Syracuse University

SURFACE

Center for Policy Research

Maxwell School of Citizenship and Public Affairs

1999

Geography, Industrial Organization, and Agglomeration

Stuart S. Rosenthal
Syracuse University. Center for Policy Research, ssrosent@maxwell.syr.edu

William C. Strange
University of Toronto, Rotman School of Management

ISSN 1525-3066

Center for Policy Research
Working Paper No. 14

GEOGRAPHY, INDUSTRIAL ORGANIZATION, AND AGGLOMERATION

Stuart S. Rosenthal and William C. Strange

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
426 Eggers Hall
Syracuse, New York 13244-1020
(315) 443-3114 | Fax (315) 443-1081
e-mail: ctrpol@syr.edu

October 1999

$5.00

CENTER FOR POLICY RESEARCH - FALL 1999

Timothy M. Smeeding, Director
Professor of Economics & Public Administration

Associate Directors

Margaret M. Austin
Associate Director,
Budget and Administration

Douglas Holtz-Eakin
Chair, Professor of Economics
Associate Director, Center for Policy Research

Douglas Wolf
Professor of Public Administration
Associate Director, Aging Studies Program

John Yinger
Professor of Economics and Public Administration
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Dan Black .... Economics
Stacy Dickert-Conlin .... Economics
William Duncombe .... Public Administration
Thomas Dunn .... Economics
Gary Engelhardt .... Economics
Deborah Freund .... Public Administration
Vernon Greene .... Public Administration
Leah Gutierrez .... Public Administration
Madonna Harrington Meyer .... Sociology
Christine Himes .... Sociology
Jacqueline Johnson .... Sociology
Bernard Jump .... Public Administration
Duke Kao .... Economics
Thomas Kniesner .... Economics
Jeff Kubik .... Economics
Jerry Miner .... Economics
John Moran .... Economics
Jan Ondrich .... Economics
John Palmer .... Public Administration
Grant Reeher .... Political Science
Stuart Rosenthal .... Economics
Jodi Sandfort .... Public Administration
Michael Wasylenko .... Economics
Douglas Wolf .... Public Administration
Assata Zerai .... Sociology

GRADUATE ASSOCIATES

Matthew Andrews .... Public Administration
Yvonne Arsenault .... Public Administration
Hakan Aykan .... Public Administration
Reagan Baughman .... Economics
Robert Bifulco .... Public Administration
Christine Caffrey .... Sociology
Meghan Collins .... Public Administration
Julie Dombrowski .... Sociology
Deb Dougherty .... Sociology
Karla English .... Public Administration
Amy Ferraro .... Public Administration
Seth Giertz .... Economics
Andrzej Grodner .... Economics
Tess Heintze .... Public Administration
Pam Herd .... Sociology
Peter Howe .... Economics
Alyssa Hundrup .... Public Administration
Kwangho Jung .... Public Administration
Young Sun Kwon .... Economics
James Laditka .... Public Administration
Xin Li .... Economics
Donald Marples .... Economics
Neddy Matshalaga .... Sociology
Adriana Sandu .... Public Administration
Shalini Sharma .... Economics
Mehmet Serkan Tosun .... Economics
James Williamson .... Economics

STAFF

Martha Bonney .... Publications and Events Coordinator
Karen Cimilluca .... Librarian/Office Coordinator
Kati Foley .... Administrative Assistant, LIS
Esther Gray .... Administrative Secretary
Anne Metzger-Wormuth .... Receptionist
Denise Paul .... Editorial Assistant, *NTJ*
Annie Pennella .... Editorial Assistant, *Gerontologist*
Mary Santy .... Secretary
Debbie Tafel .... Secretary to the Director
Ann Wicks .... Administrative Secretary
Lobrenzo Wingo .... Computer Consultant

Abstract

This paper makes two contributions to the empirical literature on agglomeration economies. First, the paper uses a unique and rich database in conjunction with mapping software to measure the geographic extent and nature of agglomerative externalities. Previous papers have been forced to assume that agglomeration economies are club goods that operate at a metropolitan scale. Second, the paper tests for the existence of organizational agglomeration economies of the kind studied qualitatively by Saxenian (1994). This is a potentially important source of increasing returns that previous empirical work has not considered. Results indicate that localization economies attenuate rapidly and that industrial organization affects the benefits of agglomeration.

Introduction

The costs of cities can be seen in the skyscrapers, highways, and aqueducts that must be built to concentrate people in a small area. The benefits of cities—known as agglomeration economies—are less concrete but are just as real. Marshall (1920) provides the first careful economic analysis of agglomeration economies, arguing that cities enhance productivity by allowing for labor market pooling, input sharing, and technological spillovers. An extensive empirical literature has considered agglomeration, including Sveikauskas (1975), Moomaw (1981), Henderson (1986), Nakamura (1985), Carlton (1983), Glaeser, Kallal, Scheinkman, and Shleifer (1992), Henderson (1994), and Ciccone and Hall (1996) to name just a few.[1] These papers focus on whether the advantages of cities depend on city size or employment in a particular industry, whether agglomerative externalities are static or dynamic, and on the importance of urban diversity.

This paper addresses two important unanswered questions about agglomeration. First, what is the geographic scope of agglomerative externalities? In contrast to explicitly geographic theoretical work,[2] empirical work on agglomeration has been almost innocent of geography, and instead has implicitly modeled the city as a club. The economy is divided into geographic units, typically states, cities (more precisely, metropolitan statistical areas—MSAs), or counties. Economic activity is then divided spatially according to the geographic partition, and the effects of the local economic environment on productivity are measured. This approach has the advantage of allowing the use of readily available aggregate data. However, it is somewhat unsatisfying, since the benefits firms get from each other through labor market pooling, shared inputs, and technological spillovers are all likely to attenuate with distance. An important gap in our understanding of agglomeration economies, therefore, is that we do not know the geographic extent of agglomerative spillovers.

The second question that the paper addresses is how the organization of economic activity within a city affects the value of agglomeration. There is reason to believe that the productivity of a local economic environment does not depend just on the quantity of available inputs, but also on the way that such inputs are organized. In Saxenian's (1994) study of the computer industry, she points out that in the mid-1970s, both Boston (especially around route 128) and the San Jose to Palo Alto corridor (Silicon Valley) were essentially equal in their positions as centers of electronics and high-technology. The next decade witnessed a movement offshore of semiconductor production, which hurt the Silicon Valley, and a shift away from minicomputers, which hurt route 128. The Silicon Valley made a transition to software and other computer related industries that has been successful enough to earn it the label of the most productive economy on the planet. Route 128 did not make the transition as successfully.

There are two explanations for this divergence. One is that either location could have become dominant in software based on its locational characteristics, but that the random hand of history selected the Silicon Valley as the industry core. The other explanation is that the locations did not have identical locational characteristics, and that the Silicon Valley offered a more productive environment. On the one hand, both locations had many of the characteristics that could be expected to attract high-technology employment including educated workforces and proximity to research universities. However, Saxenian argues that the key difference between the Silicon Valley and Route 128 is in their industrial systems. In her view (Saxenian 1994, p. 7), a local industrial system has "three dimensions: local institutions and culture, industrial structure and corporate organization." Route 128 is presented by Saxenian as being relatively rigid and hierarchical, while the Silicon Valley is presented as being flexible and entrepreneurial. This certainly seems to be the view of the industry. Saxenian quotes Jeffrey Kalb, an entrepreneurial refugee from the Digital Electronics corporation:

> There's a fundamental difference in the nature of the industry between Route 128 and [the Silicon Valley]. Route 128 is organized into large companies that do their own thing...It's very difficult for a small company to survive in that environment...The Valley is very fast-moving and start-ups have to move fast. The whole culture of the Valley is one of change. We laugh about how often people change jobs. The joke is that you can change jobs and not change parking lots. There's a culture associated with that which says that moving is okay, that rapid change is the norm, that it's not considered negative on your resume...So you have this culture of rapid decisions, rapid changes, which is exactly the environment that you find yourself in as a startup.

Saxenian's analysis complements the work of Jacobs (1969) and Chinitz (1962), both of whom also suggest that urban efficiencies depend not just on numbers (i.e., city or industry size) but also on the nature of urban interactions. In the empirical literature, this issue has been considered obliquely in Glaeser et al (1992) and Henderson, Kuncoro, and Turner (1995) by including variables such as the number of employees per firm and the degree of urban specialization. However, these variables do not really capture the degree to which a location is blessed with a creative, entrepreneurial environment rather than an inflexible, hierarchical one. A further gap in the literature, therefore, concerns the impact of industrial organization on the value of agglomeration.

This paper addresses the geographic and organizational nature of agglomeration using Dun and Bradstreet Marketplace data. The data include information on over ten million firms in the United States reporting, among other things, location at the zipcode level, corporate status (subsidiary or non-subsidiary), age of establishment, employment, and sales. Using these data we are able to study agglomeration economies by considering the location and employment decisions of new firms as a function of the economic environment when the decisions were made. Because our estimation is carried out at the zipcode level, we are able to employ metropolitan fixed effects. This has a number of advantages, which are made clear below.

The paper's most important finding is that agglomeration economies attenuate with distance. The initial attenuation is rapid, with the effect of own-industry employment two to five miles away roughly one-fourth to one-half of the effect within one mile. Beyond five miles however, attenuation is moderate. This result is consistent with both theoretical models of the internal structure of cities and stylized facts: moving away from a city center, land and house rents, building heights, and population density all decline rapidly at first and slowly thereafter. These findings suggest that agglomeration should ideally be studied at a much more refined geographic level than has been the norm.

The paper also establishes that industrial organization affects the benefits of agglomeration. For a given level of employment, small firms generate a larger agglomerative effect than do large firms, while subsidiary firms have a larger effect than do non-subsidiary firms. The firm size result is broadly consistent with arguments by Saxenian that a more competitive and entrepreneurial environment enhances growth. The subsidiary firm result, in contrast, may indicate that subsidiaries act as a window on absentee parent firms, enabling subsidiaries to channel valuable information from the parent firm to the local economy.

The paper is organized as follows. The next section presents a simple empirical model of births and of new firm employment. This is followed by a discussion of our data and a presentation of summary measures. The results are then presented followed by the conclusions.

Model

Estimating Births and New Firm Employment

If agglomeration economies are present, then new firms will be drawn to concentrations of employment. Consequently, our approach to estimating agglomeration economies is to focus on these new firms. In effect, we ask in which zipcode and at what scale will entrepreneurs choose to open new firms?

Normalizing the price of output to one, a firm's profit will be $\pi(y) = a(y)\ f(x) - c(x)$, where $a(y)$ shifts the production function $f(x)$, y is a vector of locational characteristics, the components of which will be clarified below, and x is a vector of factor inputs that cost $c(x)$. A firm will choose input quantities according to the usual first order conditions. Employment (n), for example, is chosen such that $a(y)\partial f(x)/\partial n - \partial c(x)/\partial n = 0$.

Each entrepreneur assesses whether a firm would earn nonnegative profits if all inputs were chosen at their profit-maximizing levels. If so, then the firm is born. Entrepreneurs are heterogeneous in the profitability of their ideas. We express such heterogeneity by rewriting the firm's profit function as $\pi(y,\varepsilon) = \max_x a(y)\ f(x)(1+\varepsilon) - c(x)$. We suppose that ε is independent and identically distributed across entrepreneurs according to the cumulative distribution function $\Phi(\varepsilon)$. For any y, there is a critical level $\varepsilon^*(y)$ such that $\pi(y,\varepsilon^*(y)) = 0$ and $\pi(y,\varepsilon) > (<)\ 0$ as $\varepsilon > (<) \varepsilon^*(y)$. In this case, the probability that a given entrepreneur creates a firm is $\Phi(\varepsilon^*(y))$.

We assume that new firms choose their locations from among all of the zipcodes in the United States, $j = 1, \ldots, J$. New firms make their location and employment decisions at time t-1 taking the existing economic environment as given and are born one period later at time t. We suppose that the locational characteristics of each zipcode, y_j, are partitioned into two parts, y_{zj}

and y_{mj}. The elements of y_m vary by zipcode while the elements of y_m vary by metropolitan area. Aggregating over entrepreneurs gives the number of births (B) and total new firm employment (N) in zipcode *j*, which we express as linear functions of y_z and y_m:

$$B_{j,t} = b_z y_{z,j,t-1} + b_m y_{m,j,t-1} + \varepsilon_{b,t}, \tag{1}$$

$$N_{j,t} = n_z y_{z,j,t-1} + n_m y_{m,j,t-1} + \varepsilon_{n,t}, \tag{2}$$

where ε_b and ε_n are error terms.

Any locational characteristic that increases productivity will result in both more births and in more employment by the new firms. As such, key elements of $y_{z,j}$ include the spatial distribution of employment oriented around zipcode *j* (*j* = 1, … *J*). For example, the level of employment within and outside the firm's industry within one mile, two miles, etc. of the zipcode. These variables define the level of agglomeration associated with a given zipcode and can be measured with our data.

In contrast, some of the most important elements of y_m are difficult to measure given the wide range of city-specific variables that affect productivity. Elements of y_m, for example, include local fiscal policies, climate, quality of the workforce, and wage rates for different classes of labor, to name a few.[3] Note, however, that $b_m y_{m,j}$ and $n_m y_{m,j}$ from (1) and (2) are city-specific effects. Accordingly, rewriting equations (1) and (2) we obtain

$$B_{j,t} = b_z y_{z,j,t-1} + \gamma_{m,b} + \varepsilon_{b,t}, \tag{3}$$

$$N_{j,t} = n_z y_{z,j,t-1} + \gamma_{m,n} + \varepsilon_{n,t}, \tag{4}$$

where $\gamma_{m,b}$ and $\gamma_{m,n}$ (equal to $b_m y_{m,j}$ and $n_m y_{m,j}$) control for all attributes common to a metropolitan area that affect productivity. Our data permit us to estimate (3) and (4) and in so doing, we avoid a host of potentially important omitted variables.

Other Approaches to Measuring the Benefits of Agglomeration

Our approach to measuring the benefits of agglomeration is to look at the decisions of new firms. There are, of course, other approaches. Specifically, the benefits of agglomeration have been measured using value added as a measure of economic productivity (e.g., Ciccone and Hall 1996) and the growth of total employment in an industry (e.g., Glaeser et al 1992; Henderson et al 1995).

Studying value added requires data on the market value of both output and input quantities. Although value of output and labor quantities are feasible to obtain, capital stock measures are generally quite difficult to come by at the micro level, making the value added approach difficult to implement. Instead, studying growth of total employment has been much more common, but it presents different challenges. Data on total employment are often readily available and the analysis lends itself to linear regressions. However, existing firms are constrained by prior choices, most importantly the level and kind of capital of previously installed. Those fixed factors affect how the firm values the marginal worker, and consequently how it changes its employment level in response to a change in its environment. In principle, this difficulty can be overcome by looking at changes in total employment over a sufficiently long time frame so that there are no fixed factors and all firms are effectively new. Even then, however, one still has to address a difficult endogeneity problem: not only is the growth of total employment in a given area sensitive to the composition of employment in the area (an agglomeration effect), but the reverse may hold as well. Implementing this approach, therefore, ideally requires a long panel and effective instruments to control for endogenous variables.

Focusing on the birth of new firms and their employment avoids the problems most often associated with the two approaches above. Data on capital inputs is not required, new firms are unconstrained by previous decisions, and new firms make their location and employment decisions taking the existing economic environment as exogenously given, at least as a first approximation. The principal drawback of focusing on firm births and their employment is that many locations do not receive any new births in a given period which can lead to technical challenges on the econometric side.[4]

As will become apparent, our data are especially well suited to studying the arrival of new firms and their employment decisions. In addition, robustness checks suggest that the principal results regarding new firms and their employment are extremely robust to a wide range of econometric specifications. For that reason, we focus on firm births and their employment as our measure of the benefits of agglomeration. In effect, we ask: in which location and at what scale will entrepreneurs choose to open new firms?[5]

Data and Variables

The Database

Data for the analysis were drawn from the Dun & Bradstreet Marketplace database which provides a wealth of information on over ten million firms. Details of the data are provided in Appendix A. Data from the fourth quarter of 1997 are used to construct two alternative dependent variables, new firms and their employment, where new firms are those that are listed in the data as being one year or less in age as of 1997:4. Data from the fourth quarter of 1996 are used to measure the "existing" level of employment upon which new firms are assumed to have based their location decisions.

The Variables

For each industry we calculate the existing level of employment both within and outside of the industry in question. To measure the geographic extent of agglomerative externalities we create a unique set of concentric ring employment variables for both types of employment.

These variables are calculated as follows. First, all employment in a given zipcode is treated as being located at the zipcode's geographic centroid. Then, using mapping software, circles of radius r_i, i = 1, …, 15, are drawn around the centroid of each zipcode in the United States. The level of own-industry employment contained within each circle is then calculated by adding up the own-industry employment associated with the zipcode centroids within each circle. The same procedure is used to calculate the level of other-industry employment within each circle. Differencing employment levels for adjacent circles (by employment type) yields estimates of the levels of own- and other-industry employment within a given concentric ring.[6] Thus, the 2-mile ring (r_2) reflects employment between the 1 and 2-mile circles, and so on out to 15 miles.

The other variables in our model are calculated in obvious ways. The area of each zipcode is included (in square miles) to control for differences in zipcode size. Number of firms per worker are calculated as in Glaeser et al (1992) to proxy for local competitiveness. This variable is calculated separately for own-industry employment and employment outside of the own industry. The diversity of economic activity is incorporated using a Herfindahl index of employment by 2-digit SIC industries as in Henderson (1995).[7] Both the competitiveness variables and the Herfindahl index are calculated at the zipcode level.

Finally, any other arguments of the firm's cost function that vary regionally such as wage rates and the quality of the local labor force are also pertinent. As discussed previously, these variables are controlled for using metropolitan area fixed effects. In addition, we allow for different fixed effects for non-metropolitan locations for each of the 50 states. In total, this yields 373 fixed effects in the model.

The Industries

Three criteria were used in selecting industries to study. First, we selected industries whose output is consumed nationally or internationally. Second, we selected industries with substantial numbers of new firms and consequently substantial new firm employment. Third, we selected industries that are important enough to have been the focus of other studies. Specifically, we estimate the determinants of new firm employment and births for six industries: software (SIC 7371, 7372, 7373, and 7375)[8], food processing (SIC 20), apparel (SIC 23), printing and publishing (SIC 27), fabricated metals (SIC 34), and industrial and commercial machinery (SIC 35). All of these industries meet the first two criteria. In addition, software has been studied by Saxenian (1994), while the two-digit manufacturing industries were considered by Nakamura (1985) and Henderson (1986).

The industries are a mix of traditional industries with established products and innovative industries where new products are important. Innovation in the software industry is widely known given the explosion of computer technologies and use. Apparel and to a lesser degree food processing involve fashion and are, therefore, almost by definition also innovative. Additionally, the six industries studied here are a mix of heavy and light industries, with machinery and metals both in the former category.

Summary Statistics

Table 1 provides selected summary statistics for the sample. Apart from the specific details of each industry, it is important to emphasize three points when viewing Table 1. First, there are 39,060 zipcodes and 373 identified metropolitan areas and rural zones. Because of the large number of zipcodes, many of our estimates are quite precise. Because of the large number of metropolitan/rural fixed effects, it is hoped that all regional attributes that affect productivity are controlled for.

Second, for each industry, the standard deviation of existing own-industry employment is at least ten times smaller than the standard deviation of existing employment outside of the industry in question. This is the first evidence in this paper that individual industries are more spatially concentrated than total employment.

A final important point is that of censoring. Although there are births in many zipcodes for each of the six industries (the uncensored observations), there are zero births in the majority of the zipcodes in each case (the censored observations). The large number of zeros requires nonlinear estimation. Because we have a large sample, and because births and new firm employment range into the thousands in some zipcodes, we estimate equations (3) and (4) by Tobit for each of the industries.[9] This raises a technical issue because imprecise estimates of the fixed effects in nonlinear models typically lead to inconsistent estimates of the slope coefficients [e.g., Chamberlain 1980, 1985; Hsiao 1986]. We have two principal responses to this issue.

First, bias resulting from noisy estimates of fixed effects in nonlinear models goes to zero as the number of observations per fixed effect becomes arbitrarily large. Given that our sample has over 100 zipcodes per fixed effect, inconsistency arising from noisy estimates of the fixed effects is hoped to be small. Second, in linear fixed effect models noisy estimates of the fixed effects do not bias estimates of the slope coefficients. Accordingly, as a robustness check, Appendix B presents results for each industry based on a linear (ordinary least squares) fixed effect specification in which all zipcodes with zero births are omitted. That approach suffers of course from a potential sample selection problem since most of the zipcodes are thrown out of the analysis. Nevertheless, the qualitative nature of the results in Appendix B are quite similar to results from the Tobit models. This suggests that the results in this paper are robust to any issues related to econometric specification.[10]

Results

Initial Results

Tables 2a and 2b present estimates for the Tobit fixed effect models using respectively the number of new establishments and new firm employment as dependent variables. A result that should be emphasized is that our fixed effects model—which identifies agglomeration effects based on *within* city variation in the data—yields results that are broadly consistent with previous work that was based on *between* city variation in the data. Specifically, Tables 2a and 2b show that own-industry competition encourages births and new firm employment in every industry but one. In addition, the tables show that a decrease in the diversity of employment—as measured by an increase in the Herfindahl index—decreases births and leads to less new firm employment. These two results are consistent with Glaeser et al. (1992).[11] The numerical values of our coefficients on the Herfindahl index are considerably larger than those previously reported by Glaeser et al. (1992) and Henderson et al. (1995). Those studies, however, used growth of total employment as their measure of the benefits of agglomeration. In Rosenthal and Strange (1999), we reformulate our model with growth in the total number of firms and in total employment within specific industries as the dependent variables. That modification produced substantially smaller coefficients on the Herfindahl indexes, suggesting that diversity may have more impact on firm births than on net growth.

Finally, for all six industries and for both births and new firm employment, localization effects (own-industry employment) are more important than urbanization effects (other-industry employment). Specifically, for any given concentric ring of employment (e.g., the one mile ring), the coefficient on the localization employment variable is typically at least two orders of magnitude larger than the coefficient on the corresponding urbanization employment variable. In addition, most of the localization coefficients are highly significant while most of the urbanization coefficients are not significant. The result that localization economies are more important than urbanization economies is consistent with Henderson's (1986) findings for Brazil and the United States and those of Nakamura (1985) for Japan.[12]

The Geography of Agglomeration Economies

We turn now to our most important results. In principle, one could use an arbitrarily large number of concentric rings when assessing how quickly agglomeration economies attenuate. In practice, however, it is necessary to aggregate the geographic detail in order to maintain a parsimonious specification. After some experimentation, the spatial distribution of employment was aggregated into four concentric rings: employment within one mile of the zipcode centroid, between one mile to five miles, between five and ten miles, and between ten to fifteen miles. Condensing the geographic effects to these four variables greatly facilitated both estimation and presentation without changing the qualitative nature of the geographic patterns.[13]

The key geographic result in Tables 2a and 2b is that localization economies attenuate rapidly in the first few miles but slowly thereafter. This result holds for each of the six industries studied regardless of whether the benefits of agglomeration are measured by firm births or their employment. This is most easily seen at the bottom of the table where we present the percentage change per mile (PCPM) in the localization coefficients. This is measured by the difference between the Mile1 coefficient and the coefficient on a given concentric ring, divided by the distance from one mile to the midpoint of the alternate ring.[14]

Focusing on births, the PCPM from mile 1 to 3 ranges from 25 percent for software to 47 percent for apparel. In contrast, the PCPM from mile 1 to 7 ranges from 10 to 14 percent across industries while the PCPM from mile 1 to 12 ranges from 6 to 7.5 percent across industries. To interpret this pattern, note that if localization economies attenuated at a constant rate, then the

PCPM would be constant regardless of distance. Instead, the nonlinear pattern obtained here indicates that the benefits of localization attenuate rapidly in the first few miles but slowly thereafter. This pattern is consistent with theoretical models of city structure in addition to well-known characteristics of urban areas. For example, building heights and population density both decline rapidly at first when moving away from the center of economic activity but decline slowly in more distant suburbs.

In contrast to the clear geographic pattern for the localization coefficients, the geographic pattern for the urbanization coefficients is somewhat obscure. That difference serves to highlight an important distinction between urbanization and localization economies. Localization effects are expected to be positive and to diminish monotonically with distance between firms as information spillovers and the ability to share both intermediate inputs and labor diminish. Our results support this. On the other hand, urbanization effects are not necessarily expected to be positive. They reflect the tradeoff between the benefits of locating near densely developed areas and congestion costs. To the extent that industries differ in the net benefits they derive from proximity to employment centers, some industries will prefer more densely developed areas while others will prefer more outlying locations, *ceteris paribus*. As a result, the geographic pattern of urbanization effects is expected to differ across industries and can in general be quite varied. Our results support that argument as well.[15]

Industrial Organization and Agglomeration

As noted earlier, Saxenian (1994) defines the local industrial system as having three dimensions: culture and institutions, corporate organization, and industrial structure. Absent hard data on culture and institutions, we focus on the latter two aspects of the industrial system. We will address two questions. First, does an industrial structure dominated by small firms provide a more productive environment than one dominated by large firms? Second, does a corporate organization based on parent-subsidiary links as opposed to one dominated by independent firms affect productivity?

If small firms were more open and innovative as might be inferred from Saxenian's (1994) comparison of the Silicon Valley and Route 128, then an additional worker at a small firm would enhance the productivity of neighboring firms more than an additional worker at a medium or large firm. To test that idea, we reestimated the models in Tables 2a and 2b with the localization variables divided into three types: employment at small firms (fewer than 25 employees), employment at medium firms (25 to 99 employees), and employment at large firms (100 or more employees). In addition, since localization effects attenuate rapidly, we aggregated own-industry employment from the zipcode centroid out to five miles and omitted more distant rings of own-industry employment from the regression. That simplification enables us to avoid a proliferation of localization variables and serves to highlight the effect of firm size.[16] All other regressors in Table 2 including the fixed effects were retained in the model.

Table 3a reports results for the localization variables with all other coefficients suppressed to conserve space.[17] A clear pattern emerges. For five of the six industries, employment at small firms has a larger effect on births or new-firm employment than does employment at medium firms, and medium firm employment has a larger effect than does employment at large firms. In software, for example, an additional worker at a small firm has roughly twice the effect on births and three times the effect on new firm employment than does an additional worker at a medium size firm. Medium firms have over four times as large an effect on births as do large firms, and over two times as large an effect on new firm employment.

These results have important implications for the study of localization. They suggest that efficiency arises not simply from the concentration of own-industry employment but also from the concentration of the right kind of own-industry employment. This result is consistent with

Saxenian's comparative systems analysis of the Silicon Valley and Route 128 and suggests that for software and other industries, small firms make better neighbors.[18]

In an analogous manner, because a subsidiary firm is constrained by the rest of its corporation, it might be less flexible or innovative. Because a subsidiary firm may purchase its inputs or sell its outputs within the corporation, it may not be intimately involved with its neighbors. For both of these reasons, employment at a subsidiary could have a smaller effect on the productivity of nearby firms than employment at a nonsubsidiary firm.[19] On the other hand, access to a subsidiary might provide access to resources elsewhere in the subsidiary's parent corporation, including resources at other plants in other locations. In addition, subsidiaries may generate agglomeration economies through spin-offs. Jacobs (1969, p. 66), for example, notes that breakaways from Hughes Aircraft were important sources of entrepreneurship in the Los Angeles electronics industry after World War II.[20] For both of these reasons, employment at a subsidiary could have a larger agglomerative effect than employment at a nonsubsidiary firm.

To evaluate the effect of subsidiary status on localization economies, the localization variables in Table 3a were replaced with two new localization variables: own-industry employment at subsidiaries of corporate parents and own-industry employment at nonsubsidiaries. All other features of the models estimated in Table 3a were retained. Results for this new specification are presented in Table 3b where only the coefficients on the localization variables are presented to conserve space as before.

Once again, a pattern emerges. Nonsubsidiary employment has a substantially smaller effect on both births and new firm employment in each of the industries except printing and publishing where the differences are small.[21] The largest difference is in software where the effect of an additional worker at a nonsubsidiary firm is roughly two-thirds that of a subsidiary firm in the births model. In the new firm employment model, the effect of an additional software worker at a subsidiary software firm is over ten times as large. This pattern is repeated for the other industries, except printing.

These results suggest that a corporate organization dominated by subsidiary firms is more conducive to growth than one dominated by nonsubsidiary firms. One possible interpretation of these findings is that adverse effects from increased hierarchy at subsidiary firms are more than offset by other positive effects of the sort discussed above.

Conclusion

This paper makes two important contributions to the empirical literature on agglomeration economies. First, we use a unique and rich database in conjunction with mapping software to measure the geographic extent and nature of agglomerative externalities. Previous papers have been forced to assume that agglomeration economies are club goods that operate at a metropolitan scale. Second, we test for the existence of organizational agglomeration economies of the kind studied qualitatively by Saxenian (1994). This is a potentially important source of increasing returns that previous empirical work has not considered.

Results from six industries provide compelling evidence that localization economies—agglomeration economies arising from spatial concentrations of firms within a given industry—attenuate rapidly over the first few miles and then attenuate much more slowly thereafter. While it is beyond the scope of this study to determine exactly which sources of agglomeration economies are responsible for this pattern, it is tempting to speculate. As discussed in the Introduction, three potential sources are information spillovers, labor market pooling, and shared inputs. Information spillovers that require frequent contact between workers may dissipate over a short distance as walking to a meeting place becomes difficult. On the other hand, the benefits of labor market pooling and shared inputs might extend over a much greater distance since those benefits rely more on the ability of agents to conveniently drive from one location to another. Initial rapid attenuation of information spillovers followed by a more gradual attenuation of benefits from labor market pooling and shared inputs would produce an attenuation pattern consistent with that found in this paper. Systematic empirical support for that argument, however, is left for future research.

Our results also indicate that industrial structure and corporate organization affect the benefits that arise from clustering firms together within a given industry. We find that own-industry employment at small firms has a greater effect on productivity than does a comparable level of own-industry employment at larger firms. This result lends support to recent arguments that a more entrepreneurial industrial system promotes growth. We also find that own-industry employment at subsidiary firms has a larger effect than does a comparable level of own-industry employment at nonsubsidiary firms. This result suggests that the effects of corporate organization are complicated. If hierarchy is bad for growth, as has been argued, then it is outweighed by some other benefit of having a subsidiary as a neighbor. For example, subsidiary establishments may be more likely to form spinoffs or they may channel valuable information from the parent firm to the local economy.

Finally, we should emphasize that the results above are robust to the estimation method and to the two measures of the benefits of agglomeration used in the study—births of new firms and their employment. On the whole, these findings suggest that future studies of agglomeration economies should be sensitive both to the industrial organization of the firms and especially the micro geography of agglomeration.

Appendix A: Data Description

Our principal data source is Dun and Bradstreet's Marketplace file. The data include information on over ten million firms in the United States reporting, among other things, location at the zipcode level (much smaller than a county), corporate status (subsidiary or non-subsidiary), age of establishment, employment, and sales. The "complete" D&B database includes firm specific information on over ten million firms in the United States and is based on publicly available sources and D&B phone surveys. That data set, however, was prohibitively expensive. Instead, we obtained a more limited but still enormously rich version of the database in which all of the firm specific data was aggregated up to the zipcode level.[22]

In phone conversations with analysts at D&B, we were advised that firms requesting not to be in the database are omitted from the data file. Partly for that reason, the D&B database, while immense, does not contain the entire universe of firms in the United States. Nevertheless, the D&B analysts felt that the omissions from the data set are sufficiently random that the D&B database is representative of the spatial distribution of firms in the United States.[23] Moreover, measurement error associated with the distribution of employment across industries within a given geographic zone is likely to be small if one aggregates up by even a modest amount, as to the zipcode level.[24] Accordingly, we assume that the data set provides an accurate measure of the spatial distribution of firms at the zipcode level.[25]

Appendix B: Supplemental Tables

This appendix presents two sets of supplemental tables. The first set, Tables B-1a and B-1b, reproduce Tables 2a and 2b in the text using a different estimation method: ordinary least squares including all of the locational fixed effects but dropping those zipcodes with zero new firms. As noted in the text, results are qualitatively similar to those based on the Tobit specification. This suggests that the key findings in this study are robust to econometric specification.

The second set of supplemental tables, Table B-2, presents estimates of the effects of firm size where all of the localization variables are based on employment at own-industry firms that are five years old or more. This table is included in order to examine whether correlation between firm size and age might explain the substantial effect of small firm employment that can be seen in Table 3a in the text. As is apparent from Table B-2, the results in the appendix are qualitatively the same as in Table 3a—the model based on own-industry employment at firms of all ages. Thus, it is unlikely that the small firm effect arises simply because small firms are also young.

Endnotes

*We gratefully acknowledge the financial support of the Social Sciences and Humanities Research Council of Canada, the UBC Centre for Real Estate and Urban Land Economics, the Real Estate Foundation of British Columbia, and the Center for Policy Research at Syracuse University. We thank Richard Arnott, Jan Brueckner, Iain Cockburn, Ed Coulson, Ed Glaeser, Keith Head, Vernon Henderson, Daniel McMillen, Jay Wilson, and an anonymous referee for helpful comments.

1. See Quigley (1998) for a more complete survey of this literature.

2. See O'Hara (1977), Ogawa and Fujita (1980), Imai (1982), Helsley (1990), or Krugman (1993), for example.

3. Another element of *ym* is the the "birth potential" of the city as discussed by Carlton (1983). Carlton argues that a concentration of employment in a city increases both the profitability of a new firm and the number of potential entrepreneurs who might create such a firm.

4. The principle technical issues will be clarified later in the paper.

5. In contrast, a model of employment growth that takes geography into account asks: where do net changes in employment occur? In Rosenthal and Strange (1999), we reformulate the model presented in this paper with net change over time in the total number of firms in an industry and total own-industry employment as the dependent variables. Looking across cities, we find that localization encourages growth within industries, a result that is broadly consistent with previous work on urban growth. Looking within cities, we find little evidence of a systematic change in the tendency of industries to agglomerate.

6. Various MapInfo software products were used to geocode the data and create the concentric ring variables.

7. The Herfindahl index of specialization (the inverse of diversity) is defined as $\sum_{i=1}^{90} S_i^2$, where s_i is industry *i*'s share of total employment and *i*=1,2,…,90 are the two-digit industries.

8. This definition restricts software to outputs that are sold nationally or internationally.

9 In the case of new firm employment, Tobit is especially appropriate: new firm zipcode level employment ranges from 0 to 7,392 for software (the widest range of the six industries) and from 0 to 1,200 for food products (the narrowest range). In contrast, zipcode level births range from 0 to 107 in the case of apparel (the widest range) and from 0 to 6 in the case of fabricated metals (the narrowest range). For some industries,

therefore, the range of zipcode level births might suggest a Poisson count model, while for other industries such as apparel and software Tobit still seems preferred. To simplify the analysis and presentation, we used Tobit throughout. This is an approximation, but it is one with which we are comfortable. In addition, the key results in the paper are extremely robust to alternative estimation methods and to the dependent variable used as will be clarified below.

10. Two other models were considered but rejected as solutions to the Tobit-fixed effect problem. First, Chamberlain (1980) developed a conditional logit solution to this problem in which the fixed effects are integrated out of the model allowing one to obtain consistent slope coefficients with finite T (where T is the number of observations per fixed effect). Unfortunately, the conditional logit approach is not computationally feasible for large samples such as ours in which roughly 39,000 zipcodes are spread over 373 fixed effects. As an alternative, one could omit the fixed effects and in their place include those city-specific variables that are thought to affect firm location and employment decisions directly in the model. In a sense, this is what previous agglomeration studies have done when making intermetropolitan comparisons of employment growth. That approach, however, could suffer from omitted variable bias and was not preferred for that reason.

11. The diversity result is somewhat at variance with Henderson et al.'s (1995) results for innovative industries. Although they find a positive effect of diversity when it is interacted with a dummy representing historical concentration, when diversity is included without the interaction the coefficient on diversity is negative in their paper.

12. In addition, in Tables 2a and 2b, note that printing and publishing exhibit the weakest localization economies. This is consistent with Ellison and Glaeser (1997), who found that printing and publishing exhibited little geographic localization in the United States.

13. In earlier versions we included mile-by-mile geographic variables for each employment type out to 15 miles. Results from that estimation were qualitatively similar to those in Tables 2a and 2b but were much more difficult to produce and present given the very large number of coefficients.

14. Specifically, the average percentage change per mile is calculated as (Mile1-MileD)/d*Mile1, for MileD equal to Mile2-5, Mile6-10, and Mile11-15, and d equal to 2, 7, and 12, respectively.

15. For example, the monotonic increase in urbanization coefficients for software births in Table 2a may suggest that the software industry benefits from locating in outlying areas, *ceteris paribus*. In contrast, the relatively large, positive urbanization coefficient for Printing and Publishing in the 5 to 10 mile ring may suggest that that industry does best locating 5 to 10 miles from densely developed areas, *ceteris paribus*.

16. We have also carried out parallel estimates of these models in which the localization variables reflect employment out to ten and out to fifteen miles. Results from these regressions were similar to those presented in Table 3.

17. Coefficients on the other variables were similar to those in Tables 2a and 2b.

18. If young firms tend to be small and if there are unmeasured zipcode attributes that attract new firms, then our small firm effect could reflect the influence of those unmeasured attributes. As a robustness check, we re-estimated the model in Table 3a using employment at firms at least five years old for all of the localization (own-industry) variables in the model. Results from that regression are presented in Appendix B (Table B-2) and are quite similar to those in Table 3. This suggests that the estimates in Table 3 are driven by firm size.

19. See Saxenian (1994) for a concrete discussion of the open industrial system of the Silicon Valley and its advantages relative to the closed industrial system of the Route 128 area.

20. Jacobs (p. 153) further notes that sometimes the new firm produced a product that was far removed from aircraft manufacture, such as sliding doors. In addition, Saxenian (1994, p. 52) also discusses the important role of spin-offs from Fairchild Semiconductor in the Silicon Valley.

21. There is also little difference between subsidiary and non-subsidiary effects for food products in the birth model, although the subsidiary effect in the food products employment model is twice that of the non-subsidiary effect.

22. The complete data set containing individual firm specific information costs over $600,000 for one quarter. The data aggregated to the zipcode level were available at a far lower price.

23. Additional details on the Dunn and Bradstreet (D&B) MarketPlace file are provided at the Dunn and Bradstreet web site, www.dnb.com. As described by Dunn and Bradstreet, there are several important benefits to firms from listing themselves in the D&B database and obtaining a D-U-N-S identification number. These benefits arise primarily because of the incredible size of the D&B data file. Because the D&B file is such an effective source of information on firms throughout the economy, businesses use the D&B file to do market analysis and search out potential trading partners. Individual firms therefore have an incentive to list themselves with D&B in much the way firms have an incentive to voluntarily list themselves in the yellow pages. In addition, DUNS identification numbers are rapidly becoming a standard identification device in the economy, and many companies including the Federal Government require that clients obtain a D-U-N-S number as a precondition for engaging in trade. As noted in the D&B website, "It [the D-U-N-S number] is now the standard for all United States Federal Government electronic commerce transactions to help streamline and reduce federal procurement costs."

24. Because some firms are omitted from the data set, our regressors—which reflect various measures of the existing level of employment—could suffer from an errors in variables problem which would bias the estimated coefficients towards zero. Assuming, however, that the spatial distribution of the data set is representative of the United States, then aggregating up to the zipcode level likely averages away any errors in the data, at least as regards the relative magnitude of employment in one industry versus another. Primarily for that reason, but also because many of our estimated agglomeration effects are of substantial magnitude (and significance), any errors in variables problem is likely to be slight.

25. Carlton (1983) is the only other study we are aware of that employs Dun and Bradstreet data. Carlton too concludes that the D&B data are reasonably representative.

TABLE 1
SELECTED SUMMARY STATISTICS
All Values Calculated At The Zipcode Level

	Variables Common To All Industries			
	Mean	**Std Dev**	**Min**	**Max**
Zipcode area (sq. miles)	126.95	486.14	0.00771	18,559
Herfindahl Index (2-digit)	0.23	0.24	0.0286	1

Software: SIC 7371-7373, 7375
otal Obs: 39,060 Uncensored Obs: 636

Own Industry	**Mean**	**Std Dev**	**No. 0's***	**Max**
Births	0.38	1.26	32,698	39
New firm wrkrs	2.85	54.15	32,698	7,392
Firm/workers	0.10	0.19	26,198	1
Workers	30.41	310.12	26,198	39,764
Non-sub. wrkrs	26.56	290.35	26,245	39,727
Subsidiary wrkrs	3.85	65.21	37,966	4,066
Small firm wrkrs	8.76	32.71	26,400	1,033
Med. firm wrkrs	5.84	39.23	36,818	1,405
Large firm wrkrs	15.81	284.78	38,087	39,720
Other Industry				
Firm/workers	0.17	0.13	5	1
Workers	2925.37	6114.91	5	155,906

Food Products: SIC 20
otal Obs: 39,060 Uncensored Obs: 132

Own Industry	**Mean**	**Std Dev**	**No. 0's***	**Max**
Births	0.04	0.23	37,736	10
New firm wrkrs	1.05	19.70	37,736	1,200
Firm/workers	0.06	0.15	26,355	1
Workers	44.41	240.59	26,355	14,261
Non-sub. wrkrs	38.33	211.49	26,520	14,225
Subsidiary wrkrs	6.09	78.05	37,850	5,120
Small firm wrkrs	3.90	11.06	28,114	376
Med. firm wrkrs	6.87	27.82	35,286	1,097
Large firm wrkrs	33.64	229.00	36,384	14,225
Other Industry				
Firm/workers	0.17	0.13	7	1
Workers	2911.36	6115.51	7	156,586

Apparel: SIC 23
otal Obs: 39,060 Uncensored Obs: 1,6

Own Industry	**Mean**	**Std Dev**	**No. 0's***	**Max**
Births	0.06	0.63	37,421	107
New firm wrkrs	0.62	21.13	37,421	3,300
Firm/workers	0.07	0.19	28,212	1
Workers	23.40	192.18	28,212	25,555
Non-sub. wrkrs	21.45	180.53	28,285	24,175
Subsidiary wrkrs	1.94	34.88	38,605	2,700
Small firm wrkrs	3.54	47.10	29,583	8,101
Med. firm wrkrs	5.18	66.62	36,774	10,413
Large firm wrkrs	14.68	112.91	37,426	7,041
Other Industry				
Firm/workers	0.17	0.13	3	1
Workers	2932.38	6137.83	3	156,748

Printing and Publishing: SIC 27
otal Obs: 39,060 Uncensored Obs: 5,00

Own Industry	**Mean**	**Std Dev**	**No. 0's***	**Max**
Births	0.20	0.64	34,054	13
New firm wrkrs	1.67	28.91	34,054	3,500
Firm/workers	0.11	0.19	20,805	1
Workers	49.51	264.73	20,805	14,981
Non-sub. wrkrs	41.85	212.74	20,871	12,697
Subsidiary wrkrs	7.66	105.32	37,242	8,447
Small firm wrkrs	13.46	40.08	21,233	2,519
Med. firm wrkrs	10.50	45.73	34,230	2,075
Large firm wrkrs	25.55	221.97	36,858	13,250
Other Industry				
Firm/workers	0.17	0.13	14	1
Workers	2906.26	6042.49	14	149,251

Fabricated Metal: SIC 34
otal Obs: 39,060 Uncensored Obs: 1,8

Own Industry	**Mean**	**Std Dev**	**No. 0's***	**Max**
Births	0.06	0.27	37,185	6
New firm wrkrs	1.52	26.94	37,185	3,500
Firm/workers	0.07	0.17	24,606	1
Workers	48.51	223.79	24,606	12,799
Non-sub. wrkrs	41.62	192.93	24,755	8,962
Subsidiary wrkrs	6.89	87.33	37,172	12,424
Small firm wrkrs	7.38	22.95	25,989	862
Med. firm wrkrs	12.93	49.19	33,596	1,980
Large firm wrkrs	28.20	193.75	36,306	12,775
Other Industry				
Firm/workers	0.17	0.13	4	1
Workers	2907.27	6110.13	4	157,999

Machinery: SIC 35
otal Obs: 39,060 Uncensored Obs: 3,28

Own Industry	**Mean**	**Std Dev**	**No. 0's***	**Max**
Births	0.12	0.48	35,780	16
New firm wrkrs	2.16	27.36	35,780	1,804
Firm/workers	0.09	0.19	20,869	1
Workers	76.17	422.90	20,869	26,401
Non-sub. wrkrs	64.84	386.26	20,990	26,401
Subsidiary wrkrs	11.32	106.46	36,471	8,125
Small firm wrkrs	12.44	38.15	21,762	1,293
Med. firm wrkrs	15.36	57.55	32,889	1,947
Large firm wrkrs	48.37	389.00	35,895	26,255
Other Industry				
Firm/workers	0.17	0.13	16	1
Workers	2879.61	6033.33	16	157,678

*No. 0's refers to the number of zipcodes for which the variable has a value of 0.

TABLE 2a
BIRTHS OF NEW FIRMS - GEOGRAPHIC EFFECTS
(Numbers in Parentheses are t-ratios)

	Software SIC 7371-73, 75	Food Products SIC 20	Apparel SIC 23	Printing & Publishing SIC 27	Fabricated Metal SIC 34	Machinery SIC 35
	Zipcode Area, Diversity, and Competition Effects					
Zipcode area in square miles	-4.776E-04 (-3.702)	-1.250E-05 (-0.150)	-4.519E-04 (-1.616)	-3.305E-04 (-3.328)	-8.140E-05 (-0.758)	-9.740E-05 (-1.042)
Zipcode Herfindahl Index	-1.407E+01 (-40.561)	-7.718E+00 (-16.972)	-1.159E+01 (-17.453)	-9.573E+00 (-36.798)	-7.139E+00 (-20.145)	-9.229E+00 (-28.733)
Zipcode firms per worker - other ind.	-1.174E+01 (-24.533)	-6.646E+00 (-11.729)	-7.817E+00 (-9.801)	-8.765E+00 (-24.046)	-7.582E+00 (-15.291)	-8.101E+00 (-19.303)
Zipcode firms per worker - own ind.	1.991E+00 (16.151)	1.067E+00 (6.635)	1.814E+00 (9.470)	8.379E-01 (8.190)	4.706E-01 (3.585)	1.171E-01 (0.962)
	Urbanization Effects: Other (Total - Own) Industry Employment From …					
0 to 1 Mile Ring	-1.240E-06 (-1.183)	4.490E-07 (0.440)	-5.160E-06 (-2.748)	-1.250E-06 (-1.370)	-7.730E-07 (-0.741)	-9.120E-07 (-0.809)
1 to 5 Mile Ring	-6.590E-07 (-2.507)	2.880E-07 (1.103)	9.070E-07 (1.911)	-1.340E-07 (-0.541)	-3.150E-07 (-1.215)	-8.190E-07 (-3.070)
5 to 10 Mile Ring	-1.940E-07 (-1.145)	3.200E-08 (0.170)	5.690E-07 (1.702)	5.740E-07 (3.419)	-1.890E-07 (-1.078)	-3.530E-07 (-2.153)
10 to 15 Mile Ring	-3.170E-08 (-0.212)	-1.240E-07 (-0.745)	-1.330E-07 (-0.432)	-4.860E-09 (-0.034)	-1.430E-07 (-0.874)	-6.040E-08 (-0.442)
	Localization Effects: Own Industry Employment From …					
0 to 1 Mile Ring	5.104E-04 (10.026)	3.425E-04 (6.204)	7.215E-04 (15.512)	1.202E-04 (4.391)	5.216E-04 (9.083)	2.914E-04 (11.084)
1 to 5 Mile Ring	2.576E-04 (10.146)	9.050E-05 (2.549)	4.690E-05 (2.631)	3.890E-05 (4.337)	9.520E-05 (2.773)	9.780E-05 (4.956)
5 to 10 Mile Ring	1.692E-04 (10.755)	2.020E-05 (0.711)	4.270E-05 (2.613)	1.140E-05 (1.363)	5.610E-06 (0.248)	5.060E-05 (4.820)
10 to 15 Mile Ring	1.368E-04 (8.438)	8.240E-05 (3.484)	5.560E-05 (2.773)	3.560E-05 (4.256)	1.118E-04 (5.830)	4.700E-05 (5.298)
	Average Percentage Change In Localization Effect Per Mile From ...*					
1 to 3 Miles	24.76%	36.79%	46.75%	33.82%	40.87%	33.22%
1 to 7 Miles	9.55%	13.44%	13.44%	12.93%	14.13%	11.81%
1 to 12 Miles	6.10%	6.33%	7.69%	5.87%	6.55%	6.99%
	Summary Measures					
Stnd Error	2.96	2.14	3.49	2.13	2.05	2.18
Log-L	-21,967.97	-5,916.72	-7,566.48	-17,124.28	-7,829.74	-12,518.12
Uncensored	6,362	1,324	1,639	5,006	1,875	3,280
Total Obs	39,060	39,060	39,060	39,060	39,060	39,060
Fixed Effects	373	373	373	373	373	373

*Calculated as (Mile1-MileD)/d*Mile1, for D set to 1-5, 5-10, and 10-15, and d equal to 2, 7, and 12, respectively.

TABLE 2b
NEW FIRM EMPLOYMENT - GEOGRAPHIC EFFECTS
(Numbers in Parentheses are t-ratios)

	Software SIC 7371-73, 75	Food Products SIC 20	Apparel SIC 23	Printing & Publishing SIC 27	Fabricated Metal SIC 34	Machinery SIC 35
	Zipcode Area, Diversity, and Competition Effects					
Zipcode area in square miles	-1.301E-02 (-2.321)	-1.912E-03 (-0.355)	-1.196E-02 (-1.355)	-9.362E-03 (-2.377)	-5.656E-03 (-0.734)	-3.524E-03 (-0.786)
Zipcode Herfindahl Index	-4.639E+02 (-31.410)	-3.992E+02 (-15.631)	-3.143E+02 (-16.004)	-3.015E+02 (-30.438)	-4.379E+02 (-18.686)	-3.771E+02 (-25.906)
Zipcode firms per worker - other ind.	-4.856E+02 (-21.991)	-4.054E+02 (-11.900)	-2.473E+02 (-9.737)	-3.169E+02 (-21.441)	-5.176E+02 (-15.053)	-3.916E+02 (-19.224)
Zipcode firms per worker - own ind.	8.030E+01 (5.702)	5.017E+01 (5.100)	5.023E+01 (8.039)	2.562E+01 (6.047)	2.284E+01 (9.331)	-2.048E+00 (-0.340)
	Urbanization Effects: Other Employment Outside Own Industry In The …					
0 to 1 Mile Ring	-1.453E-04 (-2.948)	8.510E-06 (0.137)	7.780E-06 (0.151)	-6.570E-05 (-1.775)	-7.670E-05 (-1.039)	-2.160E-05 (-0.405)
1 to 5 Mile Ring	-1.340E-05 (-1.123)	6.740E-06 (0.425)	5.070E-05 (3.532)	8.430E-06 (0.851)	-3.200E-05 (-1.766)	-3.940E-05 (-3.057)
5 to 10 Mile Ring	-1.530E-05 (-1.975)	-6.610E-06 (-0.580)	1.140E-05 (1.059)	1.440E-05 (2.122)	-1.740E-05 (-1.416)	-1.280E-05 (-1.626)
10 to 15 Mile Ring	1.370E-05 (2.031)	-4.590E-06 (-0.469)	-8.710E-06 (-0.869)	-5.860E-06 (-1.007)	-3.760E-06 (-0.335)	-3.390E-06 (-0.513)
	Localization Effects: Own Industry Employment In The …					
0 to 1 Mile Ring	3.439E-02 (15.139)	2.669E-02 (8.517)	8.406E-03 (6.019)	9.596E-03 (8.844)	3.981E-01 (10.452)	1.484E-02 (12.098)
1 to 5 Mile Ring	2.261E-03 (1.965)	5.476E-03 (2.630)	8.757E-04 (1.626)	-2.995E-04 (-0.840)	9.872E-03 (4.292)	2.740E-03 (2.910)
5 to 10 Mile Ring	8.100E-03 (11.475)	1.323E-03 (0.794)	1.313E-03 (2.503)	5.331E-04 (1.603)	8.291E-04 (0.542)	1.308E-03 (2.618)
10 to 15 Mile Ring	1.989E-03 (2.736)	6.515E-03 (4.800)	1.946E-03 (3.011)	1.274E-03 (3.801)	5.160E-03 (3.946)	2.583E-03 (6.188)
	Average Percentage Change In Localization Effect Per Mile From ...*					
1 to 3 Miles	46.71%	39.74%	44.79%	51.56%	48.76%	40.77%
1 to 7 Miles	10.92%	13.58%	12.05%	13.49%	14.26%	13.03%
1 to 12 Miles	7.85%	6.30%	6.40%	7.23%	8.23%	6.88%
	Summary Measures					
Stnd Error	131.47	124.19	110.53	83.21	136.91	101.25
Log-L	-44,996.18	-11,007.18	-13,017.79	-34,363.17	-15,259.83	-24,438.67
Uncensored	6,362	1,324	1,639	5,006	1,875	3,280
Total Obs	39,060	39,060	39,060	39,060	39,060	39,060
Fixed Effects	373	373	373	373	373	373

*Calculated as (Mile1-MileD)/d*Mile1, for D set to 1-5, 5-10, and 10-15, and d equal to 2, 7, and 12, respectively.

TABLE 3a*
LOCALIZATION EFFECTS: CONTROLLING FOR FIRM SIZE
(Numbers in Parentheses are t-ratios)

	Software SIC 7371-73, 75	**Food Products SIC 20**	**Apparel SIC 23**	**Printing & Publishing SIC 27**	**Fabricated Metal SIC 34**	**Machinery SIC 35**
	BIRTHS OF NEW FIRMS					
Employment at:						
Small firms	2.66E-03	3.03E-03	1.95E-04	5.64E-04	1.34E-03	1.02E-03
(1-24 workers)	(26.867)	(9.415)	(2.937)	(11.928)	(8.275)	(9.128)
Medium firms	1.29E-03	6.34E-04	2.01E-04	1.20E-04	5.97E-04	1.02E-03
(25-99 workers)	(13.634)	(4.474)	(4.284)	(4.157)	(6.936)	(12.172)
Large Firms	2.42E-04	1.79E-04	2.96E-04	3.93E-05	2.24E-04	1.45E-04
(100+ workers)	(9.305)	(6.123)	(7.174)	(4.480)	(6.106)	(8.652)
	NEW FIRM EMPLOYMENT					
Employment at:						
Small firms	8.97E-02	1.47E-01	-4.69E-03	1.76E-02	7.71E-02	3.08E-02
(1-24 workers)	(20.653)	(7.939)	(-2.530)	(9.816)	(7.262)	(5.793)
Medium firms	2.92E-02	3.31E-02	-2.94E-03	3.03E-03	2.95E-02	3.97E-02
(25-99 workers)	(7.132)	(4.051)	(-1.987)	(2.769)	(5.258)	(10.004)
Large Firms	1.15E-02	1.52E-02	1.54E-02	1.49E-03	2.41E-02	7.74E-03
(100+ workers)	(10.402)	(9.547)	(12.102)	(4.524)	(10.699)	(10.171)

*The estimates above are the coefficients on the localization variable for the indicated employment type having summed the localization concentric ring variables out to 5 miles and omitted the remaining localization concentric rings (miles 6 to 15). All other variables listed in Table 2 were included in the model: coefficients for those variables are not reported to conserve space.

TABLE 3b*
LOCALIZATION EFFECT: CONTROLLING FOR SUBSIDIARY STATUS
(Numbers in Parentheses are t-ratios)

	Software SIC 7371-73, 75	**Food Products SIC 20**	**Apparel SIC 23**	**Printing & Publishing SIC 27**	**Fabricated Metal SIC 34**	**Machinery SIC 35**
			BIRTHS OF NEW FIRMS			
Employment at:						
Non-Subsidiaries	4.67E-04	2.28E-04	1.41E-04	6.14E-05	2.76E-04	2.23E-04
	(22.577)	(8.047)	(9.000)	(7.511)	(12.322)	(16.123)
Subsidiaries	1.47E-03	2.03E-04	2.82E-03	7.03E-05	4.34E-04	5.25E-04
	(10.698)	(1.785)	(8.361)	(3.380)	(4.521)	(7.217)
			NEW FIRM EMPLOYMENT			
Employment at:						
Non-Subsidiaries	8.16E-03	1.43E-02	2.29E-03	1.94E-03	2.06E-02	7.41E-03
	(9.066)	(8.598)	(5.012)	(5.989)	(13.718)	(11.329)
Subsidiaries	1.36E-01	3.01E-02	3.66E-02	2.30E-03	3.39E-02	3.67E-02
	(23.389)	(4.713)	(3.386)	(2.799)	(5.387)	(10.953)

*The estimates above are the coefficients on the localization variable for the indicated employment type having summed the localization concentric ring variables out to 5 miles and omitted the remaining localization concentric rings (miles 6 to 15). All other variables listed in Table 2 were included in the model: coefficients for those variables are not reported to conserve space.

TABLE B-1a
BIRTHS OF NEW FIRMS - OLS FIXED EFFECTS WITH GEOGRAPHIC EFFECTS
(Numbers in Parentheses are t-ratios)

	Software SIC 7371-73, 75	Food Products SIC 20	Apparel SIC 23	Printing & Publishing SIC 27	Fabricated Metal SIC 34	Machinery SIC 35
	Diversity and Competition Effects					
Zipcode area in square miles	-2.831E-04 (-2.028)	1.772E-04 (1.888)	-1.897E-04 (-0.379)	-1.620E-04 (-1.759)	3.640E-05 (0.452)	-5.920E-05 (-0.591)
Zipcode Herfindahl Index	-3.507E+00 (-10.789)	-4.584E-01 (-2.056)	-8.610E-01 (-0.910)	-1.310E+00 (-8.934)	-2.796E-01 (-1.724)	-1.420E+00 (-6.478)
Zipcode firms per worker - other ind.	-9.943E-01 (-1.771)	-7.921E-01 (-2.234)	2.452E-01 (0.183)	-1.387E+00 (-4.160)	-2.425E-01 (-0.854)	-2.660E-01 (-0.712)
Zipcode firms per worker - own ind.	-3.745E-01 (-2.611)	-2.331E-02 (-0.230)	1.679E-01 (0.545)	-1.868E-01 (-1.875)	-7.755E-02 (-0.960)	-1.046E-01 (-0.892)
	Urbanization Effects					
0 to 1 Mile Ring	3.720E-06 (3.751)	1.070E-07 (0.227)	-4.900E-07 (-0.191)	1.210E-06 (1.685)	-3.720E-08 (-0.073)	2.470E-06 (2.964)
1 to 5 Mile Ring	-6.340E-07 (-3.102)	1.050E-07 (1.010)	-1.680E-06 (-3.212)	-2.780E-07 (-1.853)	-1.100E-07 (-0.985)	-6.590E-07 (-3.723)
5 to 10 Mile Ring	-4.670E-07 (-3.391)	9.680E-10 (0.012)	3.970E-07 (0.967)	1.370E-07 (1.315)	-9.580E-08 (-1.205)	-2.650E-07 (-2.438)
10 to 15 Mile Ring	-4.200E-08 (-0.342)	-9.320E-08 (-1.283)	-1.790E-07 (-0.413)	-1.270E-07 (-1.425)	9.510E-08 (1.232)	7.950E-08 (0.870)
	Localization Effects					
0 to 1 Mile Ring	2.702E-04 (7.711)	6.280E-05 (2.599)	7.146E-04 (16.310)	4.970E-05 (3.106)	1.580E-04 (5.911)	1.112E-04 (6.958)
1 to 5 Mile Ring	1.869E-04 (10.141)	2.520E-05 (1.982)	4.800E-05 (2.823)	3.140E-05 (6.315)	-3.590E-06 (-0.284)	7.770E-05 (7.322)
5 to 10 Mile Ring	1.258E-04 (11.477)	-2.090E-05 (-1.924)	1.470E-06 (0.083)	-4.940E-06 (-0.957)	1.410E-05 (1.766)	5.870E-05 (10.337)
10 to 15 Mile Ring	6.050E-05 (5.149)	1.290E-05 (1.393)	1.340E-05 (0.521)	1.860E-05 (3.651)	4.400E-06 (0.615)	-1.510E-05 (-3.319)
	Localization Effects: Average Percentage Change Per Mile*					
1 to 3 Miles	15.41%	29.94%	46.64%	18.41%	51.14%	15.06%
1 to 7 Miles	7.63%	19.04%	14.26%	15.71%	13.01%	6.74%
1 to 12 Miles	6.47%	6.62%	8.18%	5.21%	8.10%	9.46%
	Summary Measures					
Stnd Error	2.050	0.558	2.582	1.044	0.552	0.953
R-sq within	0.1818	0.0373	0.2741	0.1007	0.0580	0.1862
R-sq between	0.3092	0.0002	0.3450	0.0270	0.1363	0.1310
R-sq overall	0.2440	0.0189	0.2984	0.1084	0.0729	0.2425
Obs	6,362	1,324	1,639	5,006	1,875	3,280
Fixed Effects	357	274	278	350	318	341

*Calculated as (Mile1-MileD)/d*Mile1, for D set to 1-5, 5-10, and 10-15, and d equal to 2, 7, and 12, respectively.

TABLE B-1b
NEW FIRM EMPLOYMENT - OLS FIXED EFFECTS WITH GEOGRAPHIC EFFECTS
(Numbers in Parentheses are t-ratios)

	Software SIC 7371-73, 75	Food Products SIC 20	Apparel SIC 23	Printing & Publishing SIC 27	Fabricated Metal SIC 34	Machinery SIC 35
	Diversity and Competition Effects					
Zipcode area *in square miles*	-8.205E-04 (-0.086)	1.104E-04 (0.006)	5.167E-03 (0.260)	-3.168E-03 (-0.420)	-7.103E-03 (-0.362)	-1.581E-03 (-0.153)
Zipcode Herfindahl *Index*	-7.450E+00 (-0.334)	8.404E+01 (1.866)	1.641E+02 (4.372)	-4.697E+00 (-0.314)	3.965E+00 (0.100)	-2.563E-01 (-0.011)
Zipcode firms per *worker - other ind.*	-1.077E+02 (-2.794)	-2.658E+02 (-3.710)	-2.784E+01 (-0.523)	-7.299E+01 (-2.672)	-1.804E+02 (-2.600)	-1.576E+02 (-4.099)
Zipcode firms per *worker - own ind.*	-1.587E+01 (-1.611)	-5.471E+01 (-2.673)	-2.262E+01 (-1.852)	-2.629E+01 (-3.218)	-3.264E+01 (-1.653)	-4.063E+01 (-3.356)
	Urbanization Effects					
0 to 1 Mile Ring	-9.520E-05 (-1.397)	-1.007E-04 (-1.053)	7.045E-04 (6.926)	-1.100E-07 (-0.002)	-1.503E-04 (-1.204)	1.553E-04 (1.808)
1 to 5 Mile Ring	-9.350E-06 (-0.666)	-1.800E-05 (-0.855)	1.840E-05 (0.889)	1.070E-05 (0.867)	-3.920E-05 (-1.431)	-3.440E-05 (-1.879)
5 to 10 Mile Ring	-2.860E-05 (-3.026)	-2.320E-05 (-1.420)	-2.920E-06 (-0.179)	-3.980E-06 (-0.467)	-1.860E-05 (-0.958)	-6.770E-06 (-0.604)
10 to 15 Mile Ring	1.810E-05 (2.143)	3.520E-06 (0.239)	-1.160E-05 (-0.675)	-1.230E-05 (-1.689)	2.260E-05 (1.199)	-8.270E-08 (-0.009)
0 to 1 Mile Ring	2.773E-02 (11.529)	2.762E-02 (5.655)	-2.392E-03 (-1.377)	8.920E-03 (6.807)	3.291E-02 (5.038)	8.806E-03 (5.337)
1 to 5 Mile Ring	-2.960E-04 (-0.234)	3.186E-03 (1.240)	-1.032E-03 (-1.531)	-1.034E-03 (-2.538)	9.849E-03 (3.187)	1.091E-03 (0.996)
5 to 10 Mile Ring	7.249E-03 (9.635)	-1.358E-03 (-0.617)	8.250E-05 (0.117)	2.506E-04 (0.592)	3.218E-04 (0.164)	1.191E-03 (2.031)
10 to 15 Mile Ring	-8.578E-04 (-1.064)	4.844E-03 (2.592)	8.361E-04 (0.817)	7.385E-04 (1.769)	-3.934E-03 (-2.249)	2.984E-04 (0.634)
	Localization Effects: Average Percentage Change Per Mile*					
1 to 3 Miles	50.53%	44.23%	28.43%	55.79%	35.04%	43.80%
1 to 7 Miles	10.55%	14.99%	14.78%	13.88%	14.15%	12.35%
1 to 12 Miles	8.59%	6.87%	11.25%	7.64%	9.33%	8.05%
	Summary Measures					
Stnd Error	132.698	113.489	109.133	81.905	129.976	94.049
R-sq within	0.0666	0.1032	0.0841	0.0362	0.0473	0.0533
R-sq between	0.1370	0.0285	0.0041	0.0398	0.2279	0.1317
R-sq overall	0.0728	0.1014	0.0908	0.0369	0.0694	0.0708
Obs	6,362	1,324	1,639	5,006	1,875	3,280
Fixed Effects	357	274	278	350	318	341

*Calculated as (Mile1-MileD)/d*Mile1, for D set to 1-5, 5-10, and 10-15, and d equal to 2, 7, and 12, respectively.

TABLE B-2*
LOCALIZATION EFFECTS: CONTROLLING FOR FIRM SIZE BASED ON OWN-INDUSTRY FIRMS FIVE YEARS OR OLDER
(Numbers in Parentheses are t-ratios)

	Software SIC 7371-73, 75	Food Products SIC 20	Apparel SIC 23	Printing & Publishing SIC 27	Fabricated Metal SIC 34	Machinery SIC 35
	BIRTHS OF NEW FIRMS					
Small	3.39E-03	2.89E-02	1.69E-02	1.68E-02	1.37E-02	1.44E-02
(1-24 emp)	(18.753)	(12.337)	(16.013)	(26.374)	(12.669)	(23.733)
Medium	2.67E-03	5.10E-03	4.76E-03	-3.25E-04	2.68E-03	2.43E-03
(25-99 emp)	(16.275)	(5.069)	(7.241)	(-0.650)	(5.578)	(6.174)
Large	3.06E-04	9.71E-05	1.62E-04	-4.00E-05	2.78E-04	1.78E-04
(100+ emp)	(8.288)	(3.925)	(0.551)	(-0.511)	(2.404)	(3.805)
	NEW FIRM EMPLOYMENT					
Small	8.92E-02	1.48E+00	1.97E-01	3.89E-01	6.90E-01	3.95E-01
(1-24 emp)	(10.710)	(10.175)	(3.761)	(14.347)	(8.916)	(12.267)
Medium	6.19E-02	2.74E-01	5.64E-02	1.41E-02	1.67E-01	1.22E-01
(25-99 emp)	(8.156)	(4.395)	(1.738)	(0.658)	(4.827)	(5.814)
Large	1.54E-02	3.84E-02	2.73E-02	1.68E-02	3.90E-02	1.33E-02
(100+ emp)	(9.144)	(6.923)	(1.988)	(5.276)	(5.413)	(5.538)

*The estimates above are the coefficients on the localization variable for the indicated employment type having summed the localization concentric ring variables out to 5 miles and omitted the remaining localization concentric rings (miles 6 to 15). All other variables listed in Table 2 were included in the model: coefficients for those variables are not reported to conserve space.

References

Carlton, D.W. 1983. "The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuos Endogenous Variables," *Review of Economics and Statistics*, 65: 440-449.

Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data," *Review of Economic Studies* , 47: 225-238.

Chamberlain, Gary. 1984. "Panel Data". In Z. Griliches and M. Intriligator (eds.), *Handbook of Econometrics*, Chapter 22, Vol. 2. New York: Elsevier Science Publishing Company.

Chinitz, B.J. 1961. "Contrasts in Agglomeration: New York and Pittsburgh," *American Economic Review*, 51: 279-89.

Ciccone, A. and R.E. Hall. 1996. "Productivity and the Density of Economic Activity," *American Economic Review*, 86: 54-70.

Ellison, G. and E. Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 105: 889-927.

Gabriel, Stuart and Stuart S. Rosenthal. 1996. "Commute Times, Neighborhood Effects, and Earnings: An Analysis of Compensating Differentials and Racial Discrimination," *Journal of Urban Economics*, 40: 61-83.

Glaeser, E.L, H.D. Kallal, J.A. Scheinkman, and A. Shleifer. 1992. "Growth in Cities," *Journal of Political Economy*, 100: 1126-1152.

Helsley, R.W. 1990. "Knowledge and Production in the CBD," *Journal of Urban Economics*, 28: 391-403.

Henderson, J.V. 1986. "Efficiency of Resource Usage and City Size," *Journal of Urban Economics*, 19: 47-70.

Henderson, J.V. 1988. *Urban Development: Theory, Fact, and Illusion*. Oxford: Oxford University.

Henderson, J.V. 1994. "Where Does an Industry Locate?" *Journal of Urban Economics*, 35: 83-104.

Henderson, J.V., A. Kuncoro, and M. Turner. 1995. "Industrial Development in Cities," *Journal of Political Economy*, 103: 1067-1085.

Hoover, E. and R. Vernon. 1959. *Anatomy of a Metropolis*. Cambridge: Harvard University Press.

Hsiao, Cheng. 1986. *Analysis of Panel Data*. New York: Cambridge University Press.

Imai, H. 1982. "CBD Hypothesis and Economies of Agglomeration," *Journal of Economic Theory*, 8: 275-299.

Jacobs, J. 1969. *The Economy of Cities*. New York: Vintage Books.

Jaffe, A.B., M. Trajtenberg, and R. Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, CVIII: 577-598.

Krugman, P. 1993. "On the Number and Location of Cities," *European Economic Review*, 37: 293-298.

Marshall, A. 1920. *Principles of Economics*. London: MacMillan.

Moomaw, R.L. 1981. "Productivity and City Size: A Critique of the Evidence," *Quarterly Journal of Economics*, 96: 675-688.

Nakamura, R. 1985. "Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities," *Journal of Urban Economics*, 17: 108-124.

Ogawa, H. and M. Fujita. 1980. "Equilibrium Land Use Patterns in a Nonmonocentric City," *Journal of Regional Science*, 20: 455-475.

O'Hara, D.J. 1977. "Location of Firms Within a Square Central Business District," *Journal of Political Economy*, 85: 1189-1207.

Quigley, J. 1998. "Urban Diversity and Economic Growth," *Journal of Economic Perspectives*, 12: 127-138.

Romer, P.M. 1986. "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94: 1002-1037.

Rosenthal, Stuart S. and William C. Strange. 1999. "Localization and Growth," mimeo, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University.

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

Sveikauskas, L. 1975. "The Productivity of Cities," *Quarterly Journal of Economics*, 89: 393-413.

Vernon, R. 1962. *Metropolis 1985*. Cambridge, MA: Harvard University Press.