Available online at www.sciencedirect.com
ScienceDirect
Journal of Urban Economics 61 (2007) 86–111
JOURNAL OF Urban Economics
www.elsevier.com/locate/jue

Smart Café Cities: Testing human capital externalities in the Boston metropolitan area

Shihe Fu

Research Institute of Economics and Management, Southwestern University of Finance and Economics, 55 Guanghua Cun Street, Chengdu, Sichuan 610074, PR China

Received 8 February 2005; revised 1 June 2006

Available online 13 July 2006

Abstract

Existing studies have explored only one or, in some cases, two mechanisms by which human capital externalities percolate at the macrogeographic levels. This study, however, uses the 1990 Massachusetts census data to test four mechanisms at the microgeographic levels, in the Boston metropolitan area labor market. We propose that individual workers can learn from their occupational and industrial peers in the same local labor market through four channels: depth of human capital stock, Marshallian labor market externalities, Jacobs labor market externalities, and thickness of the local labor market. We find that all types of human capital externalities are significant across census blocks. Different types of externalities attenuate at different speeds over distances. For example, the effect of human capital depth decays rapidly beyond three miles away from block centroids. We conclude that knowledge spillovers are very localized within a microgeographic scope in cities that we call, "Smart Café Cities."

JEL classification: C21; J24; J31; R23

Keywords: Human capital externalities; Labor market agglomeration; Hedonic wage model; Marshallian externalities; Jacobs externalities; Spatial attenuation

E-mail address: fush@swufe.edu.cn.

0094-1190/$ – see front matter
doi:10.1016/j.jue.2006.06.002

Even walking with any two other people, I will always find a teacher among them.

Confucius, *Analects*, Book 7, 21

Most of what we know we learn from other people. We pay tuition to a few of these teachers... but most of it we get for free, and often in ways that are mutual—without a distinction between student and teacher... What can people be paying Manhattan or downtown Chicago rents for, if not for being near other people?

Robert E. Lucas, Jr., 1988, pp. 38–39

1. Introduction

A high concentration of skilled workers can promote the creation, diffusion, acquisition, and accumulation of knowledge across individual workers, geographic space, and time. Workers benefit from being close to a dense, skilled, labor market where, through different channels, they can learn from others without compensation. For example, "if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of new ideas" (Marshall [32, p. 271]). Such productivity-enhancing external benefits of labor markets are called human capital externalities, knowledge spillover effects, learning externalities, or labor market local agglomeration economies, whichever you choose. Uncompensated externalities from aggregate human capital stock have long been considered one of the important forces of economic growth (Romer [41], Lucas [29]). Further, local human capital externalities are considered to be one of the predominant reasons for the existence of cities (Henderson [22], Fujita and Ogawa [15], Lucas [30]) and urban endogenous growth (Palivos and Wang [34], Eaton and Eckstein [12], Black and Henderson [6]).

Empirically, firm data, wage data, and housing or land price data have been used to test human capital externalities. The disadvantage in using firm data is that it requires a broad set of control variables to separate it from other sources of benefit that firms obtain from being close to each other, such as natural advantage, input sharing, and forward or backward linkages. Land prices usually are not directly observable. Estimating hedonic housing models to infer human capital externalities is reasonable, but it omits information on individual workers. Other indirect methods may be possible; for example, Jaffe, Trajtenberg, and Henderson [26] used patent citation data to study the geographical localization of knowledge spillovers. The ideal direct way to identify human capital externalities is to employ wage data (assuming workers are paid by their marginal value of products). A good example of this is the paper by Wheaton and Lewis [49], where wage data is used to test labor market agglomeration across Metropolitan Statistical Areas (MSAs).

Two important questions have not yet been answered in the literature. The first question is: "What are the microfoundations of knowledge spillovers?" Or, put another way, "How do human capital externalities percolate?" Though people do exchange information and ideas and learn from each other by socializing in downtown cafés, unfortunately, we can not directly observe how knowledge spills out among buildings and across streets. We hypothesize, by distilling the existing literature, that there exist the similar mechanisms of agglomeration economies in a dense labor market as those of firms' concentration, and interpret the positive effects of local labor market agglomeration on individual earnings as the evidence of human capital externalities.[1] We propose that workers can learn from their occupational and industrial peers, who are in the same local labor market, through four channels:

(1) the depth (quality) of human capital stock in the local labor market;
(2) Marshallian labor market externalities, or the specialization and peer competition effects;
(3) Jacobs labor market externalities, or the diversity of the local labor market in terms of occupations and industries; and,
(4) the thickness (density) of the local labor market, or labor market pooling effects.

The second question is: "What is the geographic scope of human capital externalities?" For example, at what geographic level are human capital externalities the strongest and most significant? How do knowledge spillovers attenuate spatially? Geographical proximity reduces communication costs and increases the frequency of social interactions. Intuitively, knowledge spillovers through social interactions are very localized. However, most empirical papers have used data at only metropolitan (Wheaton and Lewis [49]), city or county (Timothy and Wheaton [48]), or zip code level (Rosenthal and Strange [43], [44]).

The 1990 Massachusetts census data enable us to explore labor market agglomeration down to the census tract and block levels. The data set contains detailed information on surveyed workers' individual characteristics and their workplace. We can identify at which census tract and census block a worker worked. Therefore, we can construct indices to measure different types of human capital externalities for each work tract and work block.[2] We then estimate hedonic wage models with different location fixed effects and identify that almost all types of human capital externalities are significant across blocks as well as across tracts. Furthermore, Marshallian labor market externalities and the effect of labor market thickness, in terms of industry employment density, are significant at the block level. We also estimate a spatial attenuation effects model, and find that different types of human capital externalities attenuate with distance at different speeds. For example, occupational Marshallian labor market externalities decay rapidly beyond 1.5 miles away from block centroid; the effect of human capital depth decays rapidly beyond three miles; while Jacobs externalities, in terms of industry diversity, decay very slowly within nine miles, indicating a certain degree of urbanization economies. Since human capital externalities are very localized, we call dense urban areas, "Smart Café Cities." [3]

One important issue in this study is worth investigating: if workers can free migrate to the locations where wages are higher, then, why aren't economically identical workers paid the same at locations with different human capital externalities? This is because in order to equalize individual utility levels and prevent all workers from concentrating in a particular area, land rents must increase to achieve spatial equilibrium. Theoretically, such interaction between the labor market and the housing market was modeled by Roback [40] and Rauch [39], and clearly summarized by Morettie [33] and Rosenthal and Strange [42]. All those models assumed that workers live and work in the same location (city). What if workers live and work in different census tracts?

[1] For the survey of literature on agglomeration economies from the concentration of firms or industries, see Duranton and Puga [11] and Rosenthal and Strange [45].

[2] In the rest of the paper, census tracts refer to the tracts of employment or tracts where workers worked, not the tracts where workers lived. The same rule applies to blockgroups and blocks.

[3] The term "Smart Cities" was proposed by Shapiro [46]. "Café Cities" was proposed by Professor Richard Arnott in his lectures on urban economics at Boston College.

Fortunately, there exist some other spatial equilibrium mechanisms. For example, workers working in a tract with higher wages may end up living further away from their workplace. The higher commuting costs partly offset the higher wages. However, it is beyond the scope of this paper to fully address this equilibrium issue.

The rest of the paper is organized as follows: Section 2 proposes four percolation channels of human capital externalities, Section 3 discusses the geographic scope of human capital externalities, and Section 4 specifies the econometric models and discusses the identification strategies. Section 5 introduces the data set, Section 6 presents the estimate results, and Section 7 concludes.

2. The microfoundations of human capital externalities

Existing studies have explored only one or two of the dimensions of knowledge spillover mechanisms. For example, Jovanovic and Rob [27], Rauch [39], Simon and Nardinelli [47], and Shapiro [46] studied only the depth of human capital.[4] Wheaton and Lewis [49] tested Marshallian labor market externalities and the thickness of a labor market. Ciccone and Hall [8] studied only the density of economic activity. Jacobs labor market externalities have never been tested in literature.

We propose that there are four types of percolation mechanisms of human capital externalities through local labor markets: depth (quality) of human capital stock, Marshallian labor market externalities, Jacobs labor market externalities, and thickness of labor markets. These four channels capture the four dimensions of knowledge: the vertical and horizontal differences, local specialization, and spatial density of knowledge. They can be considered the local labor market attributes that promote human capital accumulation and enhance workers' productivity.

In the following subsections, we define the four types of human capital externalities in detail and design a set of variables to measure them. Our unique data set enables us to test these four types of human capital externalities within one model specification.

2.1. The depth of human capital stock

We define the degree of advancement or sophistication of a certain type of human capital as the depth or quality of human capital, which reflects the vertical difference of knowledge. Workers can learn more and faster from others who have better human capital in their fields than from those with lower human capital levels. High-quality human capital can enhance the ability to absorb existing ideas and create new ones. Therefore, even though well-educated workers may learn less from less educated neighbors, they still learn much from the concentration of well-educated peers.

School education is the typical way to deepen human capital; therefore, the average level of education in a labor market is a good proxy for the depth of human capital stock. We use the share of workers with a college degree, or higher, in a labor market, since an increase in the share of college graduates in a labor market implies an increase in the average level of education.[5]

[4] Charlot and Duranton [7] identified that workplace communication externalities can explain only about one tenth of the effects of city size and average urban schooling on individual earnings, which hints that there must exist other channels of knowledge spillovers.

[5] Work experience is also a good measurement of human capital depth; but, Rauch [39] found that the average level of education has a much greater external effect on wages than the average level of experience.

A comprehensive literature survey on human capital externalities—in terms of the average level of education—was done by Morettie [33]. The first theoretical model incorporating human capital depth probably was constructed by Jovanovic and Rob [27]. Their model shows that knowledge spillovers depend on the vertical differences in what people know. Rauch [39] provided probably the first empirical test of the effects of human capital depth on urban wages and land rents. He used the 1980 census data on individual workers in over 200 US Standard Metropolitan Statistical Areas (SMSAs) and found that a metropolitan area with an average educational level one-year higher than another would have about a 3% productivity advantage. Simon and Nardinelli [47] found that city-aggregate and metropolitan areas with a higher percentage of college graduates grew faster over the 20th century in the United States. Shapiro [46] used the 1940, 1970, 1980, and 1990 Integrated Public Use Microdata Series (IPUMS) databases and concluded that the deepening of human capital contributes to the growth of urban employment, wages, and housing value. In his overall sample, a 10% increase in the share of college educated residents generated a 0.2%, 0.8%, and 0.7% increase in wage, urban employment, and housing value growth, respectively. It is in this sense that he called cities, "Smart."

2.2. Marshallian labor market externalities

The original idea of human capital externalities probably first dates back to Alfred Marshall [32]. Marshall emphasized that human capital externalities take place mostly between workers in the same industry and city through face-to-face interactions. He also stressed technological spillovers from one firm to another firm nearby within the same industry in a city. Marshall defined the benefits that a firm obtains from the general development of the industry as external economies ([32, p. 266]). It is within this dynamic context, recently, that urban economists developed the concept, "Marshallian externalities," meaning firms can benefit from the concentration of same-industry firms in an intertemporal context (Glaser et al. [20]).[6] Empirical work has identified that Marshallian externalities are significant in many industries. For example, Henderson [23] found that Marshallian externalities have strong productivity effects in high-tech industries.

Workers can learn from the local concentration of same-occupation and same-industry peers. We refer to this as "Marshallian labor market externalities." The mechanism is such that the concentration of specialized skilled workers generates more competition, which then provides a strong motivation for workers to learn the most up-to-date knowledge, thereby speeding the creation and diffusion of new knowledge. This is the crucial point of Porter's theory of competitive advantage in regional clusters (Porter [36,37]). Also, increased division of labor generates comparative advantage compared with under-specialized regions.

We use the degree of occupation (industry) specialization at a location to measure Marshallian labor market externalities. The occupation (industry) specialization index is the ratio of employment in a certain occupation (industry) at a location to the total employment at that location. This index measures the intensity and frequency of social interactions and knowledge spillovers

[6] In the urban economics literature, economies external to firms but internal to the industry are dubbed "localization economies" in static context; in this case, individual firms benefit from the local concentration of same-industry firms. Economies external to industries but internal to a city are called urbanization economies; here, individual firms benefit from the concentration of different industries in a city. Informally speaking, dynamic localization economies are called Marshallian or Marshall–Arrow–Romer externalities, and dynamic urbanization economies are called Jacobs externalities. One of the main sources of dynamic externalities is the accumulation of different types of human capital.

among same-occupation (same industry) workers at a location. Wheaton and Lewis [49] were most likely the first to test Marshallian labor market externalities. They used manufacturing industry wage data from the 5% Public Use Micro Sample (PUMS) of the 1990 US census and found that the differences in occupation specialization and industry specialization across MSAs could generate 23% and 30% higher wages, respectively.

2.3. *Jacobs labor market externalities*

Jacobs [24,25], with many concrete examples, emphasized that it is the variety and diversity of geographically proximate industries that promote innovation and city growth. This is why the benefits from urban diversity in the dynamic context are called, "Jacobs externalities." Firms benefit from urban diversity due to the following external economies: shared inputs, lower transaction costs, and statistical economies of scale in production and consumption. Examples would be business services and consumption amenities, labor market matching and shopping districts, and unemployment insurance in a diverse labor market (Quigley [38], Duranton and Puga [10]). New economic geography models (Fujita, Krugman, and Venables [16]) and some endogenous growth models (Barro and Sala-I-Martin [2, Chap. 6]) show that diversity and variety in producer inputs or in consumption goods can generate external scale economies.

The empirical results of testing Jacobs externalities are mixed. Glaeser et al. [20] concluded that urban diversity encourages employment growth in industries, while Henderson [23] found little evidence of Jacobs externalities in the high-tech and machinery industries.

Workers benefit from the diversity of labor markets, which we call, "Jacobs labor market externalities." One reason for this is that many inventions are interdisciplinary, due to the stimulation of "ideas" in heterogeneous surroundings in cities. Berliant, Reed, and Wang [5] assumed that individuals possess horizontally differentiated types of knowledge and randomly search for partners with whom to exchange ideas in order to improve production efficacy. Their model shows that the addition of new knowledge, through matching, results in endogenous growth.

The second reason is that labor market diversity reflects an open and tolerant social and cultural milieu that attracts different types of talented people to that area. Florida [14] constructed a Bohemia index—the bohemian population at the MSA level—and found that the presence and concentration of bohemians is highly correlated with the concentration of high human capital individuals and with innovative high-tech industries.[7] In another paper (Florida [13]), he found that the geographic distribution of talent is closely associated with diversity (meaning low entry barriers) and urban amenities.

We construct an occupation (industry) diversity index that equals one minus the Herfindahl index of occupations (industries). The Herfindahl index is the sum of squared shares of employment of different occupations (industries) at a location. It is possible that the spatial concentration of different industries may imply scale economies from forward or backward linkages, but other indices such as inputs or volume of shipments are a better measurement. Our diversity indices are based on the number of employees in different occupations and industries, which measure the broadness and horizontal differences of human capital stock in a local labor market. To the best of our knowledge, this paper is the first to test the effect of labor market diversity on individual earnings.

[7] Florida's selection of bohemian occupations included: authors; designers, musicians and composers; actors and directors; craft-artists, painters, sculptors, and artist printmakers; photographers; dancers; artists, performers, and related workers.

2.4. The thickness of a labor market

The more densely concentrated a labor market in a limited geographic area, the more luck workers will have in their random matches; that is, workers benefit from the thickness or density of a local labor market. In the literature, labor market pooling (from the viewpoint of firms) sometimes also means the thickness of a labor market, though it is used to explain both localization and urbanization economies. The importance of labor market thickness at microgeographic levels is that workers can socialize more frequently and build social networks more easily to exchange information. Bayer, Ross, and Topa [3] detected that social interactions among block neighbors help workers to build informal hiring networks, which have a significant impact on a wide range of labor market outcomes.

Employment density, the number of workers per square kilometer, is a simple index for gauging the thickness of a local labor market. An alternative is an occupation (industry) concentration index, which is the ratio of employment of an occupation (industry) at a location to the total employment of that occupation (industry) over all the locations. However, the values of concentration indices depend on the specification of geographic units.

In a labor market with imperfect information, firms and workers search for each other to form an ideal match. The larger or more dense a labor market is, the higher the probability of a better match between jobs and workers with heterogeneous human capital. This labor market pooling effect can generate agglomeration economies even without learning behavior (Helsley and Strange [21]). Ciccone and Hall [8] were the first to put density of economic activity into theory and empirical test. Their models show that spatial density results in aggregate increasing returns: a doubling of employment density in a county results in a 6% increase in average labor productivity. These locally increasing returns can explain more than half of the variance of output per worker across the United States. Wheaton and Lewis [49] found that the differences in occupation and industry concentration across MSAs could generate 12% and 16% higher wages, respectively.

3. The geographic scope of human capital externalities

The flow of knowledge across geographic space is costly. Information spillovers, which require frequent contact between workers, may dissipate over a short distance, since walking to a meeting place becomes more difficult, or random encounters become more rare, in a far away and less dense area. But what is the exact spatial scale where human capital externalities take place? How fast do knowledge spillovers decay spatially? These are empirical questions.

Most empirical work on agglomeration economies has used aggregate data, taking countries, states, MSAs, cities, or counties as geographic units. Rosenthal and Strange [44] used zip code level firm data and studied six industries. In most industries, same industry employment encourages the number of births per square mile and new establishment employment per square mile. Rosenthal and Strange interpreted this as the evidence of significant localization economies. They found that localization economies attenuate with distance: the initial attenuation is rapid, with the effect of own-industry employment in the first mile away from zip code centroid up to 10 to 1000 times larger than the effect two to five miles away; beyond five miles, attenuation is much slower. Duranton and Overman [9] also found that business agglomeration effects are very localized (at zip code level) and decay rapidly with distance.

As for knowledge spillovers, Lucas [29] argued that metropolitan areas are the most appropriate units to examine when looking for the productivity-enhancing effects of human capital abundance. Rauch [39] found evidence of human capital externalities at the SMSA level. Simon and Nardinelli [47] argued that knowledge spillovers are geographically limited to the city and much knowledge is most productive in the city within which it is acquired. They found that the estimated effects of human capital on employment growth are very large at the city-aggregate level, but smaller at the metropolitan area level. Wheaton and Lewis [49] tested Marshallian externalities and the effect of local labor market thickness at the MSA or Consolidated Metropolitan Statistical Area (CMSA) level. Beeson, Delong, and Troesken [4] identified human capital infrastructure as a determinant of population growth at the county level over the period 1840–1990.

No work has been done to identify the microgeographic scope of human capital externalities.[8] In this paper, we make a contribution to this unexplored topic. We construct labor market attribute indices at the census tract, blockgroup, and block levels, and estimate hedonic wage models with different location fixed effects. We find that all of the proposed four types of human capital externalities are strong and significant across tracts, and even across blocks. Marshallian labor market externalities and the effect of industry employment density are significant at the block level. We also find that the effects of human capital externalities attenuate at different speeds, depending on the distance away from block centroid. It is in this sense that we call cities, "Café Cities."

4. Model specification and identification

In this section we specify two types of models and discuss the identification strategies. The first is a benchmark model for testing the magnitude and significance of the types of human capital externalities at different spatial scales. The second is constructed to detect the spatial decay patterns of those externalities.

4.1. The benchmark model

The benchmark model is specified as an augmented hedonic wage model including both individual characteristics and local labor market attributes. The labor market attribute indices are constructed at the census tract level.

$$\log W_{noij} = \alpha + \lambda_c + \beta' X_n + \gamma' X_j + \epsilon_{noij}, \tag{1}$$

where W_{noij} is the imputed hourly wage of worker n, whose occupation is o and who worked in industry i at census tract j, α is a constant, λ_c is county fixed effects, X_n is the characteristics vector of worker n, X_j is the attributes vector of local labor market at census tract j, β and γ are the coefficient vectors to be estimated, and ϵ_{noij} is the error term.[9]

Variables of individual characteristics include age, age square (proxy for work experience), dummy variables for gender, marriage, race, education, English proficiency, student, veteran, and workplace disability. Industry and occupation dummies are also included to control for industry

[8] The literature on "neighborhood effects" studied social interactions at different geographical levels. Bayer, Ross, and Topa [3] detected social interaction (informal hiring referral) between workers at the census block level.

[9] Census tracts are relatively permanent statistical subdivisions of a county. A census tract includes about 2500 $\sim$ 8000 persons. Census blocks are small areas bounded on all sides by visible features such as streets, roads, and streams, and by invisible boundaries such as a city, town, etc. A census blockgroup is a cluster of blocks having the same first digit of their three-digit identifying numbers. A blockgroup includes 250 $\sim$ 550 housing units (the ideal size is 400 units). Blockgroups are nested within a tract.

Table 1
Labor market attribute variables

Variable	Definition
AveEdu	Percentage of college or higher degree-holders among all the workers at tract j, measuring the overall quality of human capital stock at tract j.
AveEduOcc	Percentage of college or higher degree-holders among all the workers of occupation o at tract j, measuring the quality of human capital stock of occupation o at tract j.
AveEduInd	Percentage of college or higher degree-holders among all the workers in industry i at tract j, proxy for the quality of human capital stock of industry i at tract j.
OccSpec	The ratio of the number of workers of occupation o at tract j to the total number of workers at that tract, proxy for Marshallian externalities of occupation o at tract j.
IndSpec	The ratio of the number of workers in industry i at tract j to the total number of workers at that tract, proxy for Marshallian externalities of industry i at tract j.
OccDiver	One minus the Herfindahl index in terms of occupations, proxy for Jacobs externalities in terms of occupations at tract j.
IndDiver	One minus the Herfindahl index in terms of industries, proxy for Jacobs externalities in terms of industries at tract j.
OccDens	Number of workers of occupation o per square kilometer at tract j.
IndDens	Number of workers in industry i per square kilometer at tract j.

and occupation specific effects. Since commuting costs must be capitalized into wages if firms are located at different points within a metropolitan area (Petitte and Ross [35], Timothy and Wheaton [48]), we also include average commuting time (minutes) to the workplace in the model.

Local labor market attributes include indices for human capital depth of local labor market and of each occupation and industry, occupation specialization, occupation diversity, occupation employment density, industry specialization, industry diversity, and industry employment density. Table 1 lists all the variable names and their definitions. To be more specific, the labor market attribute variables are constructed as follows:

AveEdu: percentage of college or higher degree-holders among all the workers at tract j, representing the overall depth of human capital stock at tract j.

AveEduOcc: percentage of college or higher degree-holders among all the workers of occupation o at tract j, representing the depth of human capital stock for occupation o employment at tract j.

AveEduInd: percentage of college or higher degree-holders among all the workers in industry i at tract j, representing the depth of human capital stock for industry i employment at tract j.

OccSpec: occupation specialization index for workers of occupation o at tract j. It is the ratio of the number of workers of occupation o at tract j to the total number of workers at that tract, used to proxy for Marshallian labor market externalities among occupational peers.

OccDiver: occupation diversity index for workers at tract j. It equals one minus the Herfindahl index of occupations, representing Jacobs labor market externalities in terms of occupation diversity. Let S_{oj} denote the ratio of occupation o employment at tract j to the total employment at tract j. Define the occupation diversity index as

$$OccDiver = 1 - \sum_o S_{oj}^2.$$

Note that if there is only one occupation in a tract, the Herfindahl index equals 1, and the diversity index equals 0. If there are many occupations in a tract, and the share of employment in each occupation is very small, then the Herfindahl index is closer to 0, and the occupation diversity index becomes closer to 1. The larger the value of the occupation diversity index, the more diverse the local labor market is, in term of occupations.

OccDens: occupation employment density index for workers of occupation o at tract j. It equals the number of workers of occupation o per square kilometer at tract j, and is used to measure the thickness of a local labor market in terms of occupation o employment. We do not use the occupation concentration index, as Wheaton and Lewis [49] did, because it is not geographically invariant.

The same methods are applied to the construction of industry indices. For example: the industry specialization index would be:

$$IndSpec = \frac{\text{number of workers in industry } i \text{ at tract } j}{\text{total employment at tract } j}.$$

The industry diversity index would be:

$$IndDiver = 1 - \sum_i S_{ij}^2,$$

where S_{ij} denotes the ratio of employment in industry i at tract j to the total employment at tract j.

IndDens: industry employment density index for workers in industry i at tract j. It is the number of industry i workers per square kilometer at tract j.

We estimate model (1) by pooling all the data, as well as by occupation and by industry. The results are reported in Sections 6.1 and 6.2.

4.2. Identification strategies

The error term captures the effects of unobservable locational attributes, unobservable individual characteristics, and measurement errors. Most likely the error terms are spatially correlated and are not identically and independently distributed. In this subsection we discuss how to deal with these problems.

A location may have better accessibility, or other natural or historical advantages, which are omitted in model (1). We use location fixed effects to capture the omitted locational attributes.[10] For example, in model (1), the county fixed effects control for county specific attributes; therefore, we can identify the significance of all types of human capital externalities within a county or across tracts. However, unobservable tract specific attributes may affect our estimation. To identify what types of human capital externalities are significant within tracts, we estimate model (1) with tract fixed effects, but we have to drop variables *AveEdu*, *OccDiver*, and *IndDiver*, since they are invariant within each tract. By the same token, to identify what types of human capital externalities are significant across (within) census blocks, we construct the labor market attribute indices at the census block level, and estimate a model with blockgroup (block) fixed effects.

Workers may have unobservable characteristics or intrinsic preferences, which correlate with the independent variables in model (1). For example, workers with high-level human capital may strongly prefer to work in the place where well-educated workers concentrate, even though no learning externalities occur (though this is highly impossible). Another example is working

[10] An important observable but heterogeneous location attribute is the employment scale. The scale effect could be positive (due to agglomeration economies) or negative (due to congestion). The location fixed effects can capture the scale heterogeneity. Also the employment density indices can reflect the net effect of employment scale.

effort. Well-educated workers may work harder or more intensely, and their efforts are compensated by higher wages (Leamer [28]). Such sorting effects cause the correlation between the error term and some labor market attribute indices. The omitted individual characteristic variables will make the coefficients of some labor market attributes overestimated. Bayer, Ross, and Topa [3] provided evidence that there is little block-level correlation in unobserved attributes within blockgroups. Therefore, sorting is not an issue in our blockgroup fixed effects models. However, sorting bias probably is an issue at large geographic scale. One way to deal with sorting problem is to use instrumental variables estimation. For example, Rosenthal and Strange [42] employed several geological variables as instruments for agglomeration measures. However, in our data we know of no such variables, which are correlated with local labor market attributes, but not with the error terms. Another approach is to use time-invariant individual fixed effects to capture all the unobservable individual characteristics (Glaeser and Maré [19]). This requires a panel data set. Since we only have a cross-section data set, we leave this endogeneity problem for future research.

Within each location, error terms may not be identically and independently distributed. We use the Huber/White estimate of variance, clustered by locations, to produce consistent standard errors.[11] However, locations nearby may share some common attributes, such as infrastructures, economic policy, and complementary industries. This may cause spatial autocorrelation across locations. Since census geographic units are classified according to the homogeneity of demographic, economic, and housing characteristics, the cross-location correlation is not a serious problem. The better alternative may be to use non-parametric or semi-parametric estimates, or spatial econometric methods by using spatial weights (Gibbson and Machin [18], Anselin [1]). We leave this on the future agenda.

No matter how carefully designed and implemented, the census data still contain measurement errors, such as undercount. The measurement error problem will be magnified when using data at lower geographic levels, but less serious at aggregate levels or at locations with large observations. Measurement errors in independent variables will cause the coefficients to be underestimated. We will take this into account when we interpret estimate results from lower geographical models. We also estimate models by selecting locations where the number of workers is greater than a certain number, in order to obtain a sense of the measurement errors.

Some indices are moderately correlated (see Table A.5 in Appendix A), which hints that our proposed human capital externalities percolation channels may interact with each other. For example, high diversity may attract highly-educated workers; high occupation diversity may imply low specialization for some occupations; a high degree of specialization for some occupations in downtown may also imply high occupation employment density. However, in this study we do not consider the interaction problem since we are particularly interested in identifying the different dimensions of human capital externalities. Our huge sample size can reduce the standard errors and partly remedy the collinearity problem.

4.3. The spatial attenuation model

To test how human capital externalities attenuate with distance, we adopt the methodology proposed by Rosenthal and Strange [44]. For each block, we construct concentric rings of various radii away from the centroid of that block; for every ring, we construct the seven indices of labor

[11] We use the STATA command “areg” with option “cluster,” which allows that the residuals are not identically distributed and are not independent within clusters.

market attributes, respectively, based on the employment in that ring.[12] We then estimate model (1) including the indices of all the rings.

Given that the census data provide individual residential and workplace information only down to block level, we cannot precisely compute distances between any two workers. We assume that all employment in a block concentrates at the block centroid and compute the distance between any two block centroids.[13] If a block centroid is within a particular ring, then the whole area of that block is considered in that ring, too. By employing this method, not only can we compute the total number of workers in each ring, but, also, we know the characteristics of each worker in each ring. Thus, we are able to construct human capital externality indices for each ring.

The alternative is to assume that employment is uniformly distributed at each block, then construct rings of certain miles away from each block centroid. In order to infer the proportion of employment from each overlapped block in a ring, Geographic Information Systems (GIS) software is needed to compute areas of all the parts of blocks that overlap with that ring. However, in our sample, each worker has a set of heterogeneous characteristics. How we assign workers in a block to a particular ring will affect the values of labor market attributes for that ring. Therefore, it is impossible to construct all indices for each ring in a consistent way. Rosenthal and Strange [44] adopted the uniform distribution assumption because they used the total number of employees in a location, which does not depend on individual specific characteristics. Compared with metropolitan areas, it does not make much difference whether one assumes that workers are uniformly distributed within a block or concentrate at block centroid.[14]

The results of the spatial attenuation model are reported in Section 6.3.

5. Data

We use the restricted version of the 1990 Massachusetts census data and access the data at the Boston Census Research Data Center. The data is also commonly referred to as the "long form." The data set includes approximately one-sixth of the population and contains detailed information on surveyed individuals' personal characteristics, family structure, geographic information of residential and work place, and housing characteristics. The main advantage of the restricted version is that it includes geographic information down to the census block level, while the PUMS is primary based on metropolitan/nonmetropolitan areas for the 1% sample, and Public Use Microeconomic Areas (PUMAs) for the 5% sample.[15]

The sample used in this paper is constructed as follows: select workers ages 16–65, working in the Boston metropolitan area (BMA), who reported non-zero wages, hours usually worked per week, and weeks worked in the previous year.[16] We exclude workers whose industry was agriculture, mining, military, or not classified. We also exclude workers who had disabilities that

[12] Here we use only *AveEdu* to measure human capital depth. The inclusion of *AveEduOcc* and *AveEduInd* does not change the results much, but makes the presentation harder since we already have too many variables.

[13] We use ArcView GIS 3.3 software to locate the longitude and latitude of each block centroid from the 1990 Massachusetts census block map, then compute the distance between any two block centroids.

[14] Actually, Rosenthal and Strange also estimated their models by assuming employment concentrate at the centroid of each zip code area. The results are very similar.

[15] PUMAs are primarily based on counties, and may be whole counties, groups of counties, or places. Each PUMA includes at least 100,000 persons.

[16] The Boston Metropolitan Area (BMA) includes all of Suffolk county, and portions of six other counties: Bristol, Essex, Middlesex, Norfolk, Plymouth, and Worcester county.

prevented them from working. For each worker, we identify his (or her) workplace at the county, census tract, blockgroup, and block levels. For each work location, we construct variables proxy for different types of human capital externalities, based on the number of workers and their characteristics at that work location. For the tract level model, we select only work tracts where the number of workers is greater than 1.[17] We apply the same rule for blockgroup and block level models. The tract level model sample includes 150,952 workers, 7 counties, 621 census tracts, 2461 blockgroups, and 11,395 blocks. Some summary statistics are listed in Table A.1 in Appendix A. Tables A.2 and A.3 list the three-digit level industries and occupations we classify in this paper and their employment shares. Table A.4 presents the mean and standard deviations across tracts for all the tract-level labor market attribute indices for some selected occupations and industries. Table A.5 presents the correlation matrix for all the labor market attribute indices.

6. Results

6.1. Benchmark model results

We first estimate model (1) with county and tract fixed effects. The results are reported in Table 2.

In Table 2, the county fixed effects model shows that estimated coefficients for all the variables of individual characteristics and local labor market attributes are significant at the 5% level, except the coefficient of occupation employment density index, which is significant at the 10% level. The effects of human capital depth at the tract level are decomposed into three components: the depth of overall, same-occupation, and same-industry human capital stock. The coefficient of *AveEdu* shows that a 1% increase in the share of workers at a tract who are college graduates increases workers' hourly wage at that tract on average by 0.15%; one standard deviation of *AveEdu* (0.126) generates 1.89% increase in wages across tracts. The effects of human capital depth of occupational and industrial peers are smaller, the semi-elasticity is about 0.05 and 0.04.

The Marshallian labor market externalities reflected by the occupation specialization index are stronger than the effects of overall human capital depth: the semi-elasticity is about 0.17. The Jacobs externalities, indicated by the occupation and industry diversity variables, are strong. One standard deviation increase in occupation diversity (0.057) across tracts generates a 1.71% increase in hourly wages; one standard deviation in industry diversity (0.105) generates 2.21% higher hourly wages. The effect of labor market thickness in terms of industry employment density is very strong: adding one more same-industry worker to a square kilometer area at a tract generates a 3.7% wage increase. Note that we controlled for individual characteristics, county, industry, and occupation specific effects; our results reflect the very general effects of human capital externalities across heterogeneous workers within a county or across tracts.

To test the stability of the model specification, we also estimate a series of models slightly different from the benchmark model. For example, drop the occupation dummies, industry dummies, or diversity indices, and select only workers in non-government sectors, and select only tracts where the number of workers is greater than 100 or 200. The coefficients are somewhat different, but, the overall patterns are similar (results are not reported here). The presented model is our preferred specification.

[17] If there is only one worker in a location, there is no social interaction in that local labor market. We also estimate models including those locations and the results are very similar.

Table 2
Benchmark model at the tract level

Independent variable	Fixed effects regression with clustered standard errors			
	County fixed effects		Tract fixed effects	
	Coefficient	Std. error	Coefficient	Std. error
Constant	0.9903	0.1285	1.4917	0.0392
Single dummy	−0.2015	0.0046	−0.2003	0.0054
Female dummy	−0.2140	0.0276	−0.2106	0.0109
Single*female	0.2283	0.0169	0.2266	0.0062
Age	0.0508	0.0021	0.0506	0.0013
Age^2	−0.0005	0.0000	−0.0005	0.0000
White dummy	0.1805	0.0129	0.1815	0.0084
White*female	−0.1213	0.0034	−0.1230	0.0100
Student dummy	−0.0745	0.0057	−0.0718	0.0068
Veteran dummy	0.0600	0.0156	0.0600	0.0057
Disability dummy	−0.1480	0.0028	−0.1469	0.0101
College dummy	0.1429	0.0028	0.1441	0.0038
Graduate dummy	0.2807	0.0032	0.2825	0.0097
English proficiency dummy	−0.0480	0.0021	−0.0448	0.0057
Commuting time	0.0017	0.0001	0.0016	0.0001
Average education (%)	0.1535	0.0149		
Av. Edu. occupation (%)	0.0474	0.0183	0.0496	0.0191
Av. Edu. industry (%)	0.0399	0.0189	0.0251^{ns}	0.0167
Occ. specialization (%)	0.1697	0.0231	0.2025	0.0504
Ind. specialization (%)	0.0643	0.0103	0.0523	0.0178
Occ. diversity (%)	0.2986	0.1107		
Ind. diversity (%)	0.2116	0.0201		
Occ. employment density	0.0035^{ns}	0.0022	−0.0099^{ns}	0.0083
Ind. employment density	0.0374	0.0054	0.0419	0.0095
R^2	0.326		0.333	
SE clusters	county(7)		tract (621)	

Notes. 17 occupation and 10 industry dummies are included. Dependent variable: log(imputed hourly wage); total observations: 150,952.
^{ns} Insignificance at the 5% level.

The county fixed effects model shows that, within a county, all the types of labor market externalities (except occupation employment density) are significant at the 5% level. To identify what types of externalities are significant at the tract level, we drop *AveEdu* and diversity indices and estimate a tract fixed effects model. The results in Table 2 show that the quality of human capital in an occupation, Marshallian labor market externalities, and industry employment density are significant at the tract level.

To identify what types of human capital externalities are significant at lower geographical levels, we construct the labor market attribute indices at the blockgroup and block level, respectively. For example, in the block level model, the overall depth of human capital is the percentage of workers who are college graduates at a block; the occupation specialization index is the ratio of same-occupation workers at a block to the total number of workers at that block. We then estimate the benchmark model at the blockgroup and block level with different locational fixed effects, respectively. Tables 3 and 4 present the results for blockgroup and block level models, where only the labor market attribute indices are listed.

Table 3
Benchmark model at the blockgroup level

Fixed effects:	County		Tract		Blockgroup	
Variable	Coefficient	Std. error	Coefficient	Std. error	Coefficient	Std. error
AveEdu	0.1174	0.0157	0.0053[ns]	0.0278		
AveEduOcc	0.0222[ns]	0.0159	0.0252	0.0130	0.0245	0.0124
AveEduInd	0.0224	0.0076	0.0127[ns]	0.0121	0.0099[ns]	0.0121
OccSpec	0.1142	0.0428	0.1385	0.0333	0.1452	0.0312
IndSpec	0.0642	0.0086	0.0563	0.0146	0.0554	0.0160
OccDiver	0.1843	0.0442	0.2004	0.0669		
IndDiver	0.1390	0.0136	0.0264[ns]	0.0279		
OccDens	0.0062	0.0024	−0.0070[ns]	0.0070	−0.0089	0.0041
IndDens	0.0159	0.0039	0.0186	0.0040	0.0183	0.0059
R^2	0.326		0.333		0.345	
SE clusters	County (7)		Tract (620)		Blockgroup (2351)	

Notes. All indices are constructed at the blockgroup level. 17 occupation and 10 industry dummies are included.
[ns] Insignificance at the 5% level.

Table 4
Benchmark model at the block level

Fixed effects: Variable	County coefficient	Tract coefficient	Blockgroup coefficient	Block coefficient
AveEdu	0.0889	0.0414	0.0383	
	(0.0059)	(0.0160)	(0.0180)	
AveEduOcc	0.0114[ns]	0.0127[ns]	0.0125[ns]	0.0063[ns]
	(0.0101)	(0.0091)	(0.0091)	(0.0095)
AveEduInd	−0.0026[ns]	−0.0102[ns]	−0.0117[ns]	−0.0108[ns]
	(0.0057)	(0.0086)	(0.0092)	(0.0097)
OccSpec	0.0545	0.0643	0.0661	0.0891
	(0.0197)	(0.0194)	(0.0196)	(0.0206)
IndSpec	0.0670	0.0603	0.0614	0.0590
	(0.0068)	(0.0120)	(0.0124)	(0.0120)
OccDiver	0.1428	0.1462	0.1451	
	(0.0205)	(0.0319)	(0.0317)	
IndDiver	0.0860	0.0372	0.0442	
	(0.0174)	(0.0175)	(0.0190)	
OccDens	0.0064	0.0039	0.0036	0.0017[ns]
	(0.0001)	(0.0011)	(0.0016)	(0.0016)
IndDens	0.0026	0.0031	0.0028	0.0024
	(0.0010)	(0.0011)	(0.0012)	(0.0011)
R^2	0.330	0.337	0.349	0.385
SE clusters	County (7)	Tract (618)	Blockgroup (2247)	Block (7779)

Notes. Standard errors in parentheses. All indices are constructed at the block level. 17 occupation and 10 industry dummies are included.
[ns] Insignificance at the 5% level.

In Table 3, the blockgroup fixed effects model identifies the types of externalities that are significant within a blockgroup, including the quality of human capital among same-occupation peers, Marshallian labor market externalities, and the thickness of the labor market, in terms of the industry employment density.

In Table 4, the blockgroup fixed effects model shows that almost all types of human capital externalities are significant across blocks. The block fixed effects model further identifies that labor market thickness, in terms of industry employment density, and Marshallian labor market externalities, are significant at the block level.

Based on Tables 2–4, Table 5 further summarizes different types of human capital externalities that are significant at the 5% percent level at different geographic scopes.

The most striking result in Table 5 is that the Marshallian labor market externalities are significant at all microgeographic levels, including at the block level. Almost all types of externalities are significant across blocks, as well as across tracts. It is in this sense that we call cities, "Café Cities."

In Table 6, we assemble the results from the benchmark model at the tract, blockgroup, and block levels, where only county fixed effects are included.

Table 6 shows some interesting spatial patterns of human capital externalities. The coefficients of occupation employment density increase, when moved down to the lower geographic levels. This hints that they decay with distance away from a block. The coefficients of industrial Marshallian externalities are very similar, which hints that they decay very slowly within a tract. All other indices have the same pattern: much stronger at the tract level, much smaller at the block level. The explanation could be as follows: these externalities are strong at the tract level, but measurement errors attenuate the coefficients at the blockgroup and block levels. To test this hypothesis, we estimate the block level model by selecting only blocks where the number of workers is greater than 10. The results show that most of the coefficients, indeed, increase significantly. Tentatively, we conclude that, without measurement errors, the coefficients of human capital depth, occupation specialization, diversity indices, and industry employment density would be similar at the tract, blockgroup, and block levels.

The above rough spatial patterns indicate that different types of human capital externalities take place at different geographic scopes, and attenuate spatially at different speeds. However, we cannot see the pattern beyond the tract level. The natural extension is to test the spatial decay patterns of different types of human capital externalities within a larger geographic scope.

Table 5
Types of human capital externalities significant at different geographic levels

	Across tracts	Within tract	Across blockgroups	Within blockgroup	Across blocks	Within block
AveEdu	Y				Y	
AveEduOcc	Y	Y	Y	Y		
AveEduInd	Y					
OccSpec	Y	Y	Y	Y	Y	Y
IndSpec	Y	Y	Y	Y	Y	Y
OccDiver	Y		Y		Y	
IndDiver	Y				Y	
OccDens					Y	
IndDens	Y	Y	Y	Y	Y	Y
Fixed effects	County	Tract	Tract	Blockgroup	Blockgroup	Block

Note. "Y" indicates the coefficient is positive and significant at the 5% level.

Table 6
Benchmark model at different microgeographic levels

	Tract level $N > 1$ coefficient	Blockgroup level $N > 1$ coefficient	Block level $N > 1$ coefficient	Block level $N > 10$ coefficient
AveEdu	0.1535	0.1174	0.0889	0.1371
AveEduOcc	0.0474	0.0222^{ns}	0.0114^{ns}	0.0256
AveEduInd	0.0399	0.0224	−0.0026^{ns}	0.0186
OccSpec	0.1697	0.1142	0.0545	0.0880
IndSpec	0.0643	0.0642	0.0670	0.0612
OccDiver	0.2986	0.1843	0.1428	0.0246^{ns}
IndDiver	0.2116	0.1390	0.0860	0.1041
OccDens	0.0035^{ns}	0.0062	0.0064	0.0054
IndDens	0.0374	0.0159	0.0026	0.0023

Notes. All models include county fixed effects, occupation and industry dummies. N is the number of workers at a location.
^{ns} Insignificance at the 5% level.

6.2. Human capital externalities by occupation and by industry

We also estimate model (1) by occupation and by industry to explore the human capital externalities of a particular occupation or industry within a labor market at the tract level. Table 7 presents the results for a few selected occupations and industries.

Table 7 shows that high-tech industry workers benefit strongly from Marshallian labor market externalities. This is consistent with Henderson's finding based on plant level data. Manufacturing industry workers benefit from both Marshallian and Jacobs externalities, which is consistent with the literature on dynamic externalities. The same pattern holds for computer industry workers. Management occupation workers benefit from all of the four types of human capital externalities by differing degrees. Computer scientists benefit strongly from same-occupation peers and industrial diversity. Artists benefit strongly from urban diversity, which is very consistent with Florida's argument based on his bohemian index. In brief, channels of knowledge spillovers vary in sub-labor markets of different occupations and industries.

6.3. Spatial attenuation model results

We construct all the labor market attribute indices in such a way that they do not depend on geographic units. If economic activities were evenly distributed over space, and if there were no spatial attenuation, the effects of human capital externalities would be the same at different locations. If estimated coefficients for an index vary with distance, we can infer its spatial pattern. We draw rings of different miles away from the centroid of each block and construct indices based on the workers in each ring. There is no prior guidance on how to determine the number and the width of rings, except through experimentation. A rule of thumb is to look at the size distribution of blocks, blockgroups, and tracts. For all the tracts, blockgroups, and blocks in BMA, if we assume they were circles, then about 95% of blocks are within circles of 0.3 mile radius, 95% of blockgroups are within circles of 1.3 miles radius, and 98% of tracts are within circles of 3 miles radius. However, the size distributions of blocks, blockgroups, and tracts are very dispersed. We first, tentatively, draw rings of 1.5, 3, 6, and 9 miles away from each block centroid. Blocks themselves comprise the inner ring. Blocks within distance (0, 1.5] miles constitute the second

Table 7
Benchmark model by industry and by occupation

Industry	High-tech		Computer		Manufacturing	
	Coefficient	Std. error	Coefficient	Std. error	Coefficient	Std. error
AveEdu	0.3351	0.0890	0.3258	0.1663	0.2901	0.0562
AveEduOcc	−0.0289[ns]	0.0577	0.0786	0.0176	0.0270[ns]	0.0282
AveEduInd	0.0527[ns]	0.0469	0.0125[ns]	0.0860	0.0107[ns]	0.0652
OccSpec	0.3996	0.0914	0.0341[ns]	0.0933	0.1664[ns]	0.0870
IndSpec	0.1527	0.0161	0.5942	0.2181	0.2243	0.0555
OccDiver	0.1122[ns]	0.2924	0.6238[ns]	0.7001	0.2649[ns]	0.3615
IndDiver	0.1345	0.0474	0.2644	0.0999	0.1470	0.0463
OccDens	−0.0245[ns]	0.0654	−0.0418	0.0150	−0.0229	0.0094
IndDens	−0.0370[ns]	0.1104	−0.0642[ns]	0.1530	0.1019[ns]	0.0597
Observations	6251		2139		21,861	
R^2	0.449		0.312		0.404	
Occupation	Writers, artists, entertainers, athletes		Mathematical, computer scientists		Managerial, professional specialty	
	Coefficient	Std. error	Coefficient	Std. error	Coefficient	Std. error
AveEdu	0.4819	0.0832	0.0029[ns]	0.1338	0.0939[ns]	0.0779
AveEduOcc	0.0377[ns]	0.0601	0.1560	0.0606	0.0973	0.0143
AveEduInd	−0.0967[ns]	0.2279	0.1427[ns]	0.1221	0.0619	0.0258
OccSpec	−0.0826[ns]	0.3485	1.6416	0.2682	0.3491	0.0901
IndSpec	0.2933[ns]	0.1922	0.0385[ns]	0.0624	0.0907	0.0343
OccDiver	1.6508	0.3804	0.2300[ns]	0.3002	−0.0021[ns]	0.1950
IndDiver	0.4721	0.1708	0.2166	0.0400	0.2524	0.0673
OccDens	0.3441	0.0955	0.4985[ns]	0.3254	0.0086	0.0032
IndDens	0.0140[ns]	0.0232	−0.0704	0.0219	0.0381	0.0065
Observations	3400		2177		26,287	
R^2	0.172		0.223		0.292	

Note. Individual variables, occupation or industry dummies, and county fixed effects are included.
[ns] Insignificance at the 5% level.

ring, which corresponds approximately to the blockgroup level. Blocks within distance (1.5, 3] miles belong to the third ring, corresponding to the tract level; (3, 6] and (6, 9] miles annulus are the fourth and fifth rings. We estimate the spatial attenuation model including the indices for all the rings. The results are reported in Table 8.

The numbers 0, 1, 3, 6, 9 at the end of each variable name indicate that the construction of that variable is based on the employment within the block, rings of 1.5, 3, 6, and 9 miles respectively. Model 1 in Table 8 selects only blocks where the number of workers is greater than one. Let us first look at the overall quality of human capital. The variable *AveEdu* is geographic-invariant. If there were no spatial attenuation, the coefficients should be the same at different rings. The coefficients of *AveEdu* are actually 0.071, 0.141, 0.103, 0.021, and 0.066, respectively, from the inner ring to the 9 miles ring. The first two coefficients are significant at the 1% level, the third one is significant at the 10% level, but the other two coefficients are not significant. The coefficient at the block level (0.071) probably is underestimated due to measurement errors. Though we do not know how strong the actual effect is at the block level, we could infer that the effects of human capital depth are positive and significant up to three miles away from each block. The effects then decay rapidly thereafter (decreasing by 5 times from the third to the fourth ring). This pattern is also consistent with Table 6.

Table 8
Spatial attenuation model (5 rings)

Variable	Model 1: $N > 1$		Model 2: $N > 10$		Model 3: $N > 20$	
	Coefficient	Std. error	Coefficient	Std. error	Coefficient	Std. error
AveEdu0	0.0708	0.0120	0.1496	0.0067	0.1736	0.0076
AveEdu1	0.1406	0.0453	0.1081	0.0476	0.0667^{ns}	0.0556
AveEdu3	0.1034	0.0584	0.0621^{ns}	0.0740	0.1012^{ns}	0.0794
AveEdu6	0.0214^{ns}	0.0567	0.0603^{ns}	0.0733	0.0213^{ns}	0.0789
AveEdu9	0.0662^{ns}	0.0584	-0.0062^{ns}	0.0494	0.0140^{ns}	0.0686
OccSpec0	0.0535	0.0213	0.0874	0.0259	0.0816	0.0462
OccSpec1	0.2227	0.0850	0.1924	0.0835	0.2475	0.0844
OccSpec3	−0.3597	0.1912	−0.3837	0.2356	−0.4632	0.2463
OccSpec6	-0.0298^{ns}	0.1417	-0.0954^{ns}	0.1456	-0.0066^{ns}	0.0774
OccSpec9	0.2900^{ns}	0.2619	0.3862^{ns}	0.3138	0.3168^{ns}	0.2807
IndSpec0	0.0602	0.0075	0.0555	0.0061	0.0626	0.0096
IndSpec1	-0.0039^{ns}	0.0102	0.0036^{ns}	0.0087	-0.0111^{ns}	0.0126
IndSpec3	0.0948	0.0537	0.0989	0.0558	0.1028	0.0565
IndSpec6	0.1773	0.0715	0.1840	0.0967	0.2067	0.0871
IndSpec9	-0.0296^{ns}	0.0494	-0.0039^{ns}	0.0394	0.0179^{ns}	0.0448
OccDiver0	0.1438	0.0271	0.0384^{ns}	0.0496	-0.0063^{ns}	0.0368
OccDiver1	0.0859^{ns}	0.1484	0.2061^{ns}	0.1820	0.1313^{ns}	0.2246
OccDiver3	−0.3742	0.2288	-0.3393^{ns}	0.2898	-0.2665^{ns}	0.3125
OccDiver6	0.1223^{ns}	0.9250	0.0040^{ns}	0.9687	0.1841^{ns}	1.1019
OccDiver9	1.1618	0.4064	0.9068	0.3867	0.8317	0.4675
IndDiver0	0.0637	0.0162	0.0810	0.0244	0.0992	0.0192
IndDiver1	0.2735	0.0813	0.2944	0.0934	0.2896	0.1059
IndDiver3	0.2203	0.0906	0.2160	0.1051	0.2364	0.1175
IndDiver6	0.3004	0.1211	0.2755	0.0943	0.3075	0.1477
IndDiver9	0.1862^{ns}	0.01486	0.1205^{ns}	0.1123	0.1797^{ns}	0.1491
OccDens0	0.0064	0.0002	0.0053	0.0001	0.0051	0.0004
OccDens1	−0.0569	0.0065	−0.0510	0.0096	−0.0484	0.0106
OccDens3	0.0784	0.0301	0.1621	0.0290	0.1382	0.0268
OccDens6	-0.0973^{ns}	0.1672	-0.2069^{ns}	0.1497	−0.2116	0.1208
OccDesn9	-0.0918^{ns}	0.3432	-0.0091^{ns}	0.2312	0.1943^{ns}	0.2654
IndDens0	0.0020	0.0006	0.0021	0.0006	0.0018	0.0006
IndDens1	0.0675	0.0190	0.0542	0.0174	0.0411	0.0205
IndDens3	−0.2428	0.0567	−0.2747	0.0619	−0.2678	0.0739
IndDens6	-0.0105^{ns}	0.0432	0.0210^{ns}	0.0565	0.0278^{ns}	0.0761
IndDens9	0.0153^{ns}	0.1106	-0.0837^{ns}	0.0924	−0.2644	0.1203
Observations	145,230		123,211		107,847	
R^2	0.332		0.345		0.354	

Notes. N: number of workers in a block. Standard error clusters: 7 counties.
[ns] Insignificance at the 10% level.

The effect of occupation specialization decays very fast beyond 1.5 miles, which implies that the occupational Marshallian externalities are very localized. The industrial specialization effects are significant up to the sixth mile, and not significant thereafter, consistent with Table 6 where industrial specialization effects are very stable within the tract level.

The coefficients of occupation employment density are large and significant at the tract level, insignificant beyond tract level. The coefficients of industry employment density show a similar pattern: large and significant at the blockgroup level, then they decay rapidly.

The pattern of diversity indices is worth noting. The coefficients of occupation diversity are the strongest and significant at the farthest ring. The coefficients of industrial diversity are positive and significant at the first four rings, and the variation of magnitude is not very large (0.064, 0.274, 0.220, 0.300 in the first four rings). These patterns hint that there exist strong urbanization economies within certain geographic scopes, due to urban diversity.

The above results show that human capital externalities in cities have obvious spatial attenuation patterns, though they may be strong at different geographic scopes and decay at different speeds. Note that the monocentric city model predicts that spatial decay effects occur with distance away from the Central Business District (CBD). However, in our model, we construct rings for each block, no matter whether the block is located in downtown or in a suburban area. Therefore, our results provide much stronger and more powerful evidence for the spatial attenuation of local agglomeration economies.

Are our results sensitive to the scale of block employment concentration? We also estimate the same model using blocks where there are more than 10 workers and 20 workers, respectively (see models 2 and 3 in Table 8). Again we find that the coefficients at the block level (the inner ring) increase significantly, probably due to measurement errors. Though the magnitude of the coefficients changed, the spatial decay patterns are almost the same.

We also estimate the spatial attenuation model with tract and blockgroup fixed effects. The results are presented in the appendix Table A.6. After controlling for tract or blockgroup specific effects, most coefficients are positive and significant only at the first and/or the second ring (block and/or blockgroup level), which makes the spatial attenuation pattern less neat. This, however, further confirms our identification strategy that the tract fixed effects model identifies agglomeration effects across blockgroups, and the blockgroup fixed effects model identifies agglomeration effects across blocks.

Are our results sensitive to the number and width of rings? We also estimate the spatial decay model by constructing 5 rings of 2 miles intervals and 6 rings of 1 or 2 miles intervals. The results (not reported here) are similar. We also estimate the spatial attenuation model by industry and occupation. The results (not reported here) are mixed since human capital externalities vary across occupations and industries.

7. Conclusion

Endogenous growth and urban economic theories assume the existence of human capital externalities. Urban theoretical models further predict that agglomeration forces attenuate spatially. In this paper, we use the 1990 Massachusetts census data and provide empirical evidence for the microfoundations and spatial attenuation patterns of knowledge spillovers at the microgeographic levels. We test four channels through which individual workers learn from their occupational and industrial peers in the same local labor market: depth of the human capital stock, Marshallian labor market externalities, Jacobs labor market externalities, and thickness of the local labor market. We find that all types of human capital externalities are significant across census tracts and blocks; Marshallian labor market externalities and the effect of labor market thickness, in terms of industry employment density, are significant at the block level. The mechanisms of knowledge spillovers vary with industries and occupations. Different types of externalities decay at different speeds over geographic distances. We conclude that knowledge spillovers are very localized within a microgeographic scope in cities that we call, "Smart Café Cities."

Some related questions may warrant further research. We estimate only the marginal effects of human capital externalities on individual earnings. It would be interesting to estimate the aggregate benefit and its attenuation for each worker by using a location potential function, such as Fujita and Ogawa [15] specified in their simulation models. Also, we estimate only a cross-section model, though we adopt the concepts of Marshallian and Jacobs externalities of labor markets. The next stage would be to test the dynamics of human capital externalities. Simon and Nardinelli [47] found considerable persistence in the effects of human capital, indicating that the distribution of human capital established in the first decade of the twentieth century played a role in the current status of American cities. They also found that the presence of human capital is less important today than in the past, perhaps reflecting the decline in the costs of transportation and communication. However, Gaspar and Glaeser [17] argued that information technology and face-to-face interactions could be complements rather than substitutes. Empirical testing of the effects of information technology on knowledge spillovers would be a very interesting topic in the near future. Finally, but not least important, since human capital externalities can be capitalized into housing values (as mentioned in the introduction section), our next research topic is to estimate a hedonic housing model, including local labor market attributes, in order to test Lucas' conjecture.[18]

Acknowledgments

I am very grateful to Richard Arnott, Stephen Ross, Marvin Kraus, Jan Brueckner, two anonymous referees, and the participants of the 18th AREUEA annual doctoral session for their very helpful comments. I also thank Jim Davis for all his help while I was working at the Boston Census Research Data Center. Financial support from the Lincoln Institute of Land Policy Dissertation Fellowship and Boston College H. Michael Mann Fellowship are gratefully acknowledged.

The research in this paper was conducted while the author was a Special Sworn Status researcher of the US Census Bureau at the Boston Census Research Data Center. Research results and conclusions expressed are those of the author and do not necessarily reflect the views of the Census Bureau. This paper has been screened to insure that no confidential data are revealed.

Appendix A

Table A.1
Some summary statistics

	Mean	Std. deviation	Min.	Max.
Number of workers in a tract	242	485		8527
Number of workers in a blockgroup	61	151		2664
Number of workers in a block	13	40		1003
Number of blockgroups in a tract	4.0	1.7	1	9
Number of blocks in a blockgroup	4.6	3.3	1	25
Mean hourly wage in a tract	14.49	3.4		
Mean hourly wage in a blockgroup	14.54	7.14		
Mean hourly wage in a block	14.42	10.42		

[18] Lucas [30, pp. 270–271] wrote "... an externality-based theory of cities might let us use the information contained in urban land price gradients to estimate the size of the externalities associated with human capital accumulation."

Table A.2
Industry code and employment share

Industry	1990 Census code	Employment share (%)
Construction	60–99	4.5
Manufacturing	100–399	14.5
Public utility	400–499	6.6
Wholesale trade	500–579	4.1
Retail trade	580–699	15.1
Finance, real estate, insurance	700–720	10.2
Business and repair services	721–760	5.0
Personal services	761–799	2.5
Entertainment	800–811	1.1
Professional services	812–899	31.5
Public administration	900–939	4.9
Computer	732–739	
High-tech industry*	181, 321–330, 352–359 362–369, 371–380	

* The classification of high-tech industry is based on Maggioni [31].

Table A.3
Occupation code and employment share

Occupation	1990 Census code	Employment share (%)
Managerial, professional specialty	1–42	17.5
Engineers, architects, surveyors	43–63	2.5
Mathematical, computer scientists	64–68	1.4
Natural scientists	69–83	0.6
Health diagnosing occupation	84–112	4.1
Teachers, librarians, archivists	113–165	6.0
Social scientists, urban planners	166–182	2.8
Writers, artists, entertainers, athletes	183–202	2.2
Technicians	203–242	4.8
Sales	243–302	11.2
Administrative	303–402	19.2
Service	403–472	11.8
Mechanics, repairers	503–552	2.3
Construction	553–612	2.8
Precision production	628–702	2.5
Machine operators, tenders	703–802	3.6
Transportation, material moving	803–863	2.3
Handlers, equipment cleaners, laborers	864–902	2.4

Table A.4
Summary statistics for selected industries and occupations

	Index	Mean	Std. dev.
Managerial, professional specialty	*AveEduOcc*	0.601	0.195
	OccSpec	0.153	0.060
	OccDens	0.039	0.138
Writers, artists, entertainers, athletes	*AveEduOcc*	0.673	0.326
	OccSpec	0.027	0.022
	OccDens	0.008	0.023
Mathematical, computer scientists	*AveEduOcc*	0.817	0.284
	OccSpec	0.017	0.017
	OccDens	0.006	0.018
Manufacturing	*AveEduInd*	0.372	0.247
	IndSpec	0.136	0.120
	IndDens	0.022	0.051
Finance, insurance, real estate	*AveEduInd*	0.438	0.260
	IndSpec	0.080	0.067
	IndDens	0.025	0.135
Retail	*AveEduInd*	0.253	0.168
	IndSpec	0.174	0.091
	IndDens	0.027	0.058
Overall	*AveEdu*	0.419	0.126
	OccDiver	0.864	0.057
	IndDiver	0.759	0.105

Table A.5
Correlation matrix

	AveEdu	*AveEduOcc*	*AveEduInd*	*OccSpec*	*IndSpec*	*OccDiver*	*IndDiver*	*OccDens*	*IndDens*
AveEdu	1.000								
AveEduOcc	0.386	1.000							
AveEduInd	0.563	0.519	1.000						
OccSpec	0.044	−0.018	0.049	1.000					
IndSpec	0.164	0.241	0.390	0.118	1.000				
OccDiver	−0.161	−0.062	−0.090	−0.271	−0.088	1.000			
IndDiver	−0.289	−0.112	−0.163	−0.042	−0.566	0.156	1.000		
OccDens	0.360	0.142	0.223	0.376	0.002	−0.313	0.041	1.000	
IndDens	0.432	0.247	0.399	0.124	0.333	−0.291	−0.175	0.670	1.000

Table A.6
Spatial attenuation models with different fixed effects

Fixed effects	Tract		Blockgroup	
	Coefficient	Std. error	Coefficient	Std. error
AveEdu0	0.0392	0.0152	0.0296	0.0172
AveEdu1	−0.1278ns	0.0868	−0.3624	0.1396
AveEdu3	0.0637ns	0.1266	−0.0449ns	0.1573
AveEdu6	−0.0725ns	0.1965	−0.0221ns	0.3237
AveEdu9	−0.0511ns	0.2193	0.0878ns	0.3345
OccSpec0	0.0578	0.0196	0.0603	0.0201
OccSpec1	0.2207	0.0792	0.2003	0.0753
OccSpec3	−0.3491	0.1385	−0.3519	0.1138
OccSpec6	−0.0915ns	0.1768	−0.1265ns	0.1647
OccSpec9	0.3527	0.1610	0.3688	0.1600
IndSpec0	0.0536	0.0124	0.0565	0.0127
IndSpec1	−0.0029ns	0.0253	−0.0172ns	0.0268
IndSpec3	0.0826	0.0430	0.0871	0.0396
IndSpec6	0.1139	0.0632	0.1493	0.0612
IndSpec9	0.0400ns	0.0699	0.0140ns	0.0675
OccDiver0	0.1351	0.0317	0.1398	0.0319
OccDiver1	−0.5065	0.2589	0.2214ns	0.3480
OccDiver3	−0.7117	0.3110	−0.4354ns	0.6218
OccDiver6	0.1882ns	1.215	1.0808ns	1.7819
OccDiver9	0.6878ns	1.3032	1.7476ns	1.7778
IndDiver0	0.0325	0.0175	0.0416	0.0192
IndDiver1	0.0616ns	0.1057	0.1576ns	0.1400
IndDiver3	0.1256ns	0.1635	0.0637ns	0.2294
IndDiver6	−0.2208ns	0.4160	−0.6970ns	0.5972
IndDiver9	0.3289ns	0.4221	0.6386ns	0.6021
OccDens0	0.0046	0.0014	0.0043	0.0017
OccDens1	−0.0500ns	0.0344	−0.0452ns	0.0334
OccDens3	0.0924ns	0.1150	0.0643ns	0.0997
OccDens6	0.0118ns	0.2232	0.0422ns	0.1988
OccDens9	−0.2439ns	0.2793	−0.1506ns	0.2800
IndDens0	0.0025	0.0009	0.0022	0.0010
IndDens1	0.0790	0.0276	0.0738	0.0289
IndDens3	−0.1935	0.0740	−0.2129	0.0700
IndDens6	0.0690ns	0.1504	0.1151ns	0.1334
IndDens9	−0.1695ns	0.2179	−0.1328ns	0.2114
R^2	0.338		0.350	
SE clusters	Tract (610)		Blockgroup (2226)	

Note. Observations: 142,530.
ns Insignificance at the 10% level.

References

[1] L. Anselin, Spatial Econometrics: Methods and Models, Kluwer Academic, Boston, MA, 1988.
[2] R. Barro, X. Sala-I-Martin, Economic Growth, McGraw-Hill, 1995.
[3] P. Bayer, S. Ross, G. Topa, Place of work and place of residence: Informal hiring networks and labor market outcomes, Working paper 11019, NBER, 2005.
[4] P. Beeson, D. Delong, W. Troesken, Population growth in US counties, 1840–1990, Regional Science and Urban Economics 31 (2001) 669–699.
[5] M. Berliant, R. Reed, P. Wang, Knowledge exchange, matching, and agglomeration, Working paper, Department of Economics, Washington University, 2002.

[6] D. Black, V. Henderson, A theory of urban growth, Journal of Political Economy 107 (1999) 252–284.
[7] S. Charlot, G. Duranton, Communication externalities in cities, Journal of Urban Economics 56 (2004) 581–613.
[8] A. Ciccone, R. Hall, Productivity and the density of economic activity, American Economic Review 86 (1996) 55–70.
[9] G. Duranton, H. Overman, Testing for localization using microgeographic data, Working paper No. 3379, CEPR, 2002.
[10] G. Duranton, D. Puga, Diversity and specialization in cities: Why, where and when does it matter? Urban Studies 37 (2000) 533–555.
[11] G. Duranton, D. Puga, Microfoundations of urban agglomeration economies, in: V. Henderson (Ed.), Handbook of Regional and Urban Economics, vol. 4, Elsevier, 2005, pp. 2063–2117.
[12] J. Eaton, Z. Eckstein, Cities and growth: Theory and evidence from France and Japan, Regional Science and Urban Economics 27 (1997) 443–474.
[13] R. Florida, The economic geography of talent, The Annals of the Association of American Geographers 92 (2002) 743–755.
[14] R. Florida, Bohemia and economic geography, Journal of Economic Geography 2 (2002) 55–71.
[15] M. Fujita, H. Ogawa, Multiple equilibria and structural transition of non-monocentric urban configurations, Regional Sciences and Urban Economics 12 (1982) 161–196.
[16] M. Fujita, P. Krugman, A. Venables, The Spatial Economy, The MIT Press, 1999.
[17] J. Gaspar, E. Glaeser, Information technology and the future of cities, Journal of Urban Economics 43 (1998) 136–156.
[18] S. Gibbons, S. Machin, Valuing English primary schools, Journal of Urban Economics 53 (2003) 197–219.
[19] E. Glaeser, D. Maré, Cities and skills, Journal of Labor Economics 19 (2001) 316–342.
[20] E. Glaeser, H. Kallal, J. Scheinkman, A. Shleifer, Growth in cities, Journal of Political Economy 100 (1992) 1126–1152.
[21] R. Helsley, W. Strange, Matching and agglomeration economies in a system of cities, Regional Science and Urban Economics 20 (1990) 189–212.
[22] V. Henderson, The sizes and types of cities, American Economic Review 64 (1974) 640–656.
[23] V. Henderson, Marshall's scale economies, Journal of Urban Economics 53 (2003) 1–28.
[24] J. Jacobs, The Death and Life of Great American Cities, Vintage Books, New York, 1961.
[25] J. Jacobs, The Economy of Cities, Vintage Books, New York, 1969.
[26] A. Jaffe, M. Trajtenberg, R. Henderson, Geographic localization of knowledge spillovers as evidenced by patent citations, Quarterly Journal of Economics 108 (1993) 577–598.
[27] B. Jovanovic, R. Rob, The growth and diffusion of knowledge, Review of Economic Studies 56 (1989) 569–582.
[28] E. Leamer, Effort, wage, and the international division of labor, Journal of Political Economy 107 (1999) 1127–1162.
[29] R.E. Lucas Jr., On the mechanics of economic development, Journal of Monetary Economics 22 (1988) 3–42.
[30] R.E. Lucas Jr., Externalities and cities, Review of Economic Dynamics 4 (2001) 245–274.
[31] M. Maggioni, Clustering Dynamics and the Location of High-Tech Firms, Physica-Verlag Heidelberg, New York, 2002.
[32] A. Marshall, Principles of Economics, eighth ed., Macmillan & Co., London, 1920.
[33] E. Morettie, Human capital externalities in cities, Working paper 9641, NBER, 2003.
[34] T. Palivos, P. Wang, Spatial agglomeration and endogenous growth, Regional Science and Urban Economics 26 (1996) 645–669.
[35] R. Petitte, S. Ross, Commutes, Neighborhood effects, and compensating differentials: Revisited, Journal of Urban Economics 46 (1999) 1–26.
[36] M. Porter, The Competitive Advantage of Nations, Free Press, New York, 1990.
[37] M. Porter, Clusters and the new economics of competition, Harvard Business Review, Nov.–Dec. (1998) 77–90.
[38] J. Quigley, Urban diversity and economic growth, Journal of Economic Perspectives 12 (1998) 127–138.
[39] J. Rauch, Productivity gains from geographic concentration of human capital: Evidence from cities, Journal of Urban Economics 34 (1993) 380–400.
[40] J. Roback, Wages, rents, and the quality of life, Journal of Political Economy 90 (1982) 1257–1278.
[41] P. Romer, Increasing returns and long-run growth, Journal of Political Economy 94 (1986) 1002–1037.
[42] S. Rosenthal, W. Strange, The attenuation of the agglomeration economies: A Manhattan skyline approach, Working paper, Department of Economics, Syracuse University, 2005.
[43] S. Rosenthal, W. Strange, The determinants of agglomeration, Journal of Urban Economics 50 (2001) 191–229.
[44] S. Rosenthal, W. Strange, Geography, industrial organization, and agglomeration, Review of Economics and Statistics 85 (2003) 377–393.

[45] S. Rosenthal, W. Strange, Evidence on the nature and sources of agglomeration economies, in: V. Henderson (Ed.), Handbook of Regional and Urban Economics, vol. 4, Elsevier, 2005, pp. 2119–2171.
[46] J. Shapiro, Smart cities: Quality of life, productivity, and the growth effects of human capital, Working Paper 11615, NBER, 2005.
[47] C. Simon, C. Nardinelli, Human capital and the rise of American cities: 1900–1990, Regional Science and Urban Economics 32 (2002) 59–96.
[48] D. Timothy, W. Wheaton, Intra-urban wage variation, employment location, and commuting times, Journal of Urban Economics 50 (2001) 338–366.
[49] W. Wheaton, M. Lewis, Urban wages and labor market agglomeration, Journal of Urban Economics 51 (2002) 542–562.