less geographically concentrated than overall manufacturing employment. At the other extreme, the locational Gini coefficient takes on values close to one when employment in a given industry is completely concentrated in one county.

Glenn Ellison and Edward Glaeser have identified a potential problem with the locational Gini coefficient. They argue that if an industry consists of a *small* number of *large* establishments, the locational Gini coefficient may take on large values, suggesting localization of the industry even if there is no agglomeration force behind the industry's location. They refer to this as the dartboard approach to geographic concentration, using the metaphor of a few darts tossed at a dartboard randomly creating a cluster. Ellison and Glaeser have developed an alternative concentration measure — called the Ellison-Glaeser, or the EG, index — that controls for an industry's organization.

Recently, Paulo Guimarães, Octávio Figueiredo, and Douglas Woodward (GFW) have generalized the EG index to include the case where the data are in the form of establishments (labs, in our case) rather than employment shares, as in the EG index. The GFW locational Gini, or the GFW index, for R&D labs is constructed just like the locational Gini for employment except each county's share of the nation's labs in a given industry is used instead of the county's employment share for the industry. As before, the GFW index for a given industry takes on a value of zero when R&D labs in the industry are not geographically more concentrated than is manufacturing employment. Following Ellison and Glaeser, Guimarães, Figueiredo, and Woodward adjust the GFW index to account for the industrial organization of the industry in question.

We use the adjusted GFW index as our measure of concentration for R&D by industry.[4] We find an adjusted GFW index of 0.0457 for R&D in the average industry at the *county* level. In studying the agglomeration patterns in the manufacturing industries, Glenn Ellison, Edward Glaeser, and William

Kerr report an average adjusted Gini coefficient of 0.03 for manufacturing in 1997 at the *metropolitan area* level. (Since metropolitan areas tend to be aggregates of counties, there are more counties than metropolitan areas.) Thus, our R&D labs appear to be more spatially concentrated, on average, than is manufacturing activity.[5]

Our findings indicate that 256, or 68 percent, of all R&D counties have an adjusted GFW index greater than zero, suggesting that R&D labs are appreciably more concentrated than manufacturing employment. Earlier we reported that the top 50 counties ranked by number of R&D labs account for 58 percent of all R&D labs, while the top 50 counties ranked by number of manufacturing establishments account for only 36 percent of all manufacturing establishments. Thus, the concentration of labs is broadly similar when looking at the top 50 counties or the adjusted GFW index.

While an adjusted GFW index for an industry could have a value greater

than zero, an important question is: Does this represent a *significant* departure from the spatial concentration of manufacturing employment? We performed a simulation procedure to determine what value of the adjusted GFW indexes constitutes a significant departure from the concentration of manufacturing employment.[6] We find R&D labs in 129 of the 376 industries considered (34.3 percent) are significantly more concentrated than is manufacturing employment. Thus, of the 256 industries with an adjusted GFW index greater than zero, only about one-half — or 129 industries — represent a significant departure from the overall concentration of manufacturing employment. This shows the importance of providing statistical tests that determine whether labs in a

---

[4] See the article by Paulo Guimarães, Octávio Figueiredo, and Douglas Woodward for details on the construction of the adjusted GFW index used in our article as well as a discussion of the EG index. Our sample consists of 376 four-digit Standard Industrial Classification industries at the county level. We chose to do our analysis based on the number of labs in a county rather than employment in these labs, since we have data on employment for only about one-half of the labs in our data set.

[5] By construction, the value of both the EG index and the adjusted GFW index is directly related to the level of aggregation of the geographic area under consideration. That is, for any given industry, the EG indexes and the adjusted GFW indexes take on larger values for metropolitan areas (aggregations of counties) than the indexes do at the county level. Thus, our finding of greater average concentration of R&D labs compared with the average concentration of manufacturing employment reported in Ellison, Glaeser, and Kerr is even more striking, given that the adjusted GFW is calculated at the county level and still exceeds the average value of the EG index calculated at the MSA level.

[6] To develop measures of statistical significance for the adjusted GFW indexes, we partitioned our industries into six nonoverlapping groups based on the number of R&D labs in a given industry. The first group consists of industries that have between two and nine labs. The second group consists of industries with 10 to 30 labs, while the third group consists of industries with between 31 and 50 labs. The fourth group consists of industries with between 51 and 100 labs, while the fifth group consists of industries with 101 to 200 labs. The final group consists of industries with more than 200 labs. For each group, we performed a simulation procedure to produce a probability distribution for the adjusted GFW index. In the simulation we randomly allocated labs to counties while maintaining the counties' share of national manufacturing employment. Therefore, if a given county has a relatively high share of the nation's manufacturing jobs, the county is more likely to randomly be assigned more R&D labs, too. For each group the simulation produces a value for the adjusted GFW index. For each group, we performed 1,000 simulations and formed a probability distribution for the adjusted GFW indexes. From the distribution we can calculate critical values (one that's positive and one that's negative) that allow us to say that we are 95 percent certain that any value that exceeds the critical value indicates that labs in that grouping are significantly more concentrated than is the actual distribution of manufacturing employment. Similarly, any value that falls below the critical value indicates that labs in that grouping are significantly more dispersed than is the actual distribution of manufacturing employment.

given industry are significantly more concentrated (significantly more dispersed) than is the actual distribution of manufacturing employment.

Our measure of concentration, the adjusted GFW index, has a maximum value of about one for R&D in five industries.[7] However, there are only two R&D labs in each of these industries, so it's not surprising to find a large value for the adjusted Gini index if the two firms are located in proximity to one another.[8] Among industries with 20 or more labs, R&D tends to be most concentrated in the oil and gas field machinery industry, the computer storage devices industry, and the electronic computer industry (see the Table).[9]

Until now, we have looked at the concentration of R&D labs relative to the concentration of manufacturing employment. We would also like to know whether labs in a particular industry (such as pharmaceuticals) are more or less concentrated than overall

---

[7] They are hog production; the production of brooms and brushes; the production of fiber cans, tubes, and drums; the bottled and canned soft drinks and carbonated waters industry; and the rolling mill machinery and equipment industry.

[8] There is a negative relationship between the size of the adjusted GFW index and the number of labs in an industry. However, this relationship is not strong: a correlation coefficient of -0.09 that is only marginally significant (at the 10 percent level).

[9] In this article, our index of concentration (the adjusted GFW index) compares the concentration of R&D labs in a given industry to the concentration of manufacturing *employment* in that industry. Instead of using manufacturing employment as the benchmark when constructing the adjusted GFW index, we could have used manufacturing *establishments* as the benchmark. In general, there's a moderate correlation (a Spearman's rank correlation coefficient of 0.56) between the industry ranking under the two alternative benchmarks for R&D industries with significant adjusted GFW indexes and with 20 or more labs. Following Guimarães, Figueiredo, and Woodward, we report the adjusted GFW index using manufacturing employment as the benchmark in this article to make our findings consistent with past studies, such as the one by Ellison and Glaeser.

TABLE

Concentration of R&D Labs for Selected Industries

Industry	Number of Labs	Adjusted GFW Index[a]
Concentrated Industries[b]		
Oil & Gas Field Machinery	22	0.33
Tires and Tubes	14	0.12
Crude Petroleum & Natural Gas	14	0.10
Computer Storage Devices	34	0.08
Motor Vehicles & Car Bodies	26	0.06
Electronic Computers	57	0.06
Semiconductors	278	0.03
Prepackaged Software	359	0.03
Motor Vehicle Parts	134	0.03
Optical Instruments and Lenses	36	0.03
Computer-Integrated Systems Design	105	0.02
Radio and TV Communication Equipment	185	0.02
Dispersed Industries[c]		
Wood Household Furniture	11	-0.01
Gaskets, Packing, and Sealing Devices	11	-0.01
Industrial Valves	14	-0.01
Plastic Plumbing Fixtures	11	-0.01
Gray and Ductile Iron Foundries	12	-0.01

---

[a] The adjusted GFW index for a given industry shows the sum of the squared differences of the share of employment in manufacturing from the share of labs in a given industry, adjusted to account for the industrial organization of the industry under consideration.

[b] R&D labs in the selected industries are significantly more concentrated than manufacturing employment (5 percent level of significance).

[c] R&D labs in the selected industries are significantly more dispersed than manufacturing employment (5 percent level of significance).

R&D labs. To get this information, we recalculated the adjusted GFW index to reflect the geographic concentration of labs in individual industries relative to the overall concentration of R&D labs (as opposed to the overall concentration of manufacturing employment). We find that 314, or 84 percent, of all R&D labs have an adjusted Gini index greater than zero; however, we find that R&D labs in only 105 of the 376 industries (28 percent) considered are significantly more concentrated than overall R&D labs.[10] It's not surprising to find less concentration of R&D by industries when the comparison is to overall R&D labs than when the comparison is to overall manufacturing employment (34.3 percent), given that R&D labs already tend to be more concentrated than manufacturing employment. Still, for the majority of industries (72 percent), labs at the industry level tend not to be more spatially concentrated than labs overall.

Maps of R&D activity for individual industries (for example, software, Figure 2; pharmaceuticals, Figure 3; and chemicals, Figure 4) confirm the findings of the adjusted GFW indexes in that the location pattern of R&D activity for the majority of industries is broadly similar to the location pattern of overall R&D activity. That is, R&D activity for most industries tends to be concentrated in the Northeast corridor, around the Great Lakes, in California's Bay Area, and in southern California.

_____

[10] We performed a simulation procedure to determine what value of the adjusted GFW indexes constitutes a significant departure from the concentration of total R&D labs. The simulation procedure is similar to the procedure used when the reference was manufacturing employment, except we now randomly allocate labs to counties while maintaining the counties' share of national R&D labs, as opposed to the counties' share of national manufacturing employment.

FIGURE 2
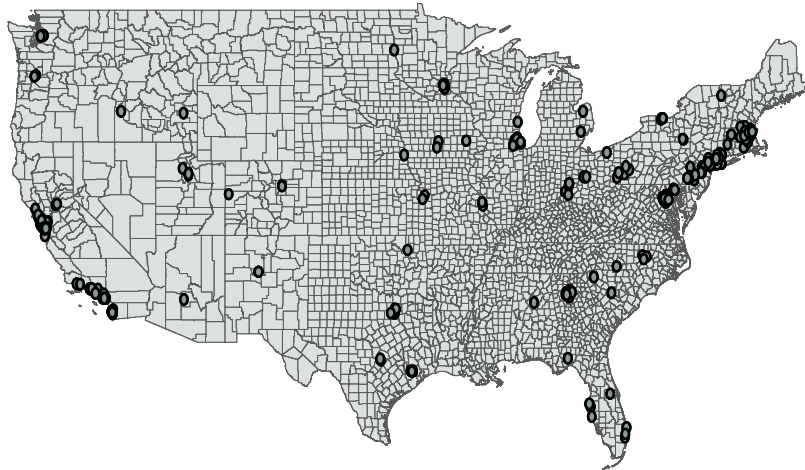
Location of Software R&D Labs
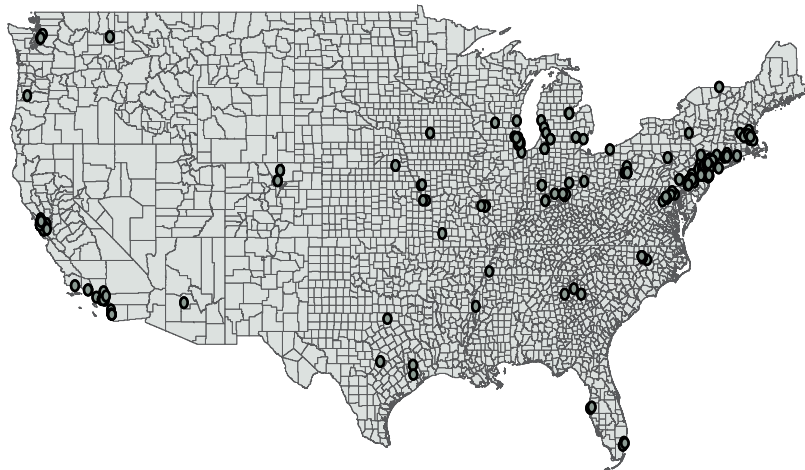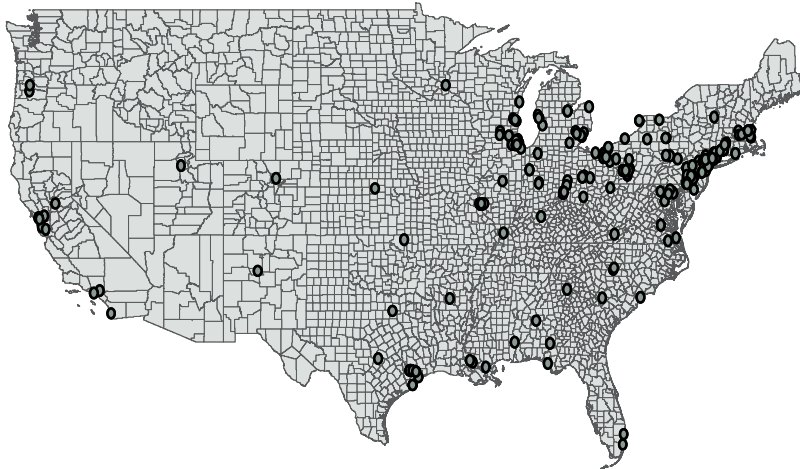

FIGURE 3

Location of Pharmaceutical R&D Labs

FIGURE 4

Location of Chemistry R&D Labs

FIGURE 5

Location of Oil and Gas Field Machinery R&D Labs

As indicated, there are a number of exceptions to the general pattern of geographic concentration just described. One exception is R&D activity in the oil and gas field machinery industry, which tends to be concentrated in Texas, especially in the Houston area, and accounts for about 60 percent of the labs doing R&D in this industry (Figure 5). Another exception is the location of R&D activity in the motor vehicle and car body industry, which tends to be concentrated in Michigan, especially in the Detroit area, and which accounts for just under 40 percent of the labs doing R&D in this industry (Figure 6). This industry comprises establishments primarily engaged in manufacturing motor vehicle parts and accessories.
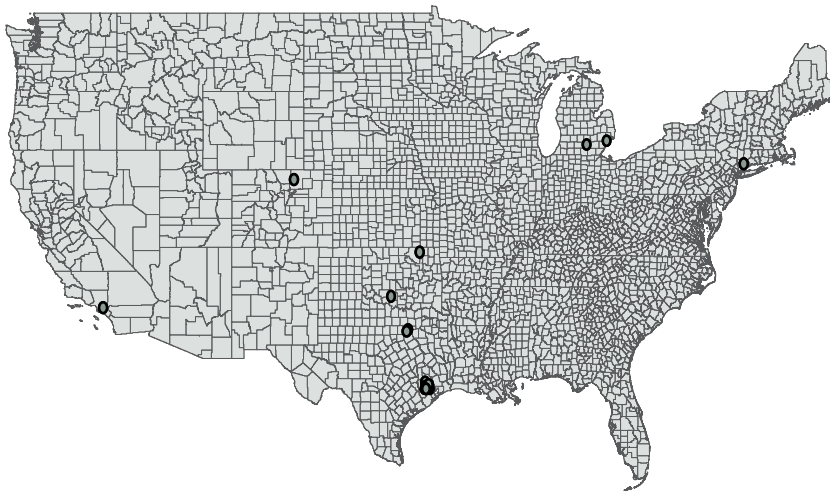
WHY DO R&D LABS CLUSTER?

Economists have developed a number of theories to explain firms' tendency (not just R&D labs) to cluster. Firms may attempt to minimize transport costs by locating close to a natural resource used as an input, or to their suppliers, or to their markets. Or firms may cluster to share inputs such as specialized workers. Finally, firms may cluster to take advantage of knowledge that "spills over" when firms are located near one another. Among these, the sharing of inputs and especially of knowledge spillovers is likely to be most important for R&D firms when choosing a location.

Knowledge Spillovers. Economists have identified two types of knowledge spillovers thought to be important in understanding the location pattern of R&D labs: MAR spillovers and Jacobs spillovers.[11] While these
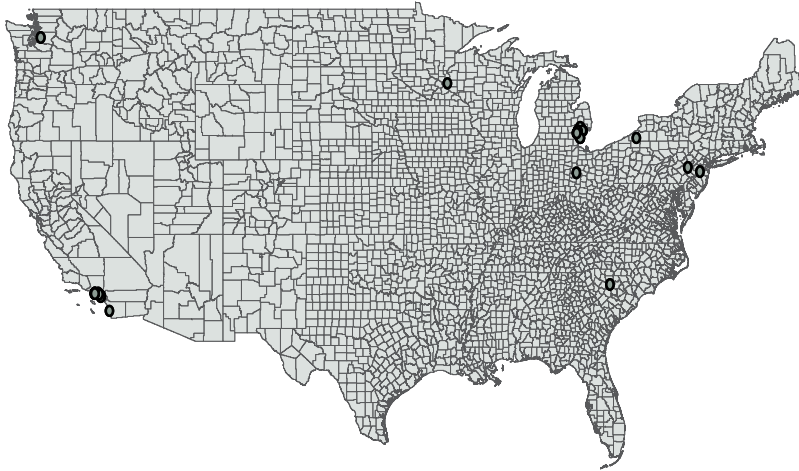
---

[11] MAR spillovers are so-called because in 1890 Alfred Marshall developed a theory of knowledge spillovers that was later extended by Kenneth Arrow and Paul Romer — hence, MAR. In 1969, Jane Jacobs developed another theory of knowledge spillovers.

FIGURE 6

Location of Motor Vehicle and Car Body R&D Labs

theories were originally developed to explain the concentration of industries in general, we think they are particularly important to an explanation of the clustering of R&D labs. More than most industries, R&D depends on new knowledge. Often, the latest knowledge about technological developments is valuable to firms but only for a short time. Thus, it behooves firms to set up shop as close as possible to the sources of information. The high spatial concentration of R&D activity facilitates the exchange of ideas among firms and aids in the creation of new goods and new ways of producing existing goods.

MAR spillovers. According to the MAR theory of spillovers, the concentration of establishments (labs in our case) in the same industry in a common area helps knowledge travel among labs and their workers and facilitates innovation and growth.[12] Employees from different establish-

ments in the same industry exchange ideas about new products or new ways to produce goods. Often, knowledge is tacit and not easily codified and therefore requires face-to-face contact to be effectively transmitted. Having firms concentrated in a particular area is an efficient way to produce new ideas, leading to innovation and growth. People's ability to receive ideas or knowledge is then influenced by their distance from the source of the ideas; communicating ideas is harder over longer distances. Stuart Rosenthal and William Strange consider the importance of input sharing, matching, and knowledge spillovers for manufacturing firms at the state, county, and ZIP code levels. They find that the effects of knowledge spillovers on the agglomeration of manufacturing firms tend to be quite localized, influencing agglomeration only at the ZIP code level.[13]

For example, many semiconductor

firms have located their R&D facilities in the Silicon Valley because the area provides an environment where semiconductor firms can develop new products and new production technologies. Often, information about current developments in the semiconductor industry is shared informally. In her 1994 book, AnnaLee Saxenian describes how gathering places, such as the Wagon Wheel Bar located only a block from Intel, Raytheon, and Fairchild Semiconductor, "served as informal recruiting centers as well as listening posts; job information flowed freely along with shop talk." Other examples include the Route 128 corridor in Massachusetts, the Research Triangle in North Carolina, and biotechnology and medical technology software firms in suburban Philadelphia.

Jacobs spillovers. Jane Jacobs believed that knowledge spillovers are related to the diversity of industries (diversity of labs in our case) in an area, in contrast to MAR spillovers, which focus on firms in a common industry. Jacobs argued that an industrially diverse environment encourages innovation. Such environments include knowledge workers with varied backgrounds and interests, thereby fa-

---

[12] Edward Glaeser, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer, who coined the term MAR spillovers, pulled these various views on knowledge spillovers together in their article.

[13] Several other studies have found that knowledge spillovers dissipate rapidly with distance. See, for example, the articles by Mohammad Arzaghi and J. Vernon Henderson; David Audretsch and Maryann Feldman; Wolfgang Keller; and Jed Kolko. The extent to which innovations in communication technologies are rendering face-to-face contacts obsolete is not so clear. Jess Gaspar and Edward Glaeser argue that improvements in telecommunications technology increase the demand for all interactions. So while technology may substitute for face-to face contact, this effect is offset by the greater desire for all kinds of interactions, including face-to-face contact.

cilitating the exchange of ideas among individuals with different perspectives. This exchange can lead to the development of new ideas, products, and processes.

As John McDonald points out, both Jane Jacobs and John Jackson have noted that Detroit's shipbuilding industry was the critical antecedent leading to the development of the auto industry in Detroit. In the 1820s, Detroit mainly exported flour. Because the industry was located north of Lake Erie along the Detroit River, small shipyards developed to build ships for the flour trade. R&D in the shipbuilding industry led to refinements and the adaptation of the internal-combustion gasoline engine to power boats on Michigan's rivers and lakes. As it turned out, the gasoline engine, rather than the steam engine, was best suited for powering the automobile. Several of Detroit's pioneers in the automobile industry had their roots in the boat engine industry. For example, Olds produced boat engines, and Dodge repaired them. In addition, a number of other industries in Michigan supported the development of the auto industry, such as the steel and machine tool industries. These firms engaged in R&D that led them to produce many of the components required to make cars.

While other factors could be at work, the adjusted GFW indexes appear to support Jacobs' diversity view, in that R&D labs for the vast majority of industries (almost three-quarters) tend to exhibit a common overlapping pattern of concentration. David Audretsch and Maryann Feldman used the U.S. Small Business Administration's innovation database and focused on innovative activity for particular industries within specific MSAs. They found less industry-specific innovation in MSAs that specialized in a given industry, a finding that also supports Jacobs' diversity thesis.

The Role of Natural Advantage. While it's tempting to argue that the broadly similar geographic clustering of R&D labs in many different industries is suggestive of Jacobs externalities, this conjecture is simply based on visual inspection of a map (Figure 1). Jacobs spillovers are one possible way to account for the common overlapping pattern of concentration among R&D labs, but other forces might be at work.

Universities are key players not only in creating new knowledge through the basic research produced by their faculties but also in supplying a pool of knowledge workers on which R&D depends.

One such source is the natural advantages an area offers to firms that locate there. An area's natural advantages, such as climate, soil, and mineral and ore deposits, could explain the location of some R&D labs. For example, oil deposits, an essential ingredient for testing equipment, may be largely responsible for the concentration of R&D labs in the oil and gas field machinery industry (one of the most highly concentrated industries, according to our adjusted GFW indexes) in Texas, especially in the Houston area. But the draw of ore deposits seems to be industry-specific and is therefore unlikely to account for the common overlapping pattern of concentration among R&D labs in many different industries. Of course, if R&D labs tend to be drawn to areas offering amenities such as pleasant weather, proximity to the ocean, and scenic views, this could explain the overlapping concentration in amenity-rich locations, such as the concentrations found in California. While local amenities might explain some of the concentrations of labs, the vast majority of R&D labs tend to be highly concentrated in the country's

Rust Belt region, an area relatively low in amenities.

Another natural advantage that an area may have lies in its workers and institutions, especially its universities. Universities are key players not only in creating new knowledge through the basic research produced by their faculties but also in supplying a pool of knowledge workers on which R&D depends. It is well known that Silicon Valley and the Route 128 corridor became important centers for R&D as a result of their proximities to Stanford and MIT. AnnaLee Saxenian describes how Stanford's support of local firms is an important reason for the Silicon Valley's success. Two of Stanford's star engineering professors, John Linvill and Fred Terman, not only drew some of the best and brightest students to Stanford, but they also trained their students (and encouraged them) to seek careers in the semiconductor industry.

There is also evidence that an area's human capital can be an important type of natural advantage. In a 2007 paper, Gerald Carlino, Satyajit Chatterjee, and Robert Hunt looked at the effect of a metropolitan area's human capital (the share of the adult population with at least a college education) on the area's ability to innovate (measured by patents per capita). Of the things these authors considered, by far the most powerful effect on local innovation is generated by local human capital. Specifically, a 10 percent increase in the share of the adult population with at least a

college degree is associated with an 8.6 percent increase in patents per capita. Since the share of a metropolitan area's population with at least a college degree varied by a factor of almost six in the sample used in Carlino, Chatterjee, and Hunt's paper, the implied gains in innovation are substantial.

There is also general evidence that R&D at local universities is important for firms' innovative activity. David Audretsch and Maryann Feldman, and Luc Anselin, Attila Varga, and Zoltan Acs found evidence of localized knowledge spillovers from university R&D to commercial innovation by private firms, even after controlling for the location of industrial R&D. However, Carlino, Chatterjee, and Hunt found that R&D at local universities has only modest effects on local innovative activity. They found that a 10 percent increase in R&D intensity of local universities is associated with less than a 1 percent increase in patent intensity.

Evidence on MAR vs. Jacobs Spillovers. To more formally address the issue of the importance of industrial diversity, or, alternatively, specialization, we conducted a simple experiment. Recall that we have only one adjusted GFW index for each industry. These industry indexes can, however, be used to construct an *overall* adjusted GFW index for each metropolitan county. This is done by weighting each industry's adjusted GFW index by the share of the county's total establishment accounted for by that industry. The industry-weighted adjusted GFW indexes for a given county are then summed to arrive at an overall adjusted GFW index for each metropolitan county. The overall adjusted GFW index for a county can be correlated with a widely used index of industrial diversity.[14] By construction, a county is said to be more highly specialized or less diversified as the value of the diversity index increases. Recall that as

the value of the adjusted GFW index increases, the extent of the spatial concentration of labs in the industry also increases. A positive correlation between the overall county adjusted GFW index and the specialization index means that as the county becomes more specialized industrially, its labs are also becoming more geographically concentrated. This evidence favors MAR spillovers.[15]

On the other hand, if the geographic concentration of labs tends to increase as the specialization index decreases — indicating that an area is more industrially diverse (or less specialized) — this negative correlation provides evidence in favor of Jacobs spillovers. We found a positive and highly significant correlation between the overall county adjusted GFW index and the specialization measure, evidence favoring MAR spillovers.

While a more definitive conclusion awaits a more complete analysis, the evidence provided in this article tends to support the importance of both Jacobs spillovers (visual inspections of maps) and MAR spillovers (statistical correlation) for R&D labs.[16]

CONCLUSION

Most countries make sustained economic growth a principal policy objective. Although many factors contribute to economic growth, recent research has found that innovation and invention play an important role. Innovation depends on R&D, and R&D depends on, among other things, the exchange of ideas among individuals. The high spatial concentration of R&D labs creates an environment in which ideas move quickly from person to person and from lab to lab. That is, locations that are dense in R&D activity encourage knowledge spillovers, thus facilitating the exchange of ideas that underlies the creation of new goods and new ways of producing

existing goods.

Finally, the study by Saxenian provides a cautionary note for policymakers who view the success of areas such as Silicon Valley as a recipe for local economic development and growth. While investing in science centers to attract R&D activity is fairly common in the U.S., Saxenian's study suggests that creating the right corporate culture to make the centers successful is more challenging. Instead of targeting industries, we suggest that policymakers consider strategies that help to establish a good business environment and which are conducive to attracting and retaining highly skilled workers. Glaeser and co-authors' study suggests that local policymakers need to focus on life-style issues because they are important in attracting and retaining high-skill workers. One such policy is providing good public schools. Other policies might focus on reducing urban crime and providing amenities such as clean streets and public parks. **BR**

---

[14] County-level specialization was measured using a Herfindahl index. A Herfindahl index measures diversification or, inversely, specialization. It is calculated by squaring and summing the share of establishments accounted for by each industry in a given county. The squaring of industry shares means that the larger industries contribute more than proportionately to the overall value of the index. Thus, as the index increases in value for a given county, this implies that the county is more highly specialized or less diversified industrially.

[15] We have 847 metropolitan counties in our sample. The correlation coefficient is 0.0148 and is significant at the 1 percent level. The coefficient is small in magnitude because the average value for the diversity index is 75 times as large as the average value for the county adjusted GFW index. Despite the relatively low value of the correlation between the county adjusted GFW index and the diversity index, the relationship between these variables is economically significant, displaying an elasticity of almost one in value.

[16] A more complete analysis of the role of MAR vs. Jacobs spillovers on the clustering of R&D labs should also control for an area's natural advantages as identified in this article.

REFERENCES

Anselin, Luc, Attila Varga, and Zoltan Acs. "Local Geographic Spillovers between University and High Technology Innovations," *Journal of Urban Economics*, 42 (1997), pp. 442-48.

Arzaghi, Mohammad, and J. Vernon Henderson. "Networking Off Madison Avenue," unpublished manuscript (2005).

Audretsch, David B., and Maryann P. Feldman. "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86 (1996), pp. 630-40.

Carlino, Gerald A. "Knowledge Spillovers: Cities' Role in the New Economy," Federal Reserve Bank of Philadelphia *Business Review* (Fourth Quarter 2001), pp. 17-23.

Carlino, Gerald A. "The Economic Role of Cities in the 21st Century," Federal Reserve Bank of Philadelphia *Business Review* (Third Quarter 2005), pp. 9-15.

Carlino, Gerald A., Satyajit Chatterjee, and Robert M. Hunt. "Urban Density and the Rate of Invention," *Journal of Urban Economics*, 61 (2007), pp. 389-419.

Directory of American Research and Technology, 23rd Edition. New York: R.R. Bowker, 1999.

Duranton, Gilles, and Henry G. Overman. "Testing for Localization Using Micro-Geographic Data," *Review of Economic Studies*, 72 (2005), pp. 1077-1106.

Ellison, Glenn, and Edward. L. Glaeser. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 105 (1997), pp. 889-927.

Ellison, Glenn, Edward L. Glaeser, and William Kerr. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," Discussion Paper 2133, Harvard Institute of Economic Research (April 2007).

Feldman, Maryann P., and David B. Audretsch. "Innovation in Cities: Science-Based Diversity, Specialization, and Localized Competition," *European Economic Review*, 43 (1999), pp. 409-29.

Gaspar, Jess, and Edward Glaeser. "Information Technology and the Future of Cities," *Journal of Urban Economics*, 43 (1998), pp. 136-56.

Glaeser, Edward, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer. "Growth in Cities," *Journal of Political Economy*, 100 (1992), pp. 1126-53.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. "Consumer City," *Journal of Economic Geography*, 1 (2001), pp. 27-50.

Guimarães, Paulo, Octávio Figueiredo, and Douglas Woodward. "Measuring the Localization of Economic Activity: A Parametric Approach," *Journal of Regional Science*, 47 (2007), pp. 753-44.

Holmes, Thomas J., and John J. Stevens. "Spatial Distribution of Economic Activities in North America," in J.V. Henderson and J.-F Thisse, eds., *Handbook of Regional and Urban Economics, Vol. IV: Cities and Geography*. Amsterdam: Elsevier, 2004.

Jackson, John. "Michigan," in R. Scott Fosler, ed., *The New Role of American States*. New York: Oxford University Press, 1988.

Jacobs, Jane. *The Economy of Cities*. New York: Vintage Books, 1961.

Keller, Wolfgang. "Geographic Localization of International Technology Diffusion," *American Economic Review*, 92 (2002), pp. 120-42.

Kolko, Jed. "Agglomeration and Co-Agglomeration of Services Industries," unpublished manuscript (April 2007).

Krugman, Paul. *Geography and Trade*. Cambridge: MIT Press, 1991.

Madden, Janice. "Creating Jobs, Keeping Jobs, and Losing Jobs: Cities and the Suburbs in the Global Economy," unpublished manuscript (2000).

Maurel, Françoise, and Béatrice Sédillot. "A Measure of the Geographic Concentration in French Manufacturing Industries," *Regional Science and Urban Economics*, 29 (1999), pp. 575-604.

McDonald, John F. *Fundamentals of Urban Economics*. Upper Saddle River, NJ: Prentice Hall, 1997.

Rosenthal, Stuart, and William C. Strange. "The Determinants of Agglomeration," *Journal of Urban Economics*, 50 (2001), pp. 191-229.

Saxenian, AnnaLee. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press, 1994.

Starr, Paul. "Review of AnnaLee Saxenian," *Regional Advantage Contemporary Sociology* (May 1995).