United Nations University

WIDER

World Institute for Development Economics Research

Discussion Paper No. 2001/2

Silicon Mountains, Silicon Molehills

Geographic Concentration and Convergence of Internet Industries in the US

Jed Kolko*

May 2001

Abstract

This paper investigates the effect of information technology on industrial patterns of concentration and convergence. Information-technology intensive industries exhibit slower employment convergence than other industries. The regression estimates suggest that the highest-IT industries exhibit employment convergence at only half the average rate for all industries. However, it is not the information-technology usage per se that is associated with slower convergence. Higher-IT industries hire more educated workers and are rapidly growing, and both of these characteristics have been argued in the literature to affect the rate of convergence. Controlling for these other characteristics reveals that the direct effect of information technology is to speed convergence. Thus, high-IT clusters persist not because they are technology-intensive per se, but because they tend to rely on high-skilled labour. Since convergence reduces the long-run efficacy of place-based economic policies, public policies that attract low-skill support functions for high-IT industries confer only short-term benefits in exchange for potentially large upfront costs.

Keywords: information technology, internet, economic geography, location, urban

JEL classification: R1, R4, O3, L2

* Forrester Research (Cambridge, MA)

This study has been prepared within the UNU/WIDER project on Production, Employment and Income Distribution in the Global Digital Economy, which is directed by Professor Matti Pohjola.

UNU/WIDER gratefully acknowledges the financial contribution to the project by the Ministry for Foreign Affairs of Finland.

Acknowledgements

The research in this paper was conducted while the author was a research associate at the Center for Economic Studies, US Bureau of the Census, and a doctoral candidate at Harvard University. The author thanks the US Small Business Administration, Office of Advocacy, for granting access to the data at the Center for Economic Studies. This work was supported by a Doctoral Dissertation Improvement Grant from the National Science Foundation. Research results and conclusions expressed are those of the author and do not necessarily indicate concurrence by the Bureau of the Census, the Center for Economic Studies, the Small Business Administration, the National Science Foundation, or Forrester Research.

1 Introduction

This paper investigates the effect of information technology on industrial patterns of concentration and convergence. Concentration is the tendency of an industry to cluster geographically; convergence is the tendency of an industry to become more uniformly distributed geographically—that is, less concentrated—over time.[1] Economic research has documented evidence of both concentration and convergence: as a rule, industries are geographically concentrated, but these concentrations weaken over time.[2]

The tendency of industries to converge geographically over time has two important policy implications. First, convergence means that the effects of local shocks diminish over time. Whether a positive (or negative) local shock comes from market forces or government intervention designed to boost a local industry, the lead that the locality takes in that industry dissipates over time as firms move away and new firms start up elsewhere. Second, convergence means that local macroeconomic fluctuations should dampen over time. To the extent that convergence diversifies a city's industry mix, and industries' shocks are less than perfectly correlated, convergence might bring greater stability to local wages and unemployment rates. Thus, convergence both reduces the long-run efficacy of place-based economic policies while reducing the local disparities that make place-based economic policies appealing in the first place.

The basic fact to be explored is that information-technology intensive industries exhibit slower employment convergence than other industries. The regression estimates suggest that the highest-IT industries exhibit employment convergence at only half the average rate for all industries. However, it is not the information-technology usage per se that is associated with slower convergence. Higher-IT industries hire more educated workers and are rapidly growing, and both of these characteristics have been argued in the literature to affect the rate of convergence. Controlling for these other characteristics reveals that the direct effect of information technology is to speed convergence. Looking more closely at those counties where technology employment is highest, one sees the persistent, though weakening, dominance of centres like Silicon Valley and Boston's northern and western suburbs.

2 Theory

Concentration at a point in time and convergence over time are distinct, but related, concepts. Concentration is the extent to which an industry's employment (or

---

1 This definition of 'convergence' comes from the economic growth literature, which uses 'convergence' to mean the tendency of poorer regions to enjoy faster income growth than richer regions. Crudely, it means that places become more alike over time. Barro and Sala-i-Martin (1998) discuss convergence extensively in the economic growth context.

2 Marshall (1892), Krugman (1991), Glaeser *et al.* (1992), and Kim (1995). Convergence need not always result in declining industrial concentration if an industry's employment shifts from a concentration in one set of cities to an equally strong concentration in a different set of cities. This issue will be taken up later in the paper.

production) is clustered in a small number of places; convergence is the extent to which an industry's employment shifts away from locations in which it is over-represented. Changes in industrial concentration over time can be decomposed into convergence and random shocks. Industry concentrations can arise either with large random shocks and sufficiently slow convergence, or with greater-than-epsilon random shocks and non-convergence or divergence.[3]

The literature on industrial concentration has taken the presence of random shocks as given and has focused on explaining the determinants of the rate of convergence (or divergence). What determines convergence (or divergence) is the relative strength of centrifugal and centripetal forces, to use the language of Krugman (1993); centrifugal forces encourage firms to locate away from other firms in their industry, contributing to convergence, while centripetal forces encourage firms to locate near other firms in their industry, contributing to divergence.

Beginning with Marshall, the literature has explored in detail possible centripetal forces. Marshall outlined several reasons why firms in the same industry locate near each other: (1) *knowledge spillovers*: to benefit from an exchange of ideas of innovations; (2) *labour pooling*: to share access to a pool of skilled labour that is willing to accept lower wages in exchange for a lower likelihood of unemployment, which arises if firms have imperfectly correlated labour demand shocks; and (3) *intermediate inputs*: to share access to intermediate inputs, the production of which might involve economies of scale. These forces affect firms' location decisions because these benefits depend on geographic proximity. Knowledge spillovers often involve planned or spontaneous face-to-face interactions, and the accuracy of the knowledge might decline if transmitted over longer distances electronically, or on paper.[4] Labour pooling works only among firms located within commuting distance of the skilled labour pool. And the benefits from sharing intermediate inputs are conferred only on firms to which the shared inputs can be cheaply transported.

Centrifugal forces prevent industries from being entirely concentrated in a single location. The cost of transporting the industry's output to geographically dispersed consumers can encourage firms to disperse, either to reduce transport costs that producers incur directly or to become dominant in an underserved market.[5]

---

3 The relationship between convergence and concentration is expressed easily mathematically. Let shr_{ix} be the share of industry i's employment in county x. Define concentration as Σshr_{ix}^2. The change in concentration over time is $\Sigma(shr_{ix}+\Delta shr_{iz})^2 - \Sigma shr_{ix}^2 = 2\Sigma shr_{ix}\Delta shr_{ix} + \Sigma(\Delta shr_{ix})^2$. The first term reflects convergence, which is the tendency of industries to grow faster where they are under-represented; the second term reflects the magnitude of random shocks. A fuller explanation appears in Dumais, Ellison, and Glaeser (1997).

4 Jaffe, Trajtenberg and Henderson (1993) show, using patent data, that geographic proximity facilitates the flow of ideas. Whyte (1988) describes the richness of in-person communication and the importance of voice, body language, and other visual cues.

5 The cost of transporting output to consumers can be a centrifugal force even when firms do not pay the direct cost for transporting their output. Krugman (1991) presents a model in which monopolistically competitive firms do not pay transport costs directly but still benefit from locating in regions where there is less competition from other producers.

Alternatively, the concentration of an industry can bid up prices for land, geographically immobile labour, or other scarce resources, causing firms to leave the concentrated area and causing new firms to start up elsewhere.

Recent advances in information technology have the potential to change the balance of centripetal and centrifugal forces and, therefore, the potential to change the rate at which an industry converges geographically. Advances in information technology have facilitated data manipulation and storage, allowing firms to create and maintain large, centralized databases and to improve record-keeping. These advances have also introduced new methods of data transmission on private networks within a firm, on private networks between firms, and on the public network—the internet. The data-manipulation-and-storage technology compresses a given amount of information into less and less electronic space, and the data-transmission technology allows a given amount of electronic data to be sent more quickly and cheaply.

In the popular mind are two different notions about the effect of information technology on industrial convergence. On the one hand is the image of the 'technopolis', of which Silicon Valley is the most obvious and best documented example.[6] These concentrations of high-technology firms create a critical mass of skilled labour, ready capital, and innovative ideas that keep existing firms there, encourage new start-ups, and attract firms from elsewhere.[7] Silicon Valley has inspired cities around the world to imitate its success, from at the very least of renaming an area Silicon-something, to major investment projects like Malaysia's Cyberjaya that offer infrastructure, tax breaks, and regulatory relief to high-tech firms. These technopoles suggest that information-technology industries thrive on concentration and that convergence is slow if at all.

The other image is the firm on the farm (or the yacht, or the mountaintop), using information technology to communicate with clients, suppliers, and competitors located elsewhere. Larger firms, like Citibank and American Express, have moved back-office functions away from major big-city banking centres to smaller towns. Numerous small firms and self-employed workers have moved or started up in Florida, the Southwest, and in other resort areas, away from the traditional centres of their industries. For these 'forty acres and a modem'[8] firms, new information technology makes it possible to locate far from existing concentrations of capital, ideas, and even labour, in order to take advantage of cheaper land or a preferred lifestyle. According to this argument, then, information technology speeds industrial convergence.

To sort out these contradictory popular images, consider how the reduction in the cost of sending data electronically affects the components of convergence outlined above.

---

6 Saxenian (1994).

7 Even the best-established technology firms outside of Silicon Valley have increased their presence there. Microsoft, based in Seattle, is building a Silicon Valley campus; America Online, based in Virginia, purchased Silicon Valley-based Netscape.

8 'Why Wall Street is Losing Out to 40 Acres and a Modem', *New York Times*, 27 December, 1998. Section 3, page 7, column 1.

First, knowledge spillovers become less dependent on geographic proximity. Ideas can be shared, written drafts exchanged, pictures and diagrams compared, and meetings conducted, all electronically: the quality of these interactions is better preserved over long distances than in the past. The regular mail system and the telephone of course also facilitate the flow of information over longer distances, but new information technologies, particularly the internet, are qualitatively different than mail and phone. The internet combines the instantaneous communication of the telephone with the regular mail's capacity to transmit documents, drawings, and images. Furthermore, the internet allows communication among many people simultaneously: it is no more difficult or expensive to send an e-mail to ten people than to one; chat rooms and newsgroups permit ongoing rapid dialogue among many participants; and the Web allows documents to be published publicly for the world to see. Finally, the internet allows one to wander and browse, opening the way for spontaneous and serendipitous discoveries; this mimics the unplanned and unexpected opportunities for sharing ideas that inevitably arise at dinner parties and street corners in areas where an industry concentrates.

New information technology, therefore, makes knowledge spillovers less dependent on geographic proximity; since knowledge spillovers are a centripetal force, this effect of information technology is to speed convergence. This supports the '40 acres and a modem' view.

The second effect of information technology is that it lowers the transport cost of many intangible goods. Industries like accounting, advertising, management consulting, and other services whose main output is intangible information are increasingly able to substitute in-person delivery to customers with electronic delivery. One can bank, take courses, and buy music entirely on-line. This reduction in the cost of transporting output weakens one of the main centrifugal forces, and therefore implies slower convergence for industries whose outputs are intangible—though should have no effect for manufacturing industries, whose outputs are tangible goods.[9] This raises the potential for high-tech service industries to concentrate—supportive of the 'technopolis' view. The overall predicted effect of information technology, therefore, is to speed convergence for manufacturing industries, while the effect on service industries is ambiguous.

Why, then, the persistence of the 'technopolis' view and the basic fact that technology-intensive industries exhibit slower convergence? This apparent contradiction arises because other characteristics of information-technology-intensive industries have to this point been overlooked. High-IT industries tend to employ highly skilled labour and are fast-growing, and both of these characteristics also affect the rate of geographic convergence. The skill level of an industry should affect convergence in two ways. First, industries with highly skilled workers by definition rely more on knowledge and ideas than other industries, so the centripetal pull caused by knowledge spillovers should be stronger in highly skilled industries. Second, skills are a scarce resource, and

---

9 Kolko (1999) offers a model that compares the role of transport costs in service and manufacturing industries, and demonstrates formally the effect of declining transport costs on industry concentration.

the benefits of labour pooling are likeliest to accrue to firms hiring labour with skills specific to the industry. The centripetal pull of labour pooling should also be stronger—and convergence slower—in highly skilled industries.[10]

Skills affect the geography of high-IT industries through other channels as well. High-skill *places* attract high-skill *industries*: since high-IT industries rely little on natural physical resources, their location decisions are driven more by the availability of scarce skilled labour. Furthermore, the scientific innovations that facilitate the growth of skilled industries often germinate first in university research centres, strengthening the advantage that high-skill places have for high-skill industries. Kolko (2000) shows that both local skill level and local university presence have independent positive effects on commercial Internet development in the US.

The rate of growth of an industry also affects convergence. Dumais, Ellison, and Glaeser (1997) separate the relative contributions to industry concentration of establishment births, the expansion/contraction of existing establishments, and establishment deaths. They find that births and expansions are more rapid outside of industrial concentrations, but deaths are as well, all contributing to the conclusion that growing industries converge faster.

The empirical section of this paper will isolate the effect of information technology on convergence and concentration by controlling for industry skill level and industry growth.

3 Data and methodology

The data used in this paper come from two sources. Data on employment location come from the Longitudinal Enterprise and Establishment Microdata (LEEM) file. Data on information technology usage and skill level of employees come from the Current Population Survey (CPS). These two data sources are matched by industry.

The LEEM tracks the employment of every private-sector establishment in the US with at least one employee. The unit of observation in the LEEM is the establishment, which is a site or plant where business is conducted. A firm consists of one or more establishments. The LEEM excludes the self-employed, farms, railroads, and private household employment. In 1996, the LEEM covers 100 million employees in almost seven million establishments. For each establishment, the LEEM reports employment, location by county, industry (using the Standard Industrial Classification), payroll, and a code that identifies the firm to which the establishment belongs. The LEEM covers the years 1989-96.

---

[10] Dumais, Ellison, and Glaeser (1997) find that industries with skilled workers are more concentrated and suggest that this effect works both through knowledge spillovers and labour pooling. Glendon shows that concentrations of skilled workers are stable over time and show less convergence that concentrations of less-skilled workers.

The LEEM is based on the Census Bureau's business register, the Standard Statistical Establishment List (SSEL).[11] The Census compiles the SSEL annually, based on Internal Revenue Service records, Social Security Administration data, and the Bureau of Labor Statistics's unemployment insurance records; the Census supplements these sources with its own annual Company Organization Survey and the quinquennial Economic Census. The Census generates the LEEM by linking establishments from the SSEL across years, resulting in a file that tracks individual establishments over time, even if the establishment changes ownership or location.[12] Because the LEEM comprises establishment-level microdata, it contains confidential information about individual firms. Thus, the LEEM can be used only at the Census Bureau's offices by researchers sworn not to disclose information about any one firm.[13]

The empirical strategy involves local industry growth regressions. Convergence is defined above as the tendency for an industry's employment to grow slower in areas where the industry is over-represented: the dependent variable in the regression analysis is net employment growth for a local industry. 'Industry' is defined using the Census's 3-digit classification into 204 categories. 'Local' is defined as metropolitan statistical areas (MSA's).[14]

The explanatory variable of interest is localization. Localization is defined as the fraction of an industry's national employment in a given MSA and can range from zero to one. A negative coefficient means that industries grow more slowly in MSA's where they are over-represented—thus, centrifugal forces dominate and industries exhibit geographic convergence.

Other variables are used as controls in the local growth regressions. Following Glaeser *et al.* (1992), who found significant effects of local industry competition and the diversity of other industries locally on local industry growth, measures of competition and diversity are used as controls. MSA employment outside the industry is included as a control for city size. These are defined precisely in the data appendix.

Local employment growth, geographic concentration, local industry competition, local diversity, and MSA employment are all measured using the establishment-level employment data in the LEEM.

---

11 County Business Patterns, an annual tabulation of employment by industry and county, is also based on the SSEL.

12 The Census Bureau's Longitudinal Research Database (LRD) also tracks establishments over time, but differs from the LEEM. The LRD is based primarily on the Economic Census and has more variables than the LEEM, but it covers only the manufacturing sector whereas the LEEM includes all sectors.

13 See Robb (1999) for full documentation of the LEEM.

14 In New England, 'local' means a New England County Metropolitan Area (NECMA), since the LEEM identifies geography by county. Consolidated Metropolitan Statistical Areas (CMSA's) are used where applicable. Counties outside of MSA's are grouped by state, and the resulting areas are included in the analysis as MSA's. There are 271 MSA's/CMSA's/NECMA's and 50 additional 'constructed' areas of counties outside of MSA's.

To assess the effect of information technology on convergence, industry-level characteristics are interacted with the localization measure. These industry-level characteristics include information-technology intensity and skill level of the industry, and they are taken from the Current Population Survey (CPS). The CPS is a monthly survey of American households, conducted by the Census Bureau and the Bureau of Labor Statistics. The CPS regularly asks respondents about their education level and industry where they work. Every four years in its October survey, the CPS also asks respondents about their computer usage at home and at work. Respondents are asked whether they use a computer and, if so, for which tasks; possible answers include 'electronic mail' and 'communications'. An industry's information technology usage is defined as the fraction of employees who report using a computer at work for either electronic mail, communications, or both, in 1989. An industry's skill level is defined as the fraction of employees who have a college degree.

The third industry characteristic interacted with localization is industry growth at the national level. National industry growth is defined as industry growth outside the MSA, in order to avoid having the dependent variable (local industry growth) included in an independent variable.

The empirical specification is:

$$\frac{\underset{'89-'96}{\Delta emp_{ix}}}{\left(\underset{'89}{emp_{ix}} + \underset{'96}{emp_{ix}}\right)\Big/2} = \alpha + \lambda_1 competition_{ix} + \lambda_2 diversity_{ix} + \lambda_3 MSAemployment_{ix} + \beta_1 localization_{ix}$$
$$+ \beta_2 (localization_{ix} * tech_i) + \beta_3 (localization_{ix} * educ_i) + \beta_4 (localization_{ix} * industrygrowth_{ix}) + \varepsilon_{ix}$$

where:

— i indexes industries;

— x indexes MSA's;

— the λ coefficients are refer to control variables;

— the β coefficients refer to the convergence variables and interactions.

The expression for local employment growth uses the average of start-year and end-year employment as the denominator. This follows Davis, Haltiwanger, and Schuh (1996) and has the advantage of having a defined value for industry-MSA cells with no employment in the start-year and positive employment in the end-year. All regressions are weighted by the local industry size (the denominator of the dependent variable), and industry fixed effects are included.

4 Results

The main results are presented in Table 1. In column one is the evidence that industries exhibit geographic convergence: a univariate regression of net employment growth on industry localization yields a negative and significant coefficient. The coefficient implies that a one standard deviation increase in the localization of an industry in an MSA leads to a 0.22 standard deviation decline in net employment growth.

The second column includes the interaction between localization and information technology usage. As mentioned in the introduction, the estimated coefficient on the interaction is positive and significant. The magnitudes of the coefficients suggest that industries with information technology usage above 83% cease to exhibit convergence and instead start to diverge. Since the highest information technology in 1989 was below this level—it was 56% in the computer manufacturing industry—the estimates suggest that IT-intensive industries still exhibit convergence, but at a much slower rate than other industries. An industry with average IT usage (12%) is estimated to show a rate of convergence that is 2.6 times that of an industry at 56% IT usage.

The main theoretical prediction about the effect of information technology is that after including other controls, IT usage should speed convergence. Column three includes the competition, diversity, and MSA size controls, and—more importantly—the interactions between localization, skill level, and industry growth rates nationally. As predicted, the effect of information technology interaction is negative and significant. The coefficient on industry skill level is positive and significant, suggesting that the apparent slow convergence among high IT industries is due to their high skill level rather than their usage of IT per se. The coefficient on the national industry growth interaction is also negative—consistent with Dumais, Ellison, and Glaeser (1997). This has little effect on the IT interaction, though, since the correlation between IT usage and national growth is weak compared to the correlation between IT usage and skill level.

The second prediction about information technology is that it should slow convergence for service industries relative to manufacturing industries, as some service outputs can be delivered electronically and need no longer be locally provided to a diverse customer base. The final column of Table 1 includes an interaction between localization, IT usage, and service sector;[15] this interaction term is negative, inconsistent with the prediction that IT might slow convergence for some service industries.

The regressions are repeated in Table 2, using growth due to new-firm births, rather than net employment growth, as the dependent variable. New-firm births mean growth due to the formation of new companies only—these new companies must have come into existence during the period 1989-96, and they must not include any establishments that were in existence prior to 1989. Focusing on new-firm births restricts the sample to location decisions unconstrained by existing firm capital or past location decision. These firms are, in theory, completely 'footloose' when they come into existence, and their optimal location is chosen during the period under study.

---

[15] The service-sector dummy equals one for industries in SIC categories 52-59 (retail trade), 6 (finance, insurance, real estate) and 7 and 8 (services).

Table 1
Convergence and IT: net employment growth

	Dependent variable: net employment growth			
	(1)	(2)	(3)	(4)
Competition = one minus local firm-level Herfindahl			0.097	0.108
			0.008	0.008
Diversity = similarity of local industry mix to national (excl. own industry)			-0.115	-0.118
			0.007	0.007
MSA employment (excl. own industry), millions			-0.029	-0.032
			0.001	0.001
Localization = industry-MSA employment as fraction of national industry	-1.647	-2.047	-0.743	-0.690
	0.028	0.047	0.072	0.079
Localization * fraction of employees using communications technology		2.451	-2.347	-1.994
		0.228	0.312	0.373
Localization * fraction of workers with college degrees			3.284	1.963
			0.252	0.275
Localization * industry growth rate (excl. own MSA), 1989-96			-1.361	-2.065
			0.128	0.142
Localization * service-sector dummy				1.109
				0.104
Localization * fraction of employees using communications technology * service-sector dummy				-0.506
				0.485
Constant	0.158	0.159	0.059	0.048
	0.001	0.001	0.008	0.008
Observations	57187	57187	57187	57187
R-squared	0.37	0.37	0.40	0.40

Notes: Unit of observation is industry-MSA cell; includes industry fixed-effects; weighted by industry-MSA size; standard errors are below coefficient estimates; for definitions of the variables, see Data Appendix.

The results are similar to those for overall net employment growth, with some important differences. First, overall convergence is much slower. One standard deviation increase in the localization of an industry reduces new-firm births by 0.05 standard deviations—less than one-fourth of the reduction in net employment growth. This is consistent with the notion that start-ups benefit greatly from proximity to existing firms, for proximity allows them to tap into existing labour pools, share existing suppliers, and even spin off from existing firms.

Second, convergence due to new-firm births is faster, instead of slower, for high-technology industries, as the coefficient on the localization-technology interaction is now negative instead of positive. But the inclusion of the skilled workers and industry growth interactions has the same effect as in the first set of regressions: the coefficient on the technology interaction falls. The coefficient on the skilled workers interaction is again positive (though insignificant), and the coefficient on the industry growth interaction is again negative. Thus, the effect of information technology on convergence, after controlling for other factors, is still negative.

Table 2
Convergence and IT: new firm births

	Dependent variable: growth in new firm births			
	(1)	(2)	(3)	(4)
Competition = one minus local firm-level Herfindahl			0.060	0.061
			0.004	0.004
Diversity = similarity of local industry mix to national (excl. own industry)			-0.068	-0.069
			0.003	0.003
MSA employment (excl. own industry), millions			0.002	0.001
			0.000	0.000
Localization = industry-MSA employment as fraction of national industry	-0.176	-0.124	-0.075	-0.042
	0.012	0.020	0.031	0.034
Localization * fraction of employees using communications technology		-0.321	-0.433	-0.585
		0.096	0.133	0.159
Localization * fraction of workers with college degrees			0.062	-0.094
			0.107	0.117
Localization * industry growth rate (excl. own MSA), 1989-96			-0.689	-0.780
			0.054	0.061
Localization * service-sector dummy				0.038
				0.044
Localization * fraction of employees using communications technology * service-sector dummy				0.391
				0.207
Constant	0.175	0.175	0.103	0.102
	0.001	0.001	0.003	0.003
Observations	57187	57187	57187	57187
R-squared	0.52	0.52	0.53	0.53

Notes: Unit of observation is industry-MSA cell; includes industry fixed-effects; weighted by industry-MSA size; standard errors are below coefficient estimates; for definitions of the variables, see Data Appendix.

The final difference in the new-firm birth regression is that the technology-and-services interaction coefficient is now positive and significant at just above the 5% level. This variable indicates the extent to which technology slows convergence for service industries relative to manufacturing industries. Its positive value is consistent with the prediction that information technology permits the electronic delivery of intangible goods, thus freeing service industries from having to locate near their customers. This weakens a key centrifugal force for services and therefore slows convergence and encourages concentration.

As outlined above, while convergence is a component of declining concentration, the two are distinct and have slightly different policy implications. Changes in concentration over time depend on both the rate of convergence and the size of random shocks. An analysis of the causes of changes in concentration over time appears in Table 3. The dependent variable is the change in an industry's concentration, as measured by a locational Herfindahl index (see appendix); the unit of observation is the industry. The explanatory variables are the same three variables that were interacted with localization in Tables 1 and 2: technology usage, skill level, and industry growth.

Table 3
Industry concentration and information technology

	Dependent variable:		
	Change in concentration	Change in HQ concentration	Change in distance from headquarters
Fraction of employees using	-0.0178	-0.0298	-86.2
communications technology	0.0066	0.0191	120.5
Fraction of workers with college degrees	0.0064	0.0066	63.9
	0.0045	0.0129	81.6
Industry growth rate (excl. own MSA),	-0.0087	0.0005	194.3
1989-96	0.0023	0.0066	41.7
Constant	-0.0019	-0.0008	-19.0
	0.0009	0.0027	16.9
R-squared	0.09	0.01	0.11
Observations	204	204	204

Notes: Unit of observation is the Census 3-digit industry; standard errors are below coefficient estimates; for definitions of the variables, see Data Appendix.

The effects of these industry-level variables on concentration are consistent with the convergence regressions.[16] The coefficients on technology usage and industry growth are negative, and the coefficient on skill level is positive (though statistically insignificant). The effect of information technology is to reduce industrial concentration.

Table 4
Summary statistics

	Mean	S.d.
For convergence regressions		
Net employment growth (dependent variable)	0.114	0.311
Growth in new firm births (dependent variable)	0.171	0.149
Competition = one minus local firm-level Herfindahl	0.894	0.181
Diversity = similarity of local industry mix to national (excl. own industry)	-0.241	0.171
MSA employment (excl. own industry), millions	1.715	2.169
Localization = industry-MSA employment as fraction of national industry	0.026	0.041
Localization * fraction of employees using communications technology	0.004	0.010
Localization * fraction of workers with college degrees	0.007	0.015
Localization * industry growth rate (excl. own MSA), 1989-96	0.003	0.013
Localization * service-sector dummy	0.014	0.030
Localization * fraction of employees using communications technology * service-sector dummy	0.002	0.007
For industry-level concentration regressions		
Concentration index, 1989	0.036	0.036
Concentration index, 1996	0.036	0.036
Change in concentration, 1989-96	-0.003	0.008
HQ concentration index, 1989	0.046	0.054
HQ concentration index, 1996	0.043	0.055
Change in HQ concentration, 1989-96	-0.003	0.021
Average distance from HQ, 1989	417.2	202.2
Average distance from HQ, 1996	410.5	191.0
Change in average distance from HQ, 1989-96	-6.7	142.6
Fraction of employees using communications technology	0.116	0.102
Fraction of workers with college degrees	0.223	0.149
Industry growth rate (excl. own MSA), 1989-96	0.117	0.184

16 A univariate regression of change in concentration on IT-intensity yields a negative, though statistically insignificant, coefficient. To reconcile this with the slower rate of convergence for high-IT industries, as shown in column 2 of Table 1, it must be that IT-intensity is associated with a smaller random-shock component.

It was hypothesized that the negative effect of technology on industrial concentration is due to the decreasing reliance of knowledge spillovers on geographic proximity. With new information technology, the argument goes, these spillovers can take place electronically, and therefore at greater distances, as well as in-person. An alternative phenomenon could generate the same empirical results, however: the decline in concentration could be due to reorganizations within firms that separate back-office functions from headquarters functions. Under this alternative hypothesis, knowledge spillovers between firms might still depend on geographic proximity, so headquarters remain concentrated; only back-offices relocate to low-cost areas while remaining connected electronically to headquarters.

To test this alternative of within-firm reorganization, the regression in column 1 of Table 3 is repeated, but only with establishments that are located in the county of their firm's headquarters.[17] The results, presented in column 2 of Table 3, are suggestive but inconclusive. The coefficient on technology usage continues to be negative, indicating that technology lowers the concentration level among headquarters establishments, but insignificant at the 5% level. The other coefficients are insignificant as well.

Another approach to testing between-firm vs. within-firm effects is to identify whether the explanatory variables have any effect on the internal geography of firms. For each industry, the average distance of an establishment from its headquarters is calculated. Establishments in headquarters counties are given a distance of zero, and the average is weighted by the number of employees in each establishment. For instance, if an industry consists entirely of single-establishment firms, then the average distance from headquarters is zero. Or, if an industry consists of a single firm, headquartered in New York, with half its employees in New York and half in San Francisco (3000 miles from New York), then the average distance from headquarters is 1500 miles.

The results in column 3 of Table 3 present industry-level regressions, with the industry's *change* in average distance from headquarters as the dependent variable. If the main effect of technology is to allow back-offices to be separated from headquarters, then the level of technology usage should be positively correlated with the average distance from headquarters. The results show that this is not the case. The coefficient on technology is negative and insignificant. The only significant effect on average distance from headquarters is the growth rate of the industry: faster growing industries tend to exhibit an increasing average distance from headquarters.

---

[17] The micro-data indicates the firm to which each establishment belongs. A separate file, the Standard Statistical Establishment List, gives the headquarters address of every firm. All establishments in the headquarters county are included because the data do not definitively identify the actual headquarters establishment. 'Establishments in the headquarters county' are identified by matching each establishment to the official mailing address of the parent firm. Establishments that constitute single-establishment firms are included in the category of 'establishments in headquarters county'.

5 High-tech agglomerations in the internet evolution

This final section explores more descriptively the concentration of high-tech employment in Silicon Valley and other technology centres over the period 1988-96. Whereas the Census restricts publication of small-cell data from the LEEM and the SSEL, the Census creates an annual tabulation of employment by industry and county in County Business Patterns (CBP).[18] Data from CBP demonstrate Silicon Valley's lead in the internet economy yet reveal a diversity of high-tech clusters emerging across the US.

Santa Clara County, California, is well-known as ground zero of the internet economy and has been the technology centre of the US for years. Table 5 shows that employment in computer and information technology industries[19] was greatest in Santa Clara County in each of the years 1988, 1992, and 1996—in spite of an initial contraction in technology employment and the development of the internet. There was little movement in the list generally: nine of the top ten counties for technology employment in 1988 were on the top ten list in 1996. The only notable movements were Fairfax County VA going from 5th to 2nd and New York County (Manhattan) going from 10th to 7th.

Table 5
Counties with highest employment in technology industries, 1988-96

1988		1992		1996	
Santa Clara, CA	72295	Santa Clara, CA	62783	Santa Clara, CA	66647
Middlesex, MA	51306	Middlesex, MA	39264	Fairfax, VA	54504
Los Angeles, CA	38341	Los Angeles, CA	32740	Middlesex, MA	51149
Orange, CA	26841	Fairfax, VA	31168	Los Angeles, CA	46563
Fairfax, VA	26500	Orange, CA	26335	Cook, IL	37971
Dallas, TX	24781	Cook, IL	23883	Dallas, TX	34915
Cook, IL	23402	Dallas, TX	22793	New York, NY	29504
Hennepin, MN	20524	San Diego, CA	18723	Orange, CA	27461
San Diego, CA	17721	Travis, TX	18214	Hennepin, MN	24057
New York, NY	17510	New York, NY	17885	Harris, TX	23956

Note: Technology employment includes SIC codes 357 and 737.

[18] CBP covers all private-sector non-farm employment in establishments with at least one paid employee. The total employment covered by CBP was around 100 million employees in 1996. The Census suppresses data in small industry-county cells where exact figures would reveal information about a particular establishment. When data are suppressed, the Census reports employment as a range, rather than as an exact figure. The imputation procedure used to make point estimates for suppressed cells is described in Kolko (1999).

[19] Computer and information technology industries are defined as those in SIC 357 and 737. SIC 357 is 'computer and office equipment' manufacturing, and SIC 737 is 'computer programming, data processing, and related services'. Over the period 1988-96 there were no redefinitions of the SIC classification.

What distinguishes Santa Clara County is not only its consistent lead in technology employment, but also the concentration of both computer manufacturing ('hardware')[20] and computer and data services ('software').[21] In 1996, Santa Clara led in hardware and was 4th in software; only Orange County CA, just south of Los Angeles, made both lists. Though software employment exceeds hardware employment in the US, it is in hardware that Silicon Valley dominates.

Table 6
Counties with highest employment in hardware and software, 1996

Hardware		Software	
Santa Clara, CA	27829	Fairfax, VA	54098
Durham, NC	15630	Middlesex, MA	45894
Boulder, CO	7605	Los Angeles, CA	43902
Madison, AL	6916	Santa Clara, CA	38818
Benton, OR	6890	Dallas, TX	34292
Fayette, KY	6877	Cook, IL	33319
Orange, CA	6559	New York, NY	29466
Hennepin, MN	5774	Oakland, MI	22989
San Diego, CA	5490	Orange, CA	20902
Harris, TX	5454	Harris, TX	18502

Notes: Hardware employment includes SIC code 357; software employment includes SIC code 737.

Separating hardware and software illustrates a clear geographic specialization: technology manufacturing is concentrated in smaller cities (like Durham County NC, in the Research Triangle; Benton County OR, in the middle of the state; and Madison County AL, around small-town Huntsville) while software and services cluster in large central cities and nearby suburbs. This geographic separation of technology manufacturing and service industries mirrors the distribution of manufacturing and services generally, with service industries over-represented in larger cities.[22]

The popular emphasis on Silicon Valley—and the attempts of other regions to create their own Silicon Whatevers—masks the diversity of roles that technology centres play. While there are other technology capitals in the Silicon Valley mould, much technology employment is centred in places that function very differently than Silicon Valley. Table 7 presents the counties with the largest technology employment, along with indicators of the local importance of technology, the diversity of technology employment, the size of the average technology firm, and recent growth of technology employment. Together, these indicators suggest a typology of four different kinds of technology-producing locations.

---

20 SIC 357 only.

21 SIC 737 only.

22 See Kolko (1999).

First are the technology capitals. In the US, there are three: Silicon Valley, Fairfax County VA (between Washington DC and Dulles Airport), and Middlesex County MA (which includes Cambridge and Route 128). In these three regions, technology is a significant share of local employment, ranging from 7-14%, though—consistent with the earlier results on convergence—only Fairfax County has seen technology employment grow in the period 1988-96. These technology capitals are lands of highway-connected office parks, each on the outskirts of major cities with educated populations, eclipsing the economic prominence that their neighbouring city traditionally had.

Table 7
The 25 largest technology centres, 1996

	Share of local employment in technology, %	Sub-sectoral technology concentration	Average technology establishment size, employees	Growth of technology employment, % (1988-96)
Santa Clara, CA	7.89	0.41	29.7	-8
Fairfax, VA	13.58	0.33	34.1	106
Middlesex, MA	6.86	0.22	28.6	0
Los Angeles, CA	1.34	0.16	16.9	21
Cook, IL	1.61	0.23	15.9	62
Dallas, TX	2.71	0.17	23.2	41
New York, NY	1.59	0.30	15.3	68
Orange, CA	2.35	0.16	17.4	2
Hennepin, MN	3.06	0.20	18.7	17
Harris, TX	1.66	0.18	17.3	44
Oakland, MI	3.31	0.34	27.2	140
San Diego, CA	2.45	0.22	19.5	21
Alameda, CA	3.41	0.26	18.7	72
Montgomery, MD	5.35	0.21	18.4	23
King, WA	2.06	0.27	14.0	79
Fulton, GA	2.83	0.17	20.6	136
Durham, NC	11.94	0.84	143.6	126
San Mateo, CA	5.00	0.35	20.7	133
DuPage, IL	2.57	0.25	12.4	32
Middlesex, NJ	3.81	0.47	15.2	79
Prince Georges, MD	5.74	0.43	40.7	50
Boulder, CO	10.22	0.53	28.9	20
Collin, TX	9.89	0.54	34.1	184
Travis, TX	3.51	0.29	19.6	97
Madison, AL	10.22	0.55	63.9	79

Note: Sub-sectoral concentration is the deviation of the local mix of technology employment across 15 4-digit SIC technology industries from the national average, measured as the sum of absolute differences divided by two; 0 means the local mix is the same as the national mix, and 1 indicates maximum divergence.

The second category is the urban tech centre, typified by Los Angeles, Cook County IL (Chicago), and New York County (which includes only Manhattan). These urban technology clusters have grown out of—and continue to support—dominant local industries: entertainment in LA and media and finance in New York. Technology production is harder to spot than in the technology capitals, obviously because technology is a smaller share of local employment (1-2%) but less obviously because the average tech establishment is around half the size (15-17 employees) of those in the technology capitals (29-34 employees). Few big-name technology companies are located here, but smaller start-ups abound.

Third are technology company towns: smaller cities dominated by one or a handful of technology employers. These towns, like Durham County NC, Boulder County CO, and Madison County AL, tend to be manufacturing centres without diversified technology bases, even though technology employment is over 10% of these local economies. Here, companies headquartered elsewhere do their high-skilled production work, aided by the presence of engineers and scientists attracted by nearby universities or military sites.

Finally, there are the suburban technology boomtowns, like Oakland County MI (outside Detroit), Montgomery County MD (outside Washington, DC), and Collin County TX (outside Dallas). Technology employment grew faster in these areas during the period 1988-96 than in other locations. On the other indicators, they look most similar to the three technology capitals, making them the possible challengers to Silicon Valley, Route 128, and Northern Virginia in the next major round of technological innovation.

6 Beyond the United States

The research presented here finds support from—and offers implications for—other countries, both rich and poor. Koski, Rouvinen, and Ylä-Anttila (2001) find strong evidence in Europe for IT-industry concentration, especially when looking at narrower industries within the broader IT classification. At the same time, they find inconclusive evidence on convergence, with countries becoming more similar over time with regard to R&D and value-added but less similar with regard to exports. Why might convergence proceed less definitively in Europe than in the US? The economic growth literature shows that convergence of incomes proceeds most rapidly among places with unimpeded flows of capital, labour, and ideas. Convergence should therefore naturally be faster among the regions of the US, which share common national laws and language, than among Europe's countries, which have different laws, cultures, and languages.[23]

[23] Concentration is more dramatic across US regions than among European countries, as Krugman (1993) points out. Trade barriers and national pride encourage European countries to develop domestic industrial clusters even when a single industrial cluster would most efficiently supply all of Europe.

For poorer countries, this research offers an important lesson for policies intended to attract high-IT industries. At first glance, such policies appear promising: since high-IT industries tend to cluster, first-mover advantage is important for determining which country or region develops a cluster, giving room for governments to offer regulatory and pecuniary incentives to lure these industries. But, as the US experience demonstrates, high-IT clusters persist not because they are technology-intensive per se, but because they tend to rely on high-skilled labour. Low-skill employment clusters in high-IT industries dissolve more quickly. Thus, public policies that attract low-skill support functions for high-IT industries confer only short-term benefits in exchange for potentially large upfront costs. A national or provincial government may be willing to offer incentives that reap short-term benefits if its goal is an economic boom before the next election, but not if its goal is long-term economic development. Unfortunately, unsustainable development is easier, since poorer countries have stronger comparative advantages for lower-skilled than higher-skilled industries. Hype aside, high-IT industries are unlikely to offer poorer countries long-term sustainable economic growth.

7 Conclusion

There is evidence to support both hypotheses of information technology's effect on business location. First, although high-IT industries exhibit slower convergence than other industries, controlling for other factors reveals that information technology raises the speed of convergence and lowers the level of industry concentration. The most plausible channel for this effect is by through knowledge spillovers taking place electronically as well as in person, reducing the need for firms to cluster together in the same location. Technology appears to have little effect on the internal organization of firms; high-IT industries did not exhibit greater separation between back-office establishments and headquarters than low-IT industries did. Because high-IT industries tend to have skilled labour forces, and because skilled industries tend to benefit more from concentration, high-IT industries are likely to remain concentrated geographically. But in the future, as low-skilled industries use information technology more intensively, their geographic concentration might decline considerably.

Second, the new-firm birth analysis suggests that information technology reduces the need for firms to locate near their clients, and might slow convergence for service industries. This result, however, is inconclusive. On balance, information technology is speeding convergence and lowering industrial concentration, with technology employment growing fastest in suburban boomtowns. The future IT-intensive economy should have more stable local wages and unemployment and less reason to use place-based economic intervention.

Data appendix

- *emp* is employment
- *i* indexes industries
- *x* indexes MSA's
- *f* indexes firms

variables for convergence regressions: Tables 1 and 2

unit of observation: industry-MSA cell

- **localization** is industry-MSA employment as fraction of industry employment

$$concentration_{ix} = \frac{emp_{ix}}{emp_i}$$

- **competition** is the competitiveness of the local industry, measured using a firm-level Herfindahl, for each industry-MSA; 1 means total competition; if no employment in 1989, then competition = 0

$$competition_{ix} = 1 - \sum_f \left(\frac{emp_{fix}}{emp_{ix}} \right)^2$$

- **diversity** is measured as the similarity of the local industry mix to the national industry mix

$$diversity_{ix} = -\sum_{j \neq i} \left(\frac{emp_{jx}}{emp_x} - \frac{emp_j}{emp} \right)^2$$

- **industrygrowth** is the employment growth rate of the industry outside the MSA

$$industrygrowth_{ix} = \ln \left(\frac{\sum_{y, y \neq x} emp_{iy}^{1996}}{\sum_{y, y \neq x} emp_{iy}^{1989}} \right)$$

- **MSA size** is total local employment in all other industries

$$MSAsize_{ix} = \sum_{j, j \neq i} emp_{jx}$$

variables for industry-level concentration regressions: Table 3

unit of observation: industry

concentration is the geographic Herfindahl of employment across MSA's

$$concentration_i = \sum_x \left(\frac{emp_{ix}}{emp_i} \right)^2$$

change in concentration = $concentration_i^{1996}$ - $concentration_i^{1989}$

industry-level HQ concentration is the geographic Herfindahl of headquarters employment across MSA's; headquarters employment includes employment in establishments located in the county where the headquarters of the establishment's parent firm is located; employment in single-establishment firms is included

$$HQconcentration_i = \sum_x \left(\frac{HQemp_{ix}}{HQemp_i} \right)^2$$

change in HQ concentration = $HQconcentration_i^{1996}$ - $HQconcentration_i^{1989}$

average distance from HQ for an industry is the weighted average of the distance, in miles, between (1) the centroid of an establishment's county and (2) the centroid of the that establishment's firm's headquarters county; the average is weighted by the establishment's employment; establishments located in the county of their parent firm's headquarters are assigned a distance of zero; establishments that are single-establishment firms are also assigned a distance of zero

change in average distance from HQ is the difference between average distance from HQ in 1996 and average distance from HQ in 1989

References

Barro, Robert, and Xavier Sala-i-Martin (1998). *Economic Growth*. Cambridge, MA: MIT Press.

Castells, Manuel (1989). *The Informational City: Information Technology, Economic Restructuring, and the Urban-Regional Process*. Oxford, UK: Blackwell.

Davis, Steven, John Haltiwanger, and Scott Schuh (1996). *Job Creation and Destruction*. Cambridge, MA: MIT Press.

Dumais, Guy, Glenn Ellison, and Edward Glaeser (1997). 'Geographic Concentration as a Dynamic Process'. *NBER Working Paper #6270*.

Ellison, Glenn, and Edward Glaeser (1997). 'Geographic Concentration in US Manufacturing Industries: A Dartboard Approach'. *Journal of Political Economy* 105 (5): 889-927.

Glaeser, Edward, Hedi Kallal, Jose Scheinkman, and Andrei Shleifer (1992). 'Growth in Cities'. *Journal of Political Economy* 100 (6): 1126-52.

Glendon, Spencer (1999). 'Urban Life Cycles'. Cambridge, MA: Harvard University. Unpublished Ph.D. dissertation.

Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson (1993). 'Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations'. *Quarterly Journal of Economics* 108: 577-98.

Kim, Sukkoo (1995). 'Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in US Regional Manufacturing Structure, 1860-1987'. *Quarterly Journal of Economics* 110 (4): 881-908.

Kolko, Jed (1999). 'Can I Get Some Service Here?' Manuscript.

Kolko, Jed (2000). 'The Death of Cities? The Death of Distance? Evidence from the Geography of Commercial Internet Usage', in Ingo Vogelsang and Benjamin M. Compaine (eds), *The Internet Upheaval*. Cambridge, MA: MIT Press.

Koski, Heli, Petri Rouvinen, and Pekka Ylä-Anttila (2001). 'ICT Clusters in Europe: The Great Central Banana and Small Nordic Potato'. WIDER Discussion Paper No. 2001/6. Helsinki: UNU/WIDER.

Krugman, Paul (1991). 'Increasing Returns and Economic Geography'. *Journal of Political Economy* 99 (3): 483-99.

Krugman, Paul (1993). *Geography and Trade*. Cambridge, MA: MIT Press.

Marshall, Alfred (1892). *Elements of the Economics of Industry*. London: Macmillan.

Robb, Alicia (1999). 'New Data for Dynamic Analysis: The Longitudinal Establishment and Enterprise Microdata (LEEM) File'. Draft.

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

Whyte, William (1988). *City: Rediscovering the Center*. New York: Anchor Books.

UNU World Institute for Development Economics Research (UNU/WIDER) was established by the United Nations University as its first research and training centre and started work in Helsinki, Finland in 1985. The purpose of the Institute is to undertake applied research and policy analysis on structural changes affecting the developing and transitional economies, to provide a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and to promote capacity strengthening and training in the field of economic and social policy making. Its work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

UNU World Institute for Development Economics Research (UNU/WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by Liisa Roponen at UNU/WIDER
Printed at UNU/WIDER, Helsinki

ISSN 1609-5774
ISBN 952-455-128-4 (printed publication)
ISBN 952-455-129-2 (internet publication)