

AGGLOMERATION OF INVENTION IN THE BAY AREA: NOT JUST ICT

By CHRIS FORMAN, AVI GOLDFARB, AND SHANE GREENSTEIN *

* Forman: Scheller College of Business, Georgia Institute of Technology 800 West Peachtree Street NW, Atlanta GA 30308 (e-mail: chris.forman@scheller.gatech.edu). Goldfarb: Rotman School of Management, University of Toronto, 105 St. George St., Toronto ON M5S 3E6, CANADA (e-mail: agoldfarb@rotman.utoronto.ca). Greenstein: Harvard Business School Harvard University, Morgan Hall 439, Boston MA 02138 (e-mail: sgreenstein@hbs.edu). Max Koven and Yasin Ozcan provided excellent research assistance. We thank Maryann Feldman and seminar participants at the University of Toronto, Stanford University, and the 2016 American Economic Association meetings for helpful comments. Avi Goldfarb thanks SSHRC for research support.

Does invention agglomerate, and if so, where does it agglomerate? In this paper we examine changes in patterns of agglomeration in invention over time, using data on patent applications from all granted US patents.

There are plenty of reasons to expect invention to agglomerate. Carlino and Kerr's (2015) recent handbook chapter summarizes many such results, emphasizing the role of input sharing, labor market matching, and knowledge spillovers, among others. Knowledge spillovers received an especially large fraction of attention in their chapter, and in the literature overall (e.g., Audretsch and Feldman 1996; Jaffe, Trajtenberg, and Henderson 1993; Moretti 2012; Kerr and Kominers 2015).

Simple economics might forecast that most invention agglomerates in the same area as the

primary using industry (Carlino and Kerr 2015). For example, patents related to automotive technology are clustered in Detroit (Hannigan, Cano-Kollmann, and Mudambi 2015). Or causality could be reversed: The location of a break-through invention can lead to industry agglomeration and localized follow-on invention (Duranton 2007). We label this “colocation” between invention and industry production.

However, other forces push away from colocation. Invention itself is an economic activity and it shares inputs, such as specialized labor institutions, particular intellectual property contracts, and information spillovers from one type of invention to another. If such forces are strong, they could lead to agglomeration of many types of invention in one place. We call this “coagglomeration of invention.” For many industries, the key inventions could be in a location distinct from the place where production for the downstream using industries reside.

Using patent data to measure invention, there are two approaches to investigate colocation and coagglomeration of invention. One is to map the agglomeration of downstream

industries and invention and measure the geographic correlation. We take another approach. We look for evidence of coagglomeration of invention – namely, invention from distinct areas appearing in the same location, irrespective of downstream using industry.

We find evidence consistent with the hypothesis of coagglomeration. We demonstrate a strong trend toward the clustering of patenting in the San Francisco Bay Area from 4 percent of US patents in 1976 to 16 percent of US patents in 2008, a time period when the fraction of the US population in the Bay Area did not increase substantially relative to the US population as a whole.¹ While this increase in Bay Area patenting is partly driven by the increasing fraction of patents in information and communication technologies (ICTs), ICTs cannot fully explain the trend. The San Francisco Bay Area has seen a substantial increase in its share of patents, even for patents that seem quite distant from ICTs, rising from 3.9 percent of such patents in 1976 to 6.9 percent in 2008.²

Our results are consistent with coagglomeration of invention in the Bay Area.

While others have documented a tendency toward agglomeration of patenting by industry, we believe we are the first to document a general tendency toward agglomeration in patenting across industries and patent classes. Further, our study is unique in its documentation of agglomeration in one particular region, the Bay Area.

Coagglomeration in production and invention has been documented in other settings and other industries. For example, Rosenberg (1963) analyzes sewing machines, bicycles, and automobiles located in Northern Ohio and southeastern Michigan and their sharing of common inventions in machine tools, and shows how growing downstream industries induced additional improvements in those innovations over time. Glaeser (2005) discusses coagglomeration of production across many industries in New York City, starting in the nineteenth century. A number of recent researchers have explored the causes and consequences of coagglomeration, including Ellison, Glaeser, and Kerr (2010), Helsley and Strange (2015), and Delgado, Porter, and Stern (2014).

¹ According to the US Census, the Bay Area grew from 2.54 percent of the US population in 1980 (5.74 million residents) to 2.65 percent of the US population in 2010 (8.15 million residents).

² One unusual aspect of patenting in the San Francisco Bay Area is that invention is not centered in the city but in Silicon Valley. Therefore, while we refer to other cities by the city names, we refer

to the “Bay Area” rather than “San Francisco” to describe the San Francisco Consolidated Metropolitan Statistical Area. We use the 2013 definition, which includes the following 12 counties and their FIPS codes: Alameda (06001), Contra Costa (06013), San Francisco (06075), San Mateo (06081), Marin (06041), Santa Clara (06085), San Benito (06069), San Joaquin (06077), Sonoma (06097), Solano (05095), Santa Cruz (06087), and Napa (06055).

At this point, our results do not provide a definitive cause of this broad increase in coagglomeration in invention. A variety of mechanisms are possible including regulation, financing, shared labor markets across invention types, and knowledge spillovers across invention types.

Of course, we are not the first to document the agglomeration of economic activity in ICT in the Bay Area. Garcia-Vicente, Garcia-Swartz, and Campbell-Kelly (2014) show that such agglomeration began to arise in the 1980s and 1990s. Our results are consistent with this timing. A variety of authors have explored the reasons behind the agglomeration of the ICT industry in the Bay Area and its dynamics in generating new firms and new ideas (e.g., Saxenian 1994; Franco and Mitchell 2008; Chen et al 2010; Marx, Singh and Fleming 2015; Kerr and Kominers 2015). Our contribution relates to the finding of the increasing role of the Bay Area in patenting overall.

I. Data and empirical strategy

We use patents granted by the US Patent and Trademark Office (USPTO) as our measure of invention. Because of the delay between patent application and grant date, we date patents using the year of application. We have data on patents granted between 1976 and 2012, and

our analysis data set includes patents with application dates between 1976 and 2008. We cut off the last four years of the data because of lags between year granted and year filed. Generally, we start to see a decline in patenting in 2008, suggesting right truncation may be an issue for the last few years of our data. The trends we identify appear long before 2008.

Patents have been shown to provide a useful measure of a firm's intangible stock of knowledge (Hall, Jaffe, and Trajtenberg 2005). Their limitations are well known. Not all patents meet the USPTO criteria for patentability (Jaffe and Trajtenberg 2002), and not all inventors seek to patent. Further the propensity to patent has changed over time during our sample (e.g., Hall and Ziedonis 2001), this was particularly the case for patents related to software which grew rapidly toward the end of our sample period due to changes that strengthened their legal rights (e.g., Graham and Mowery 2003; Hall and MacGarvie 2010). We are comfortable with using patents in this context because our primary focus is on changes in the geographic distribution of patenting within broad technology areas over time. While the propensity to patent has changed across patent classes over time, we rely on the assumption that it has not changed significantly across geographic locations.

We map inventors to counties and MSAs using the zip code of the location of the inventor. We used consolidated MSAs (CMSAs) where those were present. This will be particularly important for our analysis of the Bay Area.

For the analysis that follows, we do not weight by citations. For multi-author patents, we divide by the number of authors. For example, if a patent has 1 author in the Bay Area and 2 authors in Boston, it would count as 1/3 of a patent in the Bay Area and 2/3 of a patent in Boston. Our results are generally robust, and often stronger, using three-year and five-year citation-weighted measures. For example, using either three- or five- year citation weights, the Bay Area surpasses New York City as the location with the most patents three years earlier than with the unweighted measure.

Our analysis requires us to identify patents that represent inventions related to ICT, or inventions that draw upon the stock of knowledge related to ICT. As is well known, identifying such inventions through the patent data is difficult (see, e.g., Graham and Mowery 2003; Bessen and Hunt 2007; Hall and MacGarvie 2010). As a result, we use different definitions based on the primary class of the patent and explore the robustness of our results to alternatives.

We present our results at the year level, as aggregated means over the 33 years from 1976 to 2008 inclusive. In particular, our results are presented as graphs of time trends of the fraction of patents each year that meet some criteria such as location in the Bay Area. This is therefore a descriptive exercise that tests whether the results are consistent with increasing coagglomeration of invention in the San Francisco Bay Area over time.

II. Results

A. Patenting Across Locations

Given the overall rise in the propensity to patent, all major urban areas had an increase in the number of patents. We explore the fraction of all US patents by area, thereby controlling for the overall trend.

Figure 1 shows the increasing importance of the Bay Area as a fraction of US patenting. It compares the top 10 areas in the United States, defined by the total number of patents between 1976 and 2008. In 1976, the New York City area was the dominant center for patenting, with just under 15 percent of all patents. Generally, patenting was highly correlated with population. The Bay Area rose steadily as a fraction of patenting in the 1970s and 1980s, and then the trend increased in the 1990s before settling down at the earlier rate of increase in

the 2000s. In 1995, the Bay Area surpassed New York City as the US location with the largest number of patents.

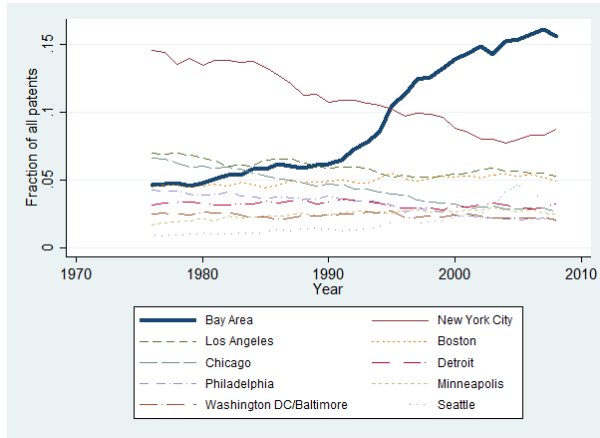


FIGURE 1. FRACTION OF PATENTS IN TOP 10 CITIES

Figure 2 combines locations into four groups: the Bay Area, New York City, the 18 other cities in the top 20, and all other locations. Generally while New York and locations outside the top 20 are falling as a proportion of patenting, the Bay Area is rising quickly, and the other 18 cities in the top 20 are rising slightly (42.6 percent in 1976 to 46.1 percent at the peak level in 2004).

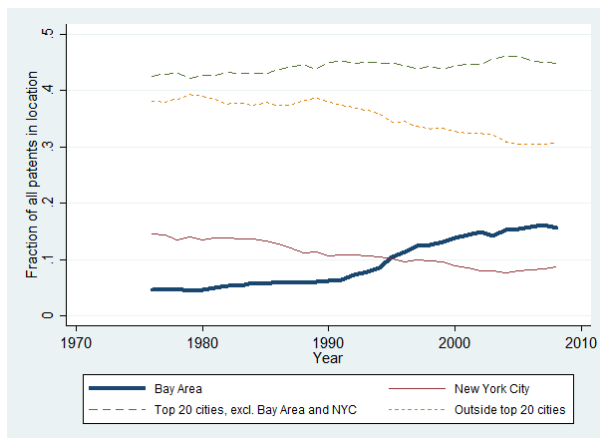


FIGURE 2. FRACTION OF PATENTS BY LOCATION TYPE

B. Patenting across types of patents

The Bay Area has had a cluster of ICT firms for many years. Therefore, one reason the Bay Area is becoming an increasing large fraction of patenting is the increase in ICT patents as a fraction of total patenting. Figure 3 displays this increase using the Hall, Jaffe, and Trajtenberg (HJT) definitions of patent classes. Computers and Communication (Class 2) went from under 10 percent of patents to over 30 percent of patents between 1980 and 2005. Some of this growth may reflect changes in the propensity to patent software and other ICT inventions (e.g., Graham and Mowery 2003, Hall and Ziedonis 2001) that have been encouraged by sympathetic treatment in the courts and the PTO. Drugs and Medical (Class 3) tracked the increase in Computers and Communication until the mid-1990s but then settled back to around 13 percent of patents.

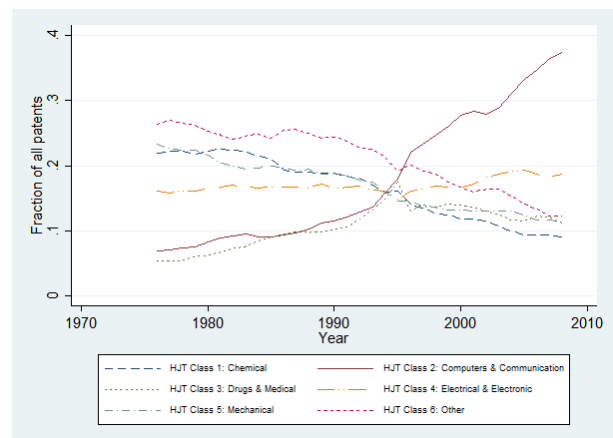


FIGURE 3. FRACTION OF PATENTS BY HJT CLASS

We offer the first evidence supporting the coagglomeration hypothesis with Figure 4,

which shows the fraction of patents that are in the Bay Area by broad class. The increase is sharpest in Computers and Communication and in Electrical and Electronic (Class 4). It is also noticeable in Chemicals (Class 1), Drugs and Medical, and Mechanical (Class 5). In Other (Class 6) the increase is smaller, rising from 3.8 percent in 1976 to a peak of 6.3 percent in 2004 before falling back to 4.4 percent in 2008. Thus, for five of six broad patent classes, we see a noticeable rise in the proportion of patents coming from the Bay Area.

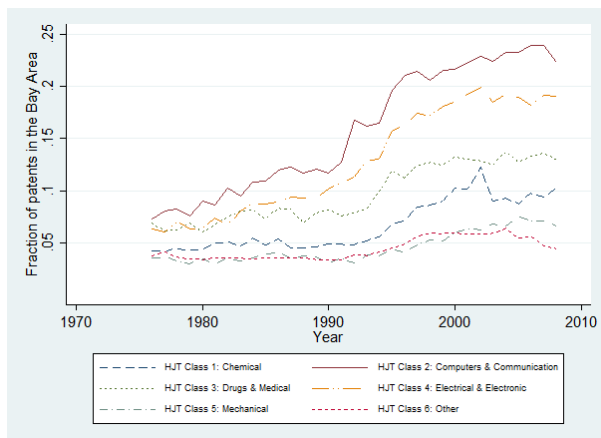


FIGURE 4. FRACTION PATENTS IN BAY AREA BY HJT CLASS

One possibility is that many of the patents in the Chemicals, Drugs and Medical, and Mechanical classes are ICT-based. Software has increasingly been used as an input into a wider array of inventions in other patent categories (Arora, Branstetter, and Drev 2013; Branstetter, Drev, and Kwon 2015). Figure 5 shows our preferred measures of ICT and non-ICT patents to account for this possibility, and examines the trend over time. We define ICT

patents as all patents in Computers and Communication (HJT Class 2) or Electrical and Electronic patents (HJT Class 4) plus software patents as defined by Graham and Mowery (2003). The general results are robust to dropping HJT Class 4 or software patents, or to including a broader definition of software patents that includes patents identified through a keyword search as in Bessen and Hunt (2007) and software patents identified in Graham and Vishnubhakat (2013).

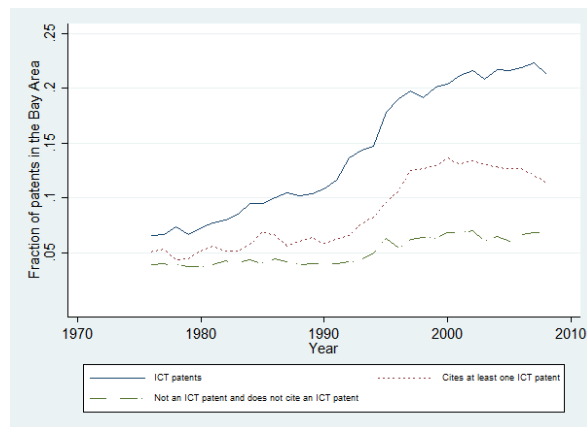


FIGURE 5. FRACTION ICT & NON-ICT PATENTS IN BAY AREA

The solid line at the top of the graph shows the increasing proportion of ICT patents that are in the Bay Area. The dotted line identifies all patents that cite ICT patents but are not explicitly categorized as ICT using the definition above. Instead, they are connected through citation and therefore build on ICT invention. There is a clear trend toward an increasing proportion of these patents in the Bay Area, providing another explanation for the rise of Bay Area patents.

Together, the above suggest the following: ICT is an increasingly large fraction of patents; the Bay Area is an increasingly large (and even dominant) fraction of ICT patents; and the Bay Area is an increasingly large fraction of patents that cite ICT patents. Given prior results on agglomeration of the ICT industry in the Bay Area, perhaps none of these results are surprising, though we believe that the results on geography of patents that cite ICT are not previously documented. These all could result from agglomeration of software invention near the location of the firms producing electronics, computing, and communications.

The evidence for coagglomeration of invention appears in the dashed line in Figure 5: The Bay Area is an increasing fraction of non-ICT US patents, rising from 3.9 percent to 6.9 percent from 1976 to 2008.

While these figures are more modest than the increase in ICT patents, they still suggest an increasingly important role for the Bay Area, relative to all other areas, in US non-ICT patenting. Figure 6 compares the Bay Area to the four other top patenting cities in the United States. The Bay Area was second behind New York for most of the period from 1997 to 2008.

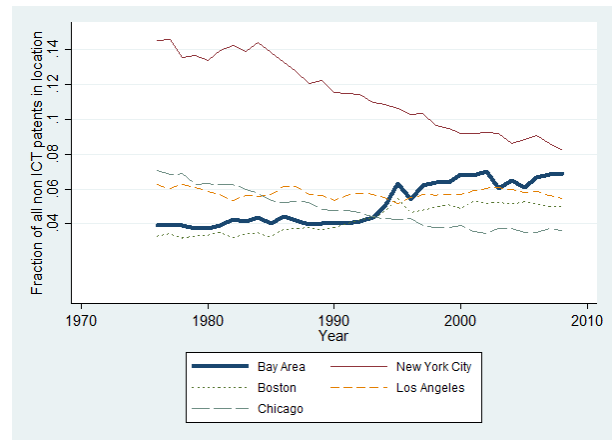


FIGURE 5. FRACTION NON-ICT PATENTS BY CITY

Overall, we interpret these results to suggest that we cannot reject coagglomeration of invention. The increase in patenting in the Bay Area is not entirely attributable to the increasing fraction of ICT patents in overall patenting.

III. Conclusions

We have documented an increase of the fraction of US patenting of all kinds that occurs the Bay Area that is disproportionate to population growth and occurs within a variety of patent classes. This partly results from the agglomeration of invention near the production of firms who use the invention, and who themselves agglomerate in one area. We also think it offers evidence of coagglomeration, the clustering of invention from many distinct types of invention into one geographic area.

While we do not know the cause of the rise in coagglomeration of many patent types in the Bay Area, our results suggest that any possible explanation must be broad-based. In particular, any explanation must account for growth in the fraction of ICT and non-ICT patents in the Bay Area.

REFERENCES

- Arora, Ashish, Lee G. Branstetter, and Matej Drev. 2013. Going Soft: How The Rise Of Software-Based Innovation Led To The Decline Of Japan's It Industry And The Resurgence Of Silicon Valley. *Review of Economics and Statistics*, 95(3): 757-775.
- Audretsch, David, Feldman, Maryann, 1996. R&D spillovers and the geography of innovation and production. *American Economic Review*. 86, 630-640.
- Bessen, James, and Robert Hunt. 2007. An Empirical Look at Software Patents. *Journal of Economics and Management Strategy*. 16(1), 157-189.
- Branstetter, Lee, Matej Drev, and Manho Kwon. 2015. Get with the Program: Software-Driven Innovation in Traditional Manufacturing. Working paper.
- Carlino, Gerald, and William Kerr. 2015. Agglomeration and Innovation. In Gilles Duranton, Vernon Henderson, and William Strange eds. *Handbook of Regional and Urban Economics Volume 5*. p. 349-404.
- Chen, Henry, Paul Gompers, Anna Kovner, and Josh Lerner. 2010. Buy local? The geography of venture capital. *Journal of Urban Economics* 67(1): 90-102.
- Delgado, Mercedes, Michael Porter, and Scott Stern, 2014. Clusters, Convergence, and Economic Performance. *Research Policy* 43(10), 1785-1799.
- Duranton, G., 2007. Urban evolutions: the fast, the slow, and the still. *American Economic Review*. 97, 197-221
- Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *American Economic Review*, 100(3): 1195-1213.
- Franco, April, and Matthew Mitchell. 2008. Covenants not to Compete: Labor Mobility and Industry Dynamics. *Journal of Economics and Management Strategy* 17(3), 581-606.
- Garcia-Vicente, Florencia, Daniel Garcia-Swartz, and Martin Campbell-Kelly. 2014. The Hydra of Lerna: Economic-Geography Perspectives on America's Early Computer Clusters. Working paper, Warwick University.
- Glaeser, Edward. 2005. Urban Colossus: Why is New York America's Largest City?

- FRBNY Economic Policy Review*. December. p. 7-24.
- Graham, Stuart, and David Mowery. 2003. Intellectual Property Protection in the U.S. Software Industry. In Wesley Cohen and Stephen Merrill eds, *Patents in the Knowledge Economy*, National Research Council, Washington DC. p. 219-258
- Graham, Stuart, and Saurabh Vishnubhakat. 2013. Of Smart Phone Wars and Software Patents. *Journal of Economic Perspectives* 27(1), 67-86.
- Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg. 2005. Market Value and Patent Citations. *RAND Journal of Economics* 36(1), 16-38.
- Hall, Bronwyn, and Megan MacGarvie. 2010. The private value of software patents. *Research Policy* 39(7), 994-1009.
- Hall, B. H., R. Ziedonis 2001. The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979–1995. *RAND Journal of Economics* 32: 101-128.
- Hannigan, Thomas J., Marcelo Cano-Kollmann, and Ram Mudambi. 2015. Thriving innovation amidst manufacturing decline: the Detroit auto cluster and the resilience of local knowledge production. *Industrial and Corporate Change* 24(6): 1-22.
- Helsley, Robert W. and William C. Strange. 2014. Coagglomeration, Clusters, and the Scale and Composition of Cities. *Journal of Political Economy*. 122(5), 1064-1093.
- Jaffe, A., M. Trajtenberg and R. Henderson 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics* 108(3): 577-598.
- Jaffe, Adam and Manuel Trajtenberg. 2002. *Patents, Citations, and Innovations: A Window on the Knowledge Economy*. MIT Press, Cambridge, MA.
- Kerr, William, and Scott Kominers., 2015. Agglomerative forces and cluster shapes. *Review of Economics and Statistics*, forthcoming
- Marx, M. J. Singh, and L. Fleming, “Regional Disadvantage? Employee Non-compete Agreements and Brain Drain.” *Research Policy* 44(2):394-404 (2015).
- Moretti, E., 2012. *The New Geography of Jobs*. Houghton Mifflin Harcourt, New York.
- Rosenberg, Nathan, 1963, Technology Change in the Machine Tool Industry, 1840-1910. *The Journal of Economic History*, 23(4), pp. 414-443.
- Saxenian, A., 1994. *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.