Software, Creativity, and Economic Geography

Software Industry Center
Carnegie Mellon University

December 2003

Richard Florida*
Heinz Professor of Regional Economic Development
H.J. Heinz III School of Public Policy and Management
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh PA 15213
Email: florida@cmu.edu

Kevin Stolarick
Researcher, Software Industry Center
H&SS, SDS/IS, Porter Hall 208
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
Email: kms@andrew.cmu.edu

Brian Knudsen
PhD Student
H.J. Heinz III School of Public Policy and Management
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh PA 15213
Email: knudsen@andrew.cmu.edu

Sam Youl Lee
PhD Student
H.J. Heinz III School of Public Policy and Management
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh PA 15213
Email: samyoul@andrew.cmu.edu

* Corresponding author

Funding was provided by the Alfred P. Sloan Foundation and the Pennsylvania Technology Investment Authority. Gary Gates collaborated on aspects of this research. Anita Sands provided helpful comments.

Abstract: This paper describes and then explores competing hypotheses explaining the economic geography of the software industry. In comparing the more traditional establishment-level data with occupational data – a method that captures people working in software outside of software firms – we find that the economic geography does not conform to the stereotype of sunbelt and bicoastal high-tech hot-spots. Correlation and regression analyses suggest that software is associated with high-tech industry, but also with a broad talent base and with a creative occupational workforce.

Keywords: Software Industry, Geography, Creativity

Introduction

Software is an industry of vital importance. The industry is large, with over $200 billion in revenues in the United States alone (U.S. DOC, 2000). The software industry in the U.S. employs about 1.5 million people directly, and counting those employed in other types of firms it is estimated that some 2.5 million or more people work as software developers or in software production in some capacity (US DOC 2000). Indeed, software now employs more people than the automobile, computer, semiconductor, and steel industries.

Moreover, despite the oft-repeated comment of the Nobel Prize winning economist Robert Solow (1987) that "computers can be found everywhere except in the productivity statistics," recent research suggests that information technology in general, and software in particular, have played a major role in U.S. productivity gains and overall economic growth in the past decade (DeVol, 1999, Jorgenson et al, 2000). But while the economic impact of software has come to be widely studied and debated, the growth dynamics of the industry itself have gotten relatively little attention and are surprisingly under-studied. For instance, it is well known that the software industry in the U.S. is clustered much more densely in some regions than in others. Many economic developers and policy makers have tried to build large Silicon Valley-style clusters in their own regions, with mixed results at best. There is a lack of understanding—and of systematic research—on the factors that underpin and shape the regional distribution of software activity.

This article is intended to help fill that gap. First, we paint a detailed picture of the regional distribution of the software industry, using both occupational and firm-level data to determine where the greatest concentrations of software activity can, in fact, be found. Then, using statistical correlations and regression analysis, we test and compare several theories – such as the industry-clustering model (Porter 1998), the human capital model (Lucas 1988; Glaeser 2000), and the creative capital model (Florida 2001, 2002a, 2002b; Florida and Gates 2001) – that might be used to explain the patterns of concentration. Our central hypothesis is as follows:

We expect software activity to arise and concentrate in places that have a broadly creative habitat that attracts creative and diverse people of all sorts. These places will have the underlying conditions – as well as the requisite talent and the market – for producing software.

We find evidence to confirm the hypothesis. We also find evidence to suggest there is merit in the "industry-clustering" model advanced by Michael Porter and others and in the "human capital" model of Robert Lucas, Edward Glaeser and others. While it hard to isolate any single casual factor that determines the concentration of software jobs and firms, it appears that regions that have a confluence of technology, talent and an overall creative habitat also do well in attracting software activities and jobs.

The Software Industry and Regional Growth

Software differs from most other industries in fundamental ways. For one thing, the basis for competitiveness is different. In older industries such as automobiles and steel — and even in many newer high-technology industries such as electronics — manufacturing productivity and quality are important, as are cost factors such as plant and equipment, materials and transportation. In software, however, competitiveness revolves principally around design. Software can be thought of as an "industry of the mind"—where knowledge, intelligence and human creativity are the crucial inputs, and people are the crucial asset.

It thus is not surprising that traditional regional economic advantages, such as natural resources or access to transportation routes, count for little in software. And many regions, well aware of this, have tried new development strategies. The most common strategy is using a variety of business incentives and programs to attract, or develop, software *firms*. But not all software activity occurs within those firms. Because software is a ubiquitous, general-purpose technology that affects virtually every sector of the economy, from farming and textile production to business services, entertainment, and medical care, software workers work across many sectors of the economy. In fact, recent research indicates that as many as three-quarters of software developers — the people who do the actual creating of software, as distinct from administrative and support people — work outside "the software industry" *per se*. Furthermore, other studies indicate that only a minority of software workers actually has degrees in computer science, software engineering, or related disciplines. Thus software appears to be a field in which *"creativity matters more than credentials."*

Older theories of economic growth and development emphasized the role of natural resources and physical assets. In recent years, several more robust theories have emerged. The first, associated with the work of Porter (1998) and others, emphasizes the role of *clusters* of related and supporting industries. According to this view, clusters operate as geographic concentrations of interrelated firms in which local sophisticated customers, along with strong competition from other firms in the same industry, drive innovation and growth. A second view associated with Lucas (1988) and Glaeser (1998) focuses on the role of human capital – that is, *highly educated people*. It argues that places with higher levels of human capital are more innovative and grow more rapidly and robustly over time.

A third view, associated with Florida (2002), emphasizes the role of creative capital. It argues that certain underlying conditions of places — such as their ability to attract *creative people* (not just highly educated people) and to be open to diversity — inform innovation and growth. In their analyses of U.S. metropolitan areas, Florida and Gates have found that the social character of these regions appears to a very large influence over their economic competitiveness. (Florida

2001, 2002a, 2002b; Florida and Gates 2001) In particular, they found that those places that offer a high quality of life and best accommodate diverse population groups enjoy the greatest success in talent attraction/retention and in the growth of their technology industries. Further independent research by Robert Cushing (2001) of the University of Texas at Austin provides a good deal of support for the creative capital view (See Florida 2002a, chapter 15). From this basic premise, we therefore argue that the creative capacities and talent required for the software industry will tend to concentrate in places that attract creative talent, and serve to mobilize and harness creativity, more generally.

Data and Methods

This report introduces a variety of data to test our hypothesis along with others. For software employment, we use not only traditional establishment-level (or firm-level) measures, but develop new measures to capture people who work in software occupations across many industries. In the past, analysts have measured industry employment by counting employees in firms or business establishments within that industry. But recent research by Markusen et al (2001) notes the problems of using establishment-level data to estimate high-tech employment. They find that occupation-level data (that is, counting workers by occupational classification regardless of employer) gives a much fuller and richer perspective — and as already noted, we would expect this to be especially true for software.

Our establishment-level data are taken from the Census County Business Patterns (CBP) dataset for 1998, the most recent year available. Our occupation-level data comes from the Bureau of Labor Statistics, Occupation and Employment Survey (OES) dataset for 2000. We also develop location quotients based on these data. This study examines these trends at the metropolitan level and uses both CMSA and PMSA classifications in different analyses.

After describing the regional concentration of the software industry, we then turn to factors that may condition this economic geography. We explore the role of factors including the presence of technology-based industry generally, commercial innovation (i.e., patenting), university presence, university R&D spending, human capital, and various new measures of creativity and diversity: the relative presence in a region of "creative class" and "super-creative" workers, artists, gays, foreign-born people and others.

We frame our analysis in terms of the "3 Ts" model of economic growth introduced by Florida (2002). The 3 Ts model posits that technology, talent and tolerance are interrelated, and that all three must be present in a region for strong growth to occur. We employ both bivariate and multivariate models to test the relative effect of technology concentration, human capital, creativity and diversity on the regional distribution of software employment.

The Economic Geography of the Software Industry

We now examine the economic geography of the software industry, looking first at employment in software occupations across all industrial sectors before turning to the regional distribution of software-industry establishments.

Software Employment

Table 1 shows the top 25 regions in terms of total employment in software occupations across all sectors of the economy. New York has the greatest absolute number of software workers followed by San Francisco, greater Washington DC, and Los Angeles. Four regions have more than 100,000 software workers, while an additional seven have more than 50,000. Boston, Chicago, Dallas, Seattle and Philadelphia round out the top 10. While there are few surprises among the top 10 software regions, the top 25 is not a listing of high-tech hotspots. As Table 1 shows, several Midwest regions, including Kansas City, St. Louis and Columbus number among the top 25, as well as two Florida regions, Miami and Orlando.

Table 1: Software Employment, by Occupation

Rank	Region	Total Software Employment	Percent Software
1	New York, NY	186,430	2.3%
2	San Francisco, CA	173,420	5.5%
3	Washington DC-Baltimore, MD	145,800	4.1%
4	Los Angeles-Long Beach, CA	102,400	1.7%
5	Boston, MA-NH	93,380	3.5%
6	Chicago, IL	88,840	2.3%
7	Dallas, TX	77,370	3.1%
8	Seattle, WA	61,320	3.9%
9	Atlanta, GA	59,730	3.1%
10	Philadelphia, PA	58,070	2.2%
11	Denver, CO	56,050	4.5%
12	Detroit, MI	43,560	2.1%
13	Minneapolis-St. Paul, MN-WI	42,680	2.8%
14	Houston, TX	41,230	2.2%
15	Phoenix-Mesa, AZ	35,790	2.5%
16	San Diego, CA	29,590	2.8%
17	Austin-San Marcos, TX	27,270	4.7%
18	St. Louis, MO-IL	26,810	2.2%
19	Raleigh-Durham-Chapel Hill, NC	25,990	4.4%
20	Kansas City, MO-KS	23,610	2.7%
21	Tampa-St. Petersburg-Clearwater, FL	22,840	2.1%
22	Portland, OR-WA	22,500	2.3%
23	Columbus, OH	20,900	2.8%
24	Miami, FL	19,770	1.4%
25	Orlando, FL	18,690	2.4%

It is also useful to examine the share of software workers as a percent of total employment (see Table 2). The San Francisco Bay Area is now in the top position, Austin 4th, the Research Triangle 6th, Greater Washington 7th, Seattle 9th and Boston 13th. The leading regions have more than 5 percent of their total employment in software occupations, while most of the top ten have in the range of 4 percent. On the percent measure, 14 of the top 25 regions are large, with populations of one million or above. But here again, we observe some interesting surprises. Huntsville, Alabama; Provo, Utah and Sioux Falls, South Dakota make the top 10. Omaha, Nebraska; Cedar Rapids, Iowa; Hartford and Kansas City also make the top 25. Seven metros have populations under 500,000, while 3 of these have fewer than 250,000.

Table 2: Percent Software Employment, by Occupation

Rank	Region	Percent Software	Total Software Employment
1	San Francisco, CA	5.5%	173,420
2	Colorado Springs, CO	5.2%	10,000
3	Huntsville, AL	4.8%	6,480
4	Austin-San Marcos, TX	4.7%	27,270
5	Denver, CO	4.5%	56,050
6	Raleigh-Durham-Chapel Hill, NC	4.4%	25,990
7	Washington DC-Baltimore, MD	4.1%	145,800
8	Provo-Orem, UT	4.0%	4,530
9	Seattle, WA	3.9%	61,320
10	Sioux Falls, SD	3.6%	3,460
11	Fort Collins-Loveland, CO	3.5%	3,250
12	Madison, WI	3.5%	7,920
13	Boston, MA-NH	3.5%	93,380
14	Omaha, NE-IA	3.5%	12,640
15	Cedar Rapids, IA	3.3%	3,210
16	Dallas, TX	3.1%	77,370
17	Atlanta, GA	3.1%	59,730
18	Hartford, CT	3.0%	16,820
19	San Diego, CA	2.8%	29,590
20	Columbus, OH	2.8%	20,900
21	Minneapolis-St. Paul, MN-WI	2.8%	42,680
22	Richmond-Petersburg, VA	2.7%	12,560
23	Albuquerque, NM	2.7%	7,610
24	Kansas City, MO-KS	2.7%	23,610
25	Charlottesville, VA	2.5%	1,410

Table 3 shows the top 25 regions in terms of software establishment employment based on our analysis of all 276 regions (MSAs and CMSAs) nationwide. To control for differences in population size, we simply look at software establishment employment as a percentage of total employment. The Greater Washington DC region (including Baltimore) is the leading region with 4.4 percent of total employment in software establishments. The 2nd and 3rd place regions may come as something of a surprise: Bryan-College Station, Texas and Provo, Utah, where software establishments account for more than 3 percent of total employment. Nor are these the only surprises: Huntsville, Alabama; Omaha, Nebraska; Columbia, South Carolina and Melbourne, Florida also make the top 10, while Lawton, Oklahoma; Detroit, Michigan and Little Rock, Arkansas are among the top 25. Two regions typically thought of as software and IT hotbeds, Austin, Texas and the San Francis Bay Area, rank 10th and 11th. The Research Triangle ranks 17th and the greater Boston region ranks 18th. Interestingly, 9 of the top 25 regions have populations under 500,000, while Bryan-College Station, Texas has fewer than 250,000 residents. The top 25 regions are widely distributed across the country and do not conform to a so-called bi-coastal pattern. This finding is consistent with Markusen et al's finding (2001) that overall, high-tech is not coastally isolated or found solely in sunbelt cities.

Table 3: Percent Software Employment, by Establishment

Rank	Region	Percent Software	Total Software Employment
1	Washington DC-Baltimore, MD	4.4%	136,065
2	Bryan-College Station, TX	3.9%	1,667
3	Provo-Orem, UT	3.3%	4,374
4	Huntsville, AL	2.9%	3,990
5	Omaha, NE-IA	2.7%	9,790
6	Columbia, SC	2.6%	6,168
7	Denver, CO	2.5%	29,024
8	Columbus, GA-AL	2.5%	2,520
9	Melbourne-Titusville-Palm Bay, FL	2.3%	3,549
10	Austin-San Marcos, TX	2.2%	10,537
11	San Francisco, CA	2.2%	68,489
12	Atlanta, GA	2.1%	39,493
13	Dallas, TX	2.0%	46,709
14	Colorado Springs, CO	1.9%	3,727
15	Lawton, OK	1.9%	523
16	Burlington, VT	1.9%	1,664
17	Raleigh-Durham-Chapel Hill, NC	1.9%	10,122
18	Boston, MA-NH	1.8%	51,437
19	Little Rock-North Little Rock, AR	1.7%	4,621
20	Minneapolis-St. Paul, MN-WI	1.7%	26,409
21	Iowa City, IA	1.7%	770
22	Detroit, MI	1.7%	38,997
23	Hartford, CT	1.7%	9,185
24	Tallahassee, FL	1.7%	1,596
25	Kansas City, MO-KS	1.7%	13,821

It is useful and interesting to compare across these two definitions of software employment (Tables 2 and 3). Three of the top 5, six of the top 10, and 15 of the top 25 regions are shared by both lists. But there are considerable differences, as noted above. While large regions are somewhat over-represented in both lists, it is clear that regions of all sizes exhibit relatively high concentrations of software employment defined both in terms of establishments and occupations.

Location Quotients for Software Employment

To get a better handle on the geography of software employment we employ location quotients. As opposed to simply looking at percentages, location quotients allow us to more fully explore the concentration of software activity at the regional level, observing whether or not software employment is over-or under-represented with respect to the national average. A location quotient of 1 equals the national average. A location quotient of 2 means that a region has double the expected national average, whole 0.5 means that it has half.

We also shift the unit of analysis to get a more fine-grained sense of the geography of software industry employment. We now use the PMSA (or primary metropolitan statistical area) instead of the CMSA (or consolidated metropolitan statistical area). This enables us to zero in on sub-units of broader regions. For example, the San Francisco Bay Area region is made up of three smaller regions—San Francisco, San Jose/ Silicon Valley, and Oakland-Berkeley. Table 4 shows the top 25 regions by location quotient.

Three regions have location quotients over 3: Boulder, San Jose/ Silicon Valley and San Francisco. An additional 12 regions have location quotients over 2, including well-known high-tech centers like Washington DC, Austin, Boston, and Seattle. Here again there are some surprises: Omaha, Cedar Rapids, Sioux Falls and Wilmington, Delaware all rank in the top 25. The regional distribution again defies the bi-coastal myth. Software regions are widely distributed across the country, with places like Sioux Falls and Omaha doing quite well. Again, while large regions are somewhat over-represented, small and medium-sized regions show an ability to compete.

Table 4: Software Employment Centers: Location Quotients

Rank	Region	Location Quotient
1	Boulder-Longmont, CO	5.73
2	San Jose, CA	4.68
3	San Francisco, CA	3.17
4	Lowell, MA-NH	2.82
5	Colorado Springs, CO	2.69
6	Washington, DC	2.66
7	Dutchess County, NY	2.56
8	Huntsville, AL	2.51
9	Austin-San Marcos, TX	2.46
10	Seattle-Bellevue-Everett, WA	2.39
11	Raleigh-Durham-Chapel Hill, NC	2.31
12	Middlesex-Somerset-Hunterdon, NJ	2.23
13	Boston, MA-NH	2.16
14	Provo-Orem, UT	2.09
15	Trenton, NJ	2.00
16	Denver, CO	1.97
17	Sioux Falls, SD	1.87
18	Fort Collins-Loveland, CO	1.84
19	Madison, WI	1.83
20	Dallas, TX	1.83
21	Omaha, NE-IA	1.81
22	Nashua, NH	1.74
23	Cedar Rapids, IA	1.74
24	Wilmington-Newark, DE-MD	1.66
25	Atlanta, GA	1.62

Explaining the Software Industry's Economic Geography

We now turn our attention to trying to explain these observed geographic patterns. As noted earlier, there are several competing theories that might be used. Some would say that software activity is related to overall levels of economic activity, or more precisely to levels of high-technology economic activity. According to this view, software employment would tend to concentrate and grow in places with relatively high levels of innovation, strong universities, or high concentrations of high-tech firms overall. Others would argue that as a knowledge-driven industry, software employment would tend to cluster in places with high levels of human capital, especially where supplies of professional and technical labor are abundant. They might further contend that this human capital effect would be enhanced by the well-documented "talent shortage" in the software industry and furthermore, as recent research shows, that the software talent pool is not drawn from specific disciplines and fields but is drawn broadly from the ranks of highly skilled people in general.

Our perspective is somewhat different. We contend that software employment will tend to be concentrated in broadly creative centers—that is, places that are open to creative people across the board and that also have low barriers to entry and thus are characterized by high levels of demographic diversity. In this section, we organize our findings along the lines of the "3 T" model of economic growth outlined in Florida (2002). This model suggests that three factors – talent, technology, and tolerance – all need to be simultaneously present in a region for growth to occur. We thus examine the relationships between software activity and indicators of each of the three T's. Our analysis is based on occupational data organized at the sub-region or PMSA level.

Technology

This section examines the relationship between software employment and various indicators of technology intensity. We consider the relationships between software employment, high-technology industry, innovation and university strength respectively.

Software and High-Technology Industry

It is likely that software industry employment will be concentrated in places that are high-technology centers broadly. Here we use the High-Tech Index, originally called the "Tech-Pole" and developed at the Milken Institute (DeVol et al 1999) as a measure of relative high-tech industry presence. Not surprisingly, we find a positive relationship between the High-Tech

Index and software employment (see Fig. 1). Clearly, software is an important component of a regional technology presence.

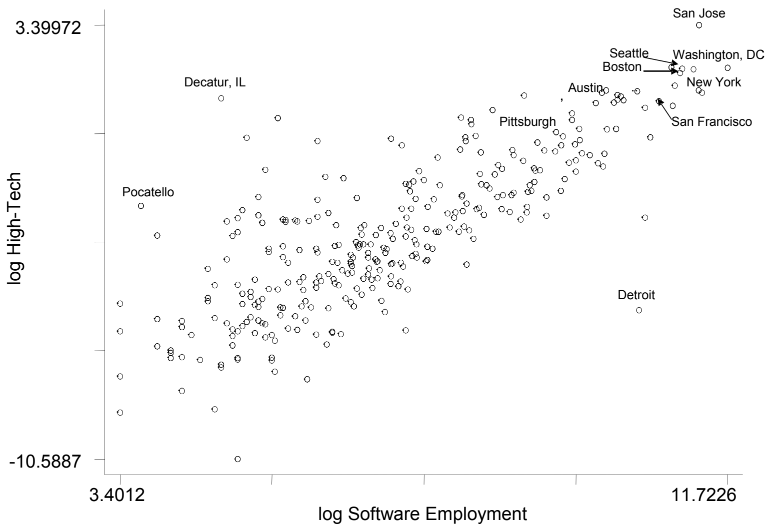

Figure 1: Software and High-Tech Industry

Software and Technological Innovation

According to many leading social theorists and economists, innovation and innovative activity contribute to regional growth. Thus, we are interested in how the regional presence of innovation (as measured by patents) relates to - software employment. As Fig.2 shows, there appears to be a positive relationship between software employment and simple patent counts.

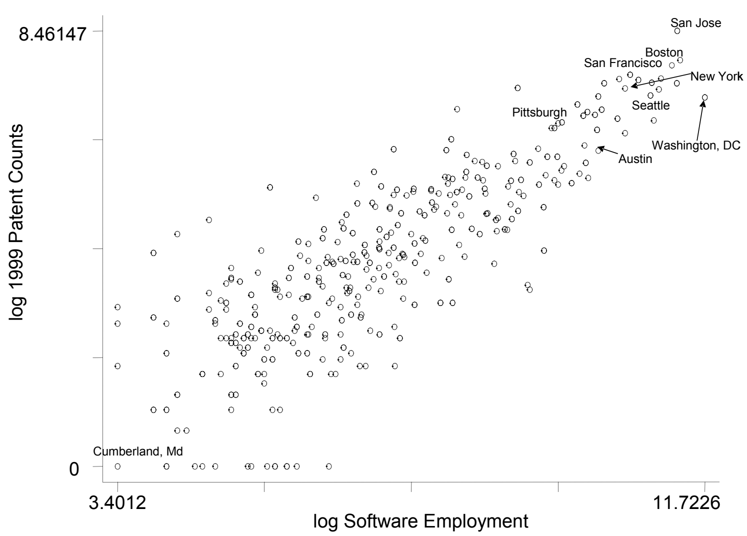

Figure 2: Software and Innovation

Software and the University

Increasingly, universities are credited with playing a central role in economic development broadly, and in technology-fueled growth more specifically. This growth is related both to the research and development capabilities of universities, and also to their capacity to produce and to attract high human capital people. It might also be thought that university strength would be positively related to software employment, given the high skill requirements and knowledge intensity of most software occupations. We employ a rather straightforward measure of university presence based on students and faculty per capita. As Fig. 3 shows, there is no discernible relationship between metro area software employment and per-capita university students and faculty. This could arise from the university measure, which is admittedly crude.

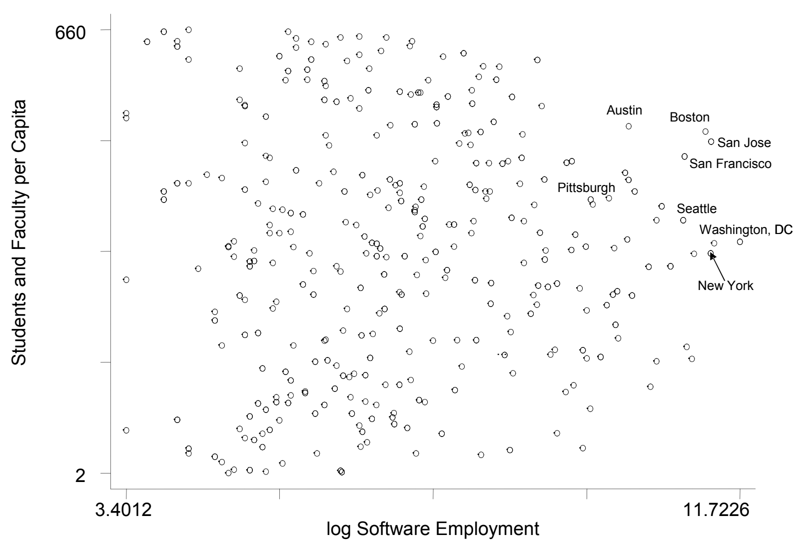

Figure 3: Software and University Presence

Talent

We now turn to the relationship between talent levels and software employment. Glaeser, Lucas, and Romer have all recently discussed the importance of human capital in promoting growth. Florida also includes it as one of his "3 Ts" of economic development – the "talent" of technology, talent, and tolerance. Therefore, we anticipate a positive relationship, since knowledge and skills are prerequisites for software development.

Software and Human Capital

We examine the effects of two widely used measure of human capital—a basic Talent Index, measured as the percent of the population with a bachelor's degree or above, and a more specialized measure of professional and technical workers. As Figure 4 shows, we find a strongly positive relationship between software employment and professional and technical workers. In fact, it is almost possible to detect a slight curvilinearity. We also observe a positive correlation between degree holders and software employment. Not surprisingly, software employment is closely associated with both a broad human capital base and the availability of more specialized human capital in the form of professional and technical workers. This is likely due to the combined effects of the talent pool and the increased market for software products and services that develops in regions with such a human capital profile.

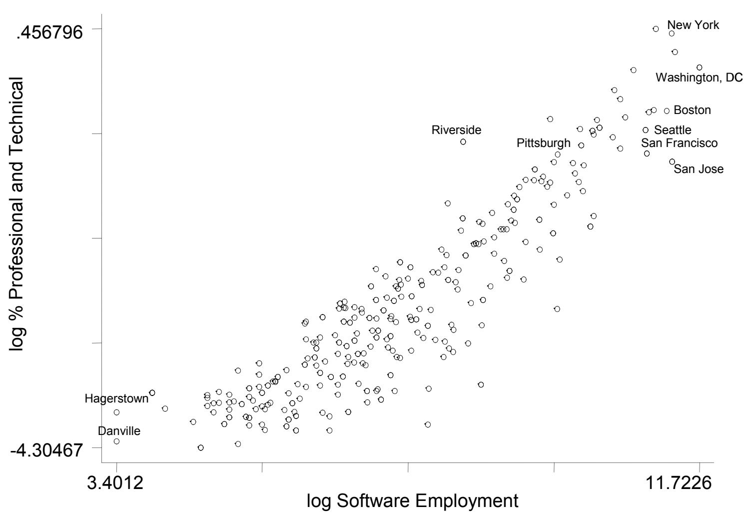

Figure 4: Software versus Professional and Technical Employment

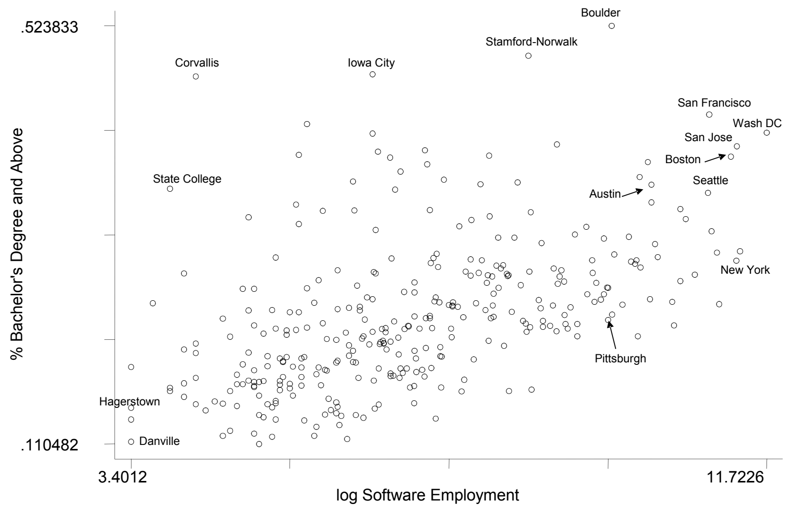

Figure 5: Software versus Talent (BA and above)

Tolerance

Tolerance is the third "T". Florida and Gates (2002, Florida 2002) have found a strong connection between high-technology industry and various measures of demographic diversity, such as their Melting Pot Index (percent of the population that is foreign born) and Gay Index (percent of gay couples, expressed as a location quotient). Florida (2001, 2002) has also found a strong relationship between high-technology industry and innovation and places that are open to artistic and cultural creativity, measured by the Bohemian Index—the percent of the population principally employed in artistic and cultural fields (again expressed as a location quotient). Places that are accepting of various lifestyles, races, and ethnicities are also places where new ideas will be welcome and will be likely to grow into innovations. Individuals whose livelihoods are centered on ideas will thus seek out tolerant places in an attempt to increase their productivity.

Software and Diversity

We now examine the relationship between software employment and four measures of diversity: the Bohemian Index, Gay Index, Melting Pot Index, and a Composite Diversity Index, which combines these three measures into an overall index (see Figs 6 through 9). We observe positive relationships between software employment and all four. These relationships align with recent theories about the role of tolerance in promoting growth.

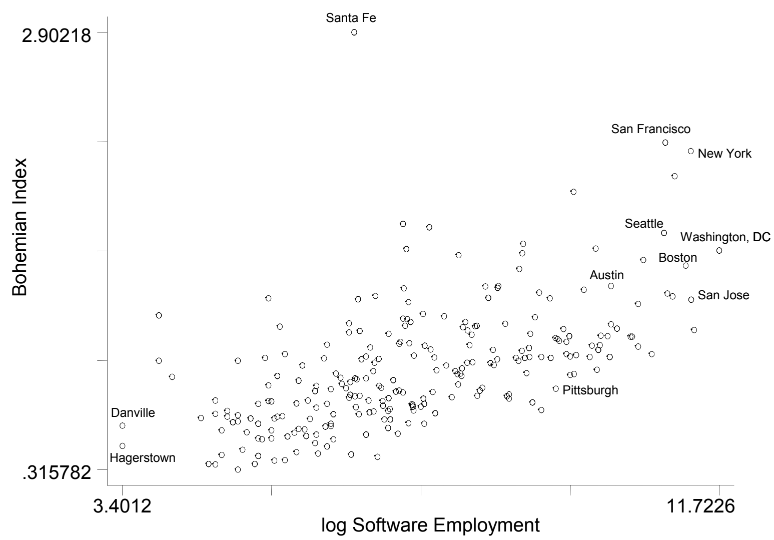

Figure 6: Software and the Bohemian Index

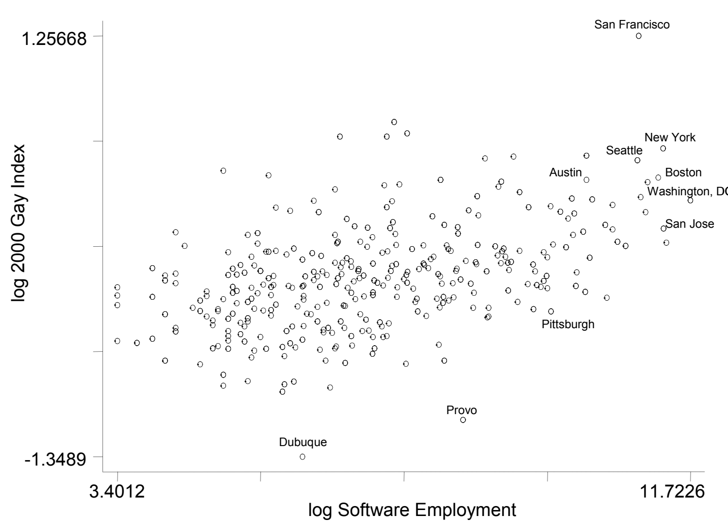

Figure 7: Software and the Gay Index

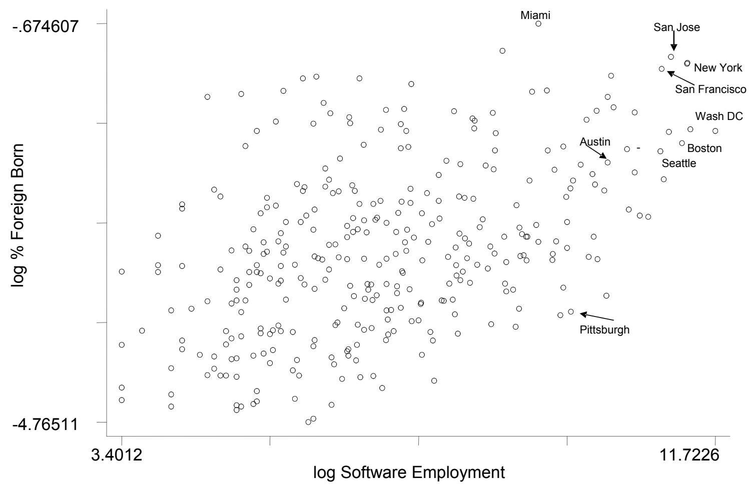

Figure 8: Software and the Melting Pot Index (Percent Foreign Born)

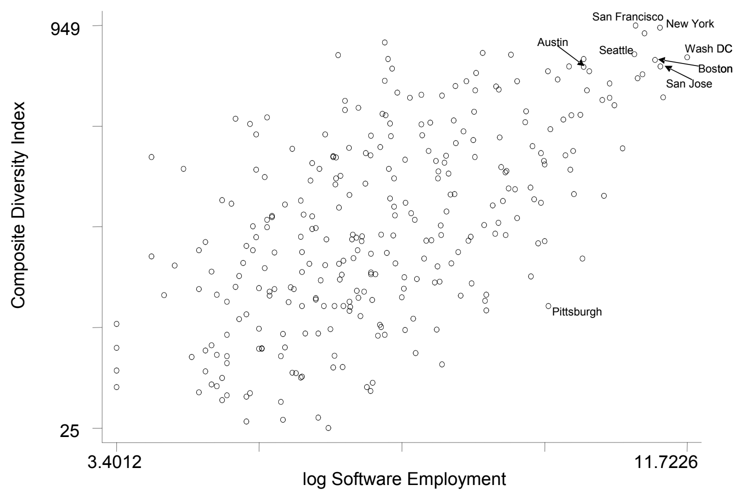

Figure 9: Software and the Composite Diversity Index

The Role of Creativity

Software is a creative industry—an industry of the mind. Florida (2002) claims that creativity has become the most important economic force. He finds that the creative sector (which includes all creative professions) now accounts for roughly half the wages and salaries paid in the U.S., and that the creative class (those who work in the creative sector) account for roughly 30 percent of all employment. In this section, we explore the relationship between software employment and regional concentrations of the creative class.

Software and the Creative Class

Florida (2002) separates the creative class into two sub-groups—the super-creative core (composed of scientists and engineers, artistic and cultural creatives, and educators), and a surrounding group of creative professionals who make up the rest of the class. Figure 10 shows the relationship between software employment and the super-creative core. The relationship is extremely positive and robust—almost perfectly linear. Figure 11 shows the relationship between software employment and the broader creative class. The relationship is also positive—but not as strong as for the super-creatives.

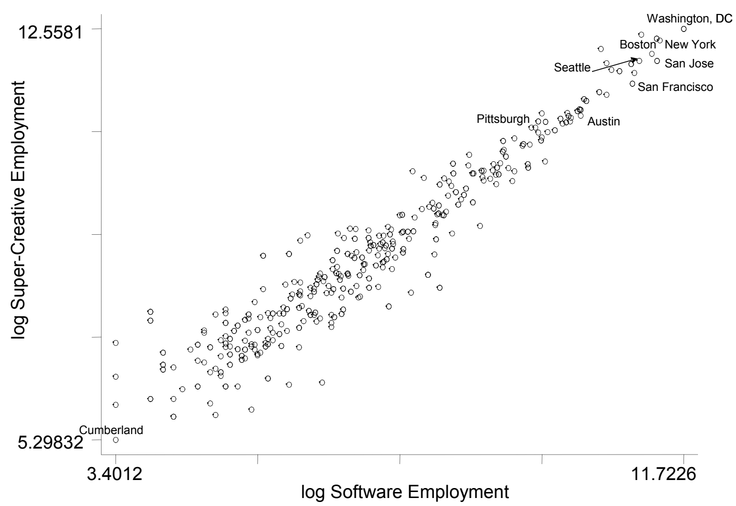

Figure 10: Software and Super-Creative Employment

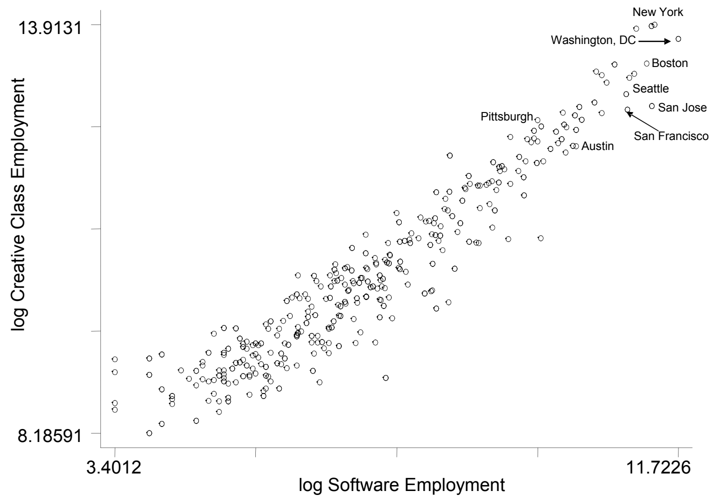

Figure 11: Software and Creative Class Employment

Statistical Findings

We have found that software employment is associated with a range of possible factors having to do with technology, talent and tolerance. In this section we try to determine the relative importance of these factors in accounting for regional concentrations.

Correlation Analyses

We begin by calculating the statistical correlations between regional software employment and the various measures and indicators outlined above (see Table 5). A correlation of 1.0 indicates a perfectly linear, fixed relationship — one in which all the points on an x-y plot would fall along a rising straight line —while a correlation of 0.0 indicates no relationship. Generally speaking correlations above 0.5 suggest a reasonably strong relationship. Negative values indicate inverse relationships.

Not surprisingly, the strongest correlation is between software and IT employment (0.988). The correlations between software employment and the two creativity measures (the super-creatives and the creative class generally) are also very strong, in the range of 0.95. The correlation between software employment and professional and technical workers is only slightly less than that for the creative class measures (0.90). The correlation with talent — the raw human-capital measure — is not as high (0.597), but still fairly strong.

Overall high-tech industry output and commercial innovation are both strongly associated with software employment. The correlations between software employment and the High-Tech Index and Patents are in the range of 0.80. Diversity is positively and significantly correlated with software employment. The highest correlation is with the Composite Diversity Index (0.69) followed by the Bohemian Index (0.59), the Gay Index (0.47) and the Melting Pot Index (0.40).

Table 5: Correlations for Software Employment

	Software Employment (log)
IT Employment (log)	0.9877
Super-Creative Core (log)	0.9584
Creative Class (log)	0.9475
Professional & Technical (log)	0.9030
Patents 1999 (log)	0.8262
High-Tech Index 2000 (log)	0.7927
Composite Diversity Index	0.6864
Talent Index (% BA and above)	0.5971
Bohemian Index	0.5857
Gay Index (log)	0.4656
Melting Pot Index (log % Foreign Born)	0.3957
University Presence	0.0276

Note: All correlations except that with University Presence are *significant* at p = 0.0001. (This means the probability that such a relationship would occur randomly, due to mere chance, is no more than 0.0001, or 1 in 10,000). The correlation with University Strength is not significant at p = 0.05.

Software employment is not significantly correlated with University Presence. This tends to agree with other research we have conducted, in which we find that the economic benefits of a university do not necessarily accrue locally. The students educated at a university, and the technologies developed there, can migrate elsewhere, and in fact some regions appear to serve as net "donors" or "breeders" of university talent and technology while other regions reap the harvest (Florida et al 2003, in process). But here it must be noted that our University Presence measure is a fairly crude one, capturing only the overall numbers of students and faculty per capita. It does not reflect whether the university has a strong computer science program, with a high level of related research — and many believe that these factors play a large role in driving software industry growth and employment in a region. Thus in the next step of our analysis we introduce a new measure, University R&D, which is sponsored research and development dollars at universities in a particular PMSA. Although this measure does not isolate software-intensive R&D, most university R&D dollars go into scientific and technical fields, and the measure may serve as at least a rough proxy for relative levels of software research.

Regression Analyses

To get a more precise handle on the factors that affect the economic geography of the software industry, we conducted several regression analyses. Regression analysis is a complex tool used to estimate, or confirm, functional cause-and effect relationships between a dependent variable and one or more independent variables that are thought to be driving it. The idea is to try to capture the contribution of *each* independent variable, while holding the others constant and thus "controlling for" them.

Our dependent variable is software employment (expressed as a percent of overall regional employment). Our independent variables are all the factors examined thus far. We also include a measure of regional population to account for market size, plus the University R&D measure just discussed.

An important cautionary note is that our regressions are intended to be *exploratory*. Our goal here is not to arrive at actual equations for expressing (and thus predicting) the movement of software employment in response to the independent variables. We are simply trying to better gauge the effects of creativity and diversity variables.

Recall that we contend a region's ability to generate and/or attract software employment is related to its underlying creative capabilities and its openness to diversity. We thus believe that our creativity and diversity measures should complement measures like market size, technology, and/or human capital in explaining the economic geography of software employment.

Table 6 summarizes our regression results. The table needs a bit of explanation beforehand. Always using software employment as our *dependent* variable, we examined its relationship to each of the various independent factors under each of three different models. In model (1), we included the overall Creative Class measure among our independent variables. In model (2), we left this out and included only the Super-Creative Core. In model (3), we left out the Creative Class altogether — in order to try to isolate the effects of diversity variables, which we believe affect Creative Class concentrations in the first place. In model (4), we substitute the Composite Diversity Index (CDI) for our three individual diversity measures. The CDI is a combination of these three measures. The findings presented in Table 6 are based on 164 metro areas, that being the number for which all desired data were available.

The creative class variables are significantly associated with software employment. Most noticeable is the effect of super-creative employment on software employment (Model 2 above). Talent also appears to also be a statistically significant predictor of software employment. The basic Talent Index (percent BA and above) enters positive and significant into all four regressions. Somewhat surprisingly, the role of scientists and engineers in explaining software employment is negative in all four models. "Scientists and Engineers" is a significant variable in all four models as well.

The role of diversity is mixed. The coefficient for the Bohemian Index is positive in all three models but insignificant, while the coefficient for the Gay Index is positive in model 2, negative in models 1 and 3, but also insignificant. The Melting Pot Index is positive in models 1 and 2, but insignificant. The CDI appears to be negatively, but insignificantly related to software employment.

Table 6: Regression Results

Dependent Variable: Regional Software Employment (as % of total employment)

	Regression Model (1)	Regression Model (2)	Regression Model (3)	Regression Model (4)
Independent Variables:				
High-Tech Index (Milken 2000)	0.002***	0.001***	0.002***	0.002***
Creative Class	0.105***	—	—	0.095***
Super-Creative Core	—	0.324***	—	
Gay Index	-0.001		-0.001	
Bohemian Index	0.004		0.004	
Melting Pot Index (% Foreign Born)	0.008	0.003	-0.005	
Composite Diversity Index	—	—	—	-0.000002
Talent Index (% Bachelor's Degree +)	0.105***	0.074***	0.142***	0.122***
Scientists and Engineers	-0.199**	-0.175***	-0.173*	-0.179*
University Presence (Students & Faculty Per Cap)	-0.00002***	-0.00001***	-0.00002***	-0.00002
University R&D ($ 2000)	-0.0002	-0.0002	-0.0001	-0.0002
Metro Population	0.001	0.001	0.003	0.002
Constant	-0.042***	-0.026***	-0.048***	-0.046***
Adjusted R-Squared	0.7505	0.8302	0.7158	0.7483

* significant at 0.05
** significant at 0.01
*** significant at 0.005 N = 164 metro areas, all three models

Overall high-tech industry output (as measured by the High-Tech Index) is positively and significantly associated with software employment. Neither University Presence nor University R&D is positively associated with software employment. In fact, for University Presence, the association is *significantly* negative. Market size (measured as 1990 population) has a negligible effect on software employment.

Conclusions

Our analysis has led us to a number of interesting findings. Using occupational data along with establishment-level data — a method that allows us to capture people working in software outside software firms — we find that the economic geography of the software industry defies simple patterns or imagery. On the one hand, regions like Silicon Valley, the broader San Francisco Bay Area, Austin, Seattle, Boston and Washington DC are clearly important centers of software employment. But places like Omaha, Nebraska; Sioux Falls, South Dakota and Cedar Rapids, Iowa are significant centers as well. The economic geography of this industry does *not* conform to the stereotype of sunbelt and bicoastal high-tech hot-spots. Regions throughout the country—small and large alike—have shown an ability to compete as centers of software employment.

Our statistical analyses also inform some deeper findings that are perhaps more interesting. First, our results support the conventional wisdom that software employment is associated with regional concentrations of high-technology industry generally. This lends credence to the clustering theory of Porter and others, which says that similar or related firms tend to cluster and thus gain productive efficiencies and spur innovation.

But our analyses add to, and enrich, this view. They indicate that software employment is also associated with the presence of a broad regional talent base, thus also lending support to the human capital theory of regional development associated with Lucas and Glaeser. Here we observe that software workers, who are skilled high-human-capital people, tend to locate in places with many other educated, professional workers.

We find, further, that software employment is indeed related to regional creativity. Regions with high concentrations of creative workforce tend to also have high levels of software employment. This, we believe, stems from their broader creative milieu or habitat which functions to attract software workers along with other creative people. The process of creating software, as mentioned earlier, is unlike production processes in other industries. Most of the value created in this process stems from the skill, knowledge, and creativity of the workers. Thus, as Florida (2002) concludes, people who incorporate creativity into their employment will also seek out creative environments in which to live and play. Places that are broadly creative will consequently tend to experience higher rates of software employment, along with higher rates of innovation and growth.

Taken together, our findings suggest that the economic geography of the software industry is driven by the broad talent base and creative capacities of regions in addition to technology. The overall results strongly support the creativity thesis. They suggest that creativity is an important factor in the economic geography of software employment; and also that human capital is very important. Thus, our findings suggest that the economic geography of software employment is a function of a broad talent base and/ or underlying creative capacity, as opposed to being solely a function of a more specialized scientific and engineering base.

These findings have several implications for those who wish to stimulate software clusters as part of a regional economic development strategy. They suggest that there is good reason to question conventional approaches which seek to invest in and promote specific high-tech sectors and clusters, or which seek to increase the volume of university technology transfer and commercialization as a mechanism for tech-based economic development. The findings indicate, rather, that regional competitiveness in software stems from broader capacities to generate, attract and mobilize talented and creative people across the board. This in turn suggests that regional development may be better served by broader approaches to developing a habitat or ecosystem that can feed these underlying human creative capabilities. This habitat can then provide the talent base, capacities, user-base, and market demand, in turn stimulating technological innovation, new firm formation, and business growth not just in the software industry but also across the emerging creative economy.

References

Cortwright, J., and Mayer, H. 2001. *High-Tech Specialization: A Comparison of High Technology Centers*. Washington DC: Center on Urban and Metropolitan Policy, The Brookings Institution.

-------------. 2002. Signs of Life: The Growth of Biotechnology Centers in the U.S." The Brookings Institution, Center on Urban and Metropolitan Policy.

DeVol, R., Wong, P., Catapano, J., and Robitshek, G. 1999. America's High-Tech Economy: Growth, Development, and Risks for Metropolitan Areas". Santa Monica, CA: Milken Institute.

Florida, R. 2002a. *The Rise of the Creative Class*. New York: Basic Books.

-------------. 2002b. Bohemia and economic geography. *Journal of Economic Geography*, 2: 55-71.

-------------. 2002c. The economic geography of talent. *Annals of the American Association of Geography* 92: 743-755.

Florida, R. and Gates, G. 2001. *Technology and Tolerance: The Importance of Diversity to High-Technology Growth.* Washington, DC: Center on Urban & Metropolitan Policy, The Brookings Institution.

Florida, R., Stolarick, K., Gates, G., and Knudsen, B. 2003. The Role of the University in the Creative Economy. Working paper, Heinz School of Public Policy and Management, Carnegie Mellon University, Pittsburgh.

Glaeser, E. 1998. Are cities dying? *Journal of Economic Perspectives* 12:139-160.

-------------. 2000. The New Economics of Urban and Regional Growth. In *The Oxford Handbook of Economic Geography*, ed. G. Clark, M. Gertler, M. Feldman, 83-98. Oxford: Oxford University Press.

Hall, B., Jaffe, A., and Tratjenberg, M. 2001. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. NBER Working Paper 8498.

Jorgenson, D., and Stiroh, K. 2000. Raising the Speed Limit: U.S. Economic Growth in the Information Age. *Brookings Papers on Economic Activity*, Vol. 1.

Lucas, R. 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22:1-42.

Markusen, A., Chapple, K., Schrock, G., Yamamoto, D., and Yu, P. 2001. High-Tech and I-Tech: How Metros Rank and Specialize. Working Paper, Project on Regional and Industrial Economics, Humphrey Institute of Public Affairs, University of Minnesota.

Porter, M. November-December 1998. Clusters and the New Economics of Competition. *Harvard Business Review* 77-90.

Simon, C. 1998. Human Capital and Metropolitan Employment Growth. *Journal of Urban Economics* 43: 223–43.

Solow, R. 1987. We'd better watch out. *New York Times Book Review* (July 12): 36.

U.S. Department of Commerce, International Trade Administration. U.S. Industry and Trade Outlook, 2000. McGraw Hill and Company.