

Larsson, Johan P

Conference Paper

Disaggregating Density A Knowledge Spillover Analysis Using Geo-Coded Data

52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Larsson, Johan P (2012) : Disaggregating Density A Knowledge Spillover Analysis Using Geo-Coded Data, 52nd Congress of the European Regional Science Association: "Regions in Motion - Breaking the Path", 21-25 August 2012, Bratislava, Slovakia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/120724>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Neighborhood Effects on Productivity

A Knowledge Spillover Analysis Using Geocoded Data

Johan P. Larsson^Ø

Abstract

This paper analyses the effects of neighborhood density on worker productivity. Because of the attenuation of human capital externalities and localized clustering, there is theoretical scope for examining this *neighborhood effect*, as distinct from region-wide external effects. Using a geocoded, publicly audited, full population dataset on employment and wages in Sweden's city areas, the analysis is based on 250 meter (about 0.15 miles), 1 000 meter, and 10 000 meter squares, representing the neighborhood level. Individual-level wage regressions confirm that wages are positively affected by proximity to economic activity on the neighborhood level, but that the results are dependent on spatial resolution, with the elasticity being the highest in the smaller squares. Controlling for economic density (market potential) at the level of urban regions decreases the elasticities in the larger squares, suggesting that the initial results are largely driven by region-wide effects. The coefficients associated with the 250 meter neighborhood squares are robust to such controls. Here, the elasticity is still in the 2 percent range when controlling for spillovers at the regional level. Further, IV estimations using geological variables indicate that the results are not driven by reverse causality. The results are consistent with i) the existence of a localized neighborhood effect and ii) quite sharp attenuation of human capital spillovers.

JEL:

Keywords: density, productivity, spatial dependence, geo-coded data, human capital, agglomeration economies

^Ø Center for Entrepreneurship and Spatial Economics (CEnSE), Jönköping International Business School. E-mail: johan.p.larsson@jibs.hj.se

Acknowledgements: This research has been supported by a financial grant from the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS).

1. INTRODUCTION

This paper provides a conceptual discussion and empirical estimations of the magnitude of the *neighborhood effect* in spatial economics. Numerous approaches to understanding the relationship between worker productivity and the density of economic activity have taken cities or regions as starting points. Lucas (1988) notes that nation states are large, arbitrary entities and - drawing on Jacobs (1969) - suggests that cities may be a better basis for an analysis of knowledge spillovers. Research in this area has to a large extent been driven by data availability. The proposition that cities and regions are heterogeneous is not controversial. In fact, many authors have, at least conceptually, dealt with districts (e.g. Lucas, 1988). Jacobs (1961), who is often cited in studies of urbanization economies, notes how different parts of the same city often fare differently at the same point in time. In this paper, I propose a division of Alfred Marshall's (1920) forces of agglomeration - knowledge spillovers, input sharing, and labor market pooling - into two levels: one representing a region-wide effect, and one representing a neighborhood effect, to a high extent driven by knowledge spillovers. The extent and existence of the neighborhood effect is, at present, an empirical issue.

The question addressed is to what extent the economic density of a work place's immediate surroundings influences the productivity of workers - an effect that is potentially obscured by high levels of aggregation. Marshall's (1920) forces of agglomeration still broadly account for why firms and individuals choose to cluster to each other, but there are a multitude of reasons to believe that these effects are not equally concentrated in space. Specifically, knowledge spillovers are highly localized in space; since knowledge is not effortlessly codified, copied and transmitted, but driven by face to face interaction, it is subject to significant distance decay (Å. E. Andersson & Beckmann, 2009; Rosenthal & Strange, 2008). This effect makes geographically large observation units problematic when addressing the issue.

Ciccone and Hall (1996, p. 54) define economic density as "*the intensity of labor, human capital and physical capital relative to physical space*". The definition is reasonable, but applied at a high level of aggregation, such as states, counties, or metropolitan statistical areas, it produces only a crude indication of the actual density incurred to any single establishment or individual. Measures of economic density are not intrinsically useful. The idea behind the use of economic density in regional economics is to proxy the potential for productive interactions, and such interactions are not uniformly spread out across space.

A common practice in regional economics is to assume that workers in a city, region or metropolitan area are uniformly affected, e.g. by overall level of economic activity or average levels of schooling in the area under observation. But bearing in mind that knowledge spillovers are driven by direct

interaction and that search costs increase quickly as an area is enlarged¹, this assumption appears strong. A worker in a central business district is faced with substantially different opportunities compared to a worker in a suburb, both in terms of density and in terms of learning opportunities. In fact, if opportunities and accessibility were uniform across space within a city, then a similar pattern should be observed in land prices. The non-uniform prevalence of human capital externalities is undoubtedly one reason that they are not. Hence, this article lines up with the approaches to regional economics that deal with the implications of face-to-face meetings and ‘talk’ in the economy, up to and including direct persuasion (see e.g. McCloskey & Klammer, 1995).

A large part of the literature has dealt with the so-called urbanization versus localization issue. The dichotomy is not necessarily true: there is an observed tendency for similar industries to cluster to each other *within a larger agglomeration of firms*, e.g. financial districts located at the core of many large metropolitan areas. An agglomerated region may in this sense be seen as the sum of many neighborhoods, each of which may exhibit a large degree of localization. Each such localization can sensibly be thought of as a specialized part of a diversified city or region. Capello (2002) elaborates on these issues in some detail and provides a useful taxonomy for how to think about the distinction.

Localized knowledge spillovers as well as possible sub-city specialization are both arguments why an analysis is warranted at the very disaggregated level. The smallest observational units in the United States that have been used for a similar purpose are the American Community Survey’s PUMS areas (PUMAs). Gordon and Ikeda (2011) analyze the densities of these areas with some very promising results. Even though the smallest of the PUMAs are approaching neighborhood size - the New York City PUMAs are typically 7-12 square kilometers - their mean size remains quite large. In this paper, external economies are assessed at an even finer level of disaggregation. The analysis in this paper is based on a full-population, geo-coded, publicly audited dataset, pinpointing each individual worker in Sweden’s private workforce to a square, representing the neighborhood where he or she is working. Using individual-level wage regressions, the density-productivity relationship is then assessed at the neighborhood level.

The fact that density, as defined above, is a quotient suggests that an area may be of high density either because it is small or because it is rich in economic activity. The structure of the data restricts changes in density to the changes in intra-neighborhood mass. The uniform-sized grid implies that any difference in values of some characteristic between two squares equals the difference in density with respect to that characteristic. The contribution of this paper is firstly to assess the productivity-density relationship at a very low level of disaggregation, observing areal units of exactly the same size. Second, the relationship is for different levels of aggregation to test the hypothesis of localized

¹ Search costs increase rapidly with enlargement, since the size of the area under observation is equal to the square of the distance to the point or origin (as noted in Andersson & Beckmann, 2009)

spillover effects. The regressions indicate that there are productivity spillovers within neighborhoods, as well as within regions, but also that these vary with chosen spatial resolution. Small neighborhood squares capture an effect robust to considerations of the surrounding region, while the effect picked up by the larger squares to a larger extent proxies a large region.

2. THE NATURE AND SCOPE OF AGGLOMERATION ECONOMIES

Firms will tend to agglomerate when the linkages in the market lower the costs or raise the revenue of the individual firms, i.e. where external economies are present. A distinction is commonly made between localization economies, stemming from the benefits of locating firms in similar industries in the same place and urbanization economies, drawing on the benefits of a large market in general. (Dicken & Lloyd, 1990)

The spillover strand of the literature begins with Marshall (1920), who argues that short proximity to other workers increases learning and renders workers more productive, but also that firms benefit from labor market pooling and input sharing. An important implication of the non-rival nature of knowledge is noted in Romer (1990); if an input is characterized by non-rivalry and has productive value, then constant-returns-to-scale is not a necessity. The production function does not have to be homogenous of degree one if all inputs in the production chain do not require replication.

In general, the forces of agglomeration are appropriately analyzed in a regional context. E.g. viewing a region as an integrated labor market (or part thereof), productivity is enhanced in thicker markets by matching of specialized employees with suitable employers (Helsley & Strange, 1990). Willingness to commute within a region is generally high, and poses no major problems for this reasoning in terms of attenuation of the effect within the region, while commuting outside of the region is subject to substantial distance decay (Johansson, Klaesson, & Olsson, 2003). While parts of the mechanisms behind localized external economies are intrinsically regional (or city) phenomena, and should be modeled as such, human capital externalities are of a different kind altogether.

The attenuation of human capital spillovers is increasingly treated as something of a stylized fact in the literature. An early finding to support this claim is found in Rosenthal and Strange (2001); their proxy for knowledge spillovers is positive and significant at the zip code level, but not at the state and county levels. They conclude that knowledge spillovers contribute to the forces of agglomeration at the local level, while noting that their results should be interpreted with caution. More recently, the effect has been documented more sophisticatedly, with the help of concentric rings around workers in Rosenthal and Strange (2008). In their study, the smallest rings extend 5 miles from the place of employment, and they find a substantial attenuation of the effect outside of the first distance band.

Jovanovic and Rob (1989) model the transmission of knowledge as an interactive process that, unlike an invention, is unique in the sense that its diffusion is inseparable from its creation. They model the diffusion of knowledge through agent meetings and propose that the higher the level of diversity of agents, the more beneficial meetings. Diversity is itself tightly linked with market size (M. Andersson & Klaesson, 2009). Similar thoughts have been prevalent in sociology; Granovetter (1973) argues that interpersonal connections is the bridge between micro and macro phenomena.

Common to the arguments outlined above is the notion of a spatially delimited effect that by any standard must be considered smaller than a region in the normal sense of the word. Because of the importance of face-to-face meetings, learning by interacting with other workers, and encounters of the serendipitous kind, there is plenty of theoretical scope to investigate the effects of neighborhoods on productivity directly.

Glaeser (1999) formalizes Marshallian knowledge spillovers and models learning as a function of agent meetings, which increases with the number of people working in the central business district (CBD) and with city density. Agents in the model commute to a CBD from their home, in order to take part of the higher rate of learning present at this particular location; the agents also have the option of settling in a hinterland with fixed wages and no learning from neighbors. Hence the wage formation may be thought of as based upon within-city heterogeneities. The key proposition is that the degree of learning is not uniform across space within cities.

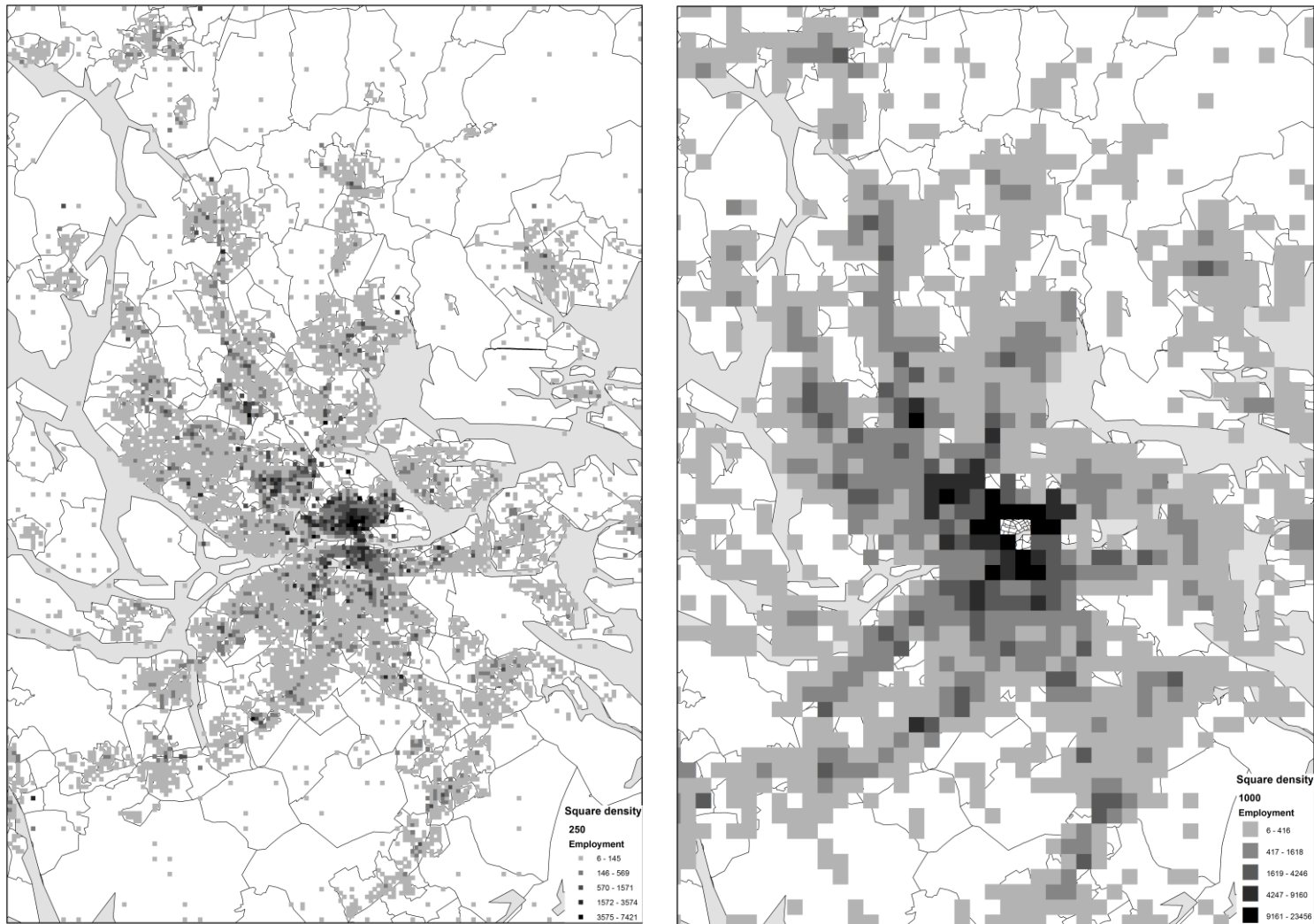
Using individual-level wage regressions, the theoretical underpinnings for this analysis rests on Roback (1982); it is assumed that workers, on average, are paid the value of their marginal product and that since firms in denser areas must pay higher nominal wages to equate consumption opportunities over space, they must be more productive than in less dense areas.

Following Lucas (1988) it is assumed that one worker with productive capacity $h(t)$ is the productive equivalent of two workers with productive capacity $\frac{1}{2}h(t)$. In this view, the number of employees in a given geographical area, one traditional way of measuring density, will not necessarily reflect the extent of economic activity in that area to the extent that workers are equipped with different levels of human capital. Thus, in the empirical part, economic density is measured as *quality adjusted employment*, and represented by aggregated wage sums in each square, rather than by number of employees. This operation also minimizes a problem with measuring density given the data source, namely that number of hours worked is a missing variable. This point is elaborated upon below. Most of the variance between squares is nevertheless driven by employment density and using number of employees as density measure does not upset the results presented in the empirical part of this paper.

Even though this analysis is mainly focused on localized agglomeration economies, the analysis incorporates ‘localized’ localization economies in the empirical section, by considering the possibility of localization economies within neighborhoods.

A remaining question is how to determine the size of the observational units. The extent of a possible neighborhood effect is not easily determined based on theory. In the absence of empirical approaches at the very disaggregated levels, the spatial reach must at this point be regarded as an empirical question. In this paper three different sizes are proposed: squares with sides of 250 meters, 1 000 meters and 10 000 meters are used to trace the effect. The smallest squares are the size of a few blocks, while the largest squares approximately resemble the size of a small urban region in terms of square kilometers. Figure 1 shows employment density in the Stockholm area for 250 and 1 000 meter squares.

Figure 1. The larger Stockholm area. Private workforce employment density on the neighborhood level, using 250-by-250 meter squares (left) and 1000-by-1000 meter squares (right).



3. DATA, MODEL AND VARIABLES

The data source is a publicly audited, matched employer-employee dataset, maintained by Statistics Sweden and covering the period from 1991 to 2008. Employees are pinned down geographically to a 250 by 250 meter square via their work place within the grid. By matching each individual worker to a square via the employer, wage sums for each square are aggregated. These measures of wage density are then matched back to each individual worker. Since all squares are of equal size, no further normalization is needed to get an exact measure of density. This feature makes interpretation straightforward and eliminates the problem that some geographically large regions contain only small areas of spatially concentrated activity.

After selecting wage workers 20-64 years old, removing publicly employed workers and primary industry workers, the remaining population is about 1.5 million individuals per year, although increasing in later periods. The data source has one clear weakness. Even though the selected individuals are primarily wage workers, number of hours worked is a missing variable. Even though most of the issues are resolved by the within estimation technique, problems may still arise if the number of individuals that e.g. are moving from part-time to full-time work are disproportionately allocated to dense neighborhoods. To eliminate outliers, workers that spend excessively much (or little) time working, and miss-reported observations, the individuals with wages in the 5 percent upper and lower tails are excluded from the estimations. Empirically, cutting the tails decreases the neighborhood elasticities somewhat, but the estimates presented below all exhibit strong robustness to exclusion of further individuals.

The applied model is an augmented version of the wage equation in Mincer (1974), wherein logarithmic earnings are modeled as a function of schooling and experience. An important issue in this regard is sorting of workers of unobserved ability across space. Since workers of higher unobserved ability tend to self-select to dense areas (see e.g. Combes, Duranton, & Gobillon, 2008) OLS estimates are likely to suffer from an omitted variable problem and are likely to be biased upwards. Two features of the data are important to note at this point. First, the panel follows the same individuals over time and, second, the density areas exhibit substantial within-variation over time, which is not always true for larger regions. Hence, the data is well suited for within estimations in order to account for time-invariant unobserved worker characteristics without losing the effects of economic density. Conclusively, the Hausman test strongly favors fixed effects.

Indexing workers by i , firms by r , neighborhoods by j , and urban regions by k , the model employed is:

(1)

where w_{it} is the wage of individual i in year t , \mathbf{X} is a matrix of neighborhood j characteristics, \mathbf{Y} is a matrix of region k economic density, \mathbf{Z} is a matrix of individual i control variables, and \mathbf{F} is a matrix of firm r characteristics. The individual-specific error term ϵ_{it} contains unobserved, time-invariant, worker characteristics (e.g. ability) from the fixed effects estimations and is seen alongside the normally distributed error term η_{it} . Table 1 outlines the details for all variables used on the neighborhood, regional, individual, and firm levels.

Table 1. Description of right hand side variables, separated by level of aggregation.

Neighborhood level variables (\mathbf{X})		Exp. sign
Density	Wage sums pertaining to the square where an individual is working.	+
Localization	Share of wage sums generated in the industry where the individual is working ² .	+
Regional level variables (\mathbf{Y})		
Density (accessibility)	Distance decay weighted accessibility to local, inter-regional, and intra-regional economic activity. Used to control for spatial autocorrelation in the second part of the analysis.	+
Individual level variables (\mathbf{Z})		
Schooling	Theoretical number of years of the completed level of education.	+
Experience squared	Number of years working after graduation. Technically age less years of schooling less 6.	+
Experience ²	The square of the experience variable.	-
Tenure	Number of years that the worker has been employed by the same company, since 1991.	+/-
Firm level variables (\mathbf{F})		
Firm size	Number of workers employed by the company.	+
Average years of schooling	Average length of education at the firm level.	+
Industry	Dummy variable on the 2 digit SIC level.	

The neighborhood level variables are assessed for 250, 1 000 and 10 000 meter squares. The density variable tracks economic activity over time by recording aggregated wage sums in the squares. Density is the main variable of interest; it measures the extent of economic production at the *neighborhood* level and proxies the potential for productive interaction at the very local level. As the

² The industries are manufacturing, knowledge intensive business services, and other services.

squares become larger, the hypothetical ‘neighborhoods’ increase in size; even though they are all referred to as neighborhoods for consistency reasons, the larger squares are approximately the size of a small urban region. The industry concentration variable is included to control for localization economies within squares, i.e. benefits within clusters of industries. This variable reflects the degree of neighborhood specific localization, depicting the potential for neighborhoods to be specialized within the confinements of larger economic agglomerations. This variable pushes the density variable closer to being an indicator of localized agglomeration economies. The male share variable is included to remedy some of the problems with the inability of distinguishing part-time from full-time work.

The regional density, or accessibility, variable is a time travel distance weighted measure of market potential in each of Sweden’s 290 urban regions (municipalities). The weighting alleviates some of the problems with some regions being geographically large, but containing only small spatial concentrations of economic activity. The measure represents a step towards a continuous view of geography, where effects from surrounding regions are taken into account, and where activities across the geography matter, but matter less the further away they are. The accessibility measure is elaborated upon in some detail under the empirical section. The variable reflects the potential for productive interaction at the *regional* level. A large accessibility variable reflects a thick market, with the corresponding benefits in terms of labor market pooling and input sharing.

The schooling and experience variables are standard in the literature and are both expected to be positive since they indicate accumulation of human capital. The increasing effect of experience is expected to attenuate with time, as is represented by the experience squared term. The tenure variable is measuring the number of consecutive years that each individual has spent with the same employer and is an indication of the quality of the match between the two parties³. The idea is that in case of a good match (thereby higher productivity and wage) an employee would be more inclined to stay with the same employer over time.

Firm size is included to control for higher productivity via increased division of labor and specialization in larger firms. This variable appears particularly important in smaller squares, since some of them can tend to be dominated by a few large firms, or even by one single establishment. Average years of schooling at the firm level is added to assess “peer effects” and spillovers within firms. There is some empirical evidence to support the notion that an inflow of well-educated individuals at the firm-level may boost the productivity of other employees (see e.g. Mas & Moretti, 2009).

The industry dummy variables controls for worker shifts between industries.

³ Since the estimations are performed using within variation, the interpretation in the empirical part relates to whether the employer is the same as the year before or not.

Table 2. Descriptive statistics (all figures are on the individual level) 1991-2008.

Variable	Mean	St.dev. (overall)	St.dev. (between)	St.dev. (within)	Min	Max
Yearly wage (SEK)	229,123	93,281	80,906	61,557	60,100	495,700
Density 250 (log)	13.07	1.82	1.72	.91	6.40	17.92
Density 1 000 (log)	14.52	1.78	1.68	.90	6.40	19.34
Density 10 000 (log)	16.94	2.02	1.93	.82	6.41	20.55
Density region (accessibility, log)	19.16	1.27	1.23	.40	14.48	21.32
Localization 250	.75	.27	.23	.17	.00	1
Localization 1 000	.65	.27	.23	.15	.00	1
Localization 10 000	.47	.20	.18	.10	.00	1
Schooling	11.79	1.94	1.99	.34	9	22
Experience	21.07	11.59	12.04	4.02	-7	49
Tenure	4.19	3.64	2.12	2.74	1	18
Firm size (log)	3.98	1.86	1.70	.89	0	9.41
Schooling (firm average)	11.84	1.24	1.17	.52	9	22

Total population size: 25,538,091.

The average 1 000 meter square associated with an individual in the dataset is about 4.5 times larger in terms of economic activity compared to the average 250 meter square. In turn, the average 10 000 meter square is home to about 10 times as much economic production as the average 1 000 meter. It is important to note that these are not differences in ‘density’, since the denominator in the square fraction differs between square sizes. The fact that the difference is not proportional to the difference in square sizes is easiest thought of as further evidence that economic activity is not evenly spread out across space. The larger squares contain larger proportions of ‘idle lands’: areas where there is no economic activity. Since economic production is clustered, smaller squares are overall made denser by a self-determining mechanism in the data generating process; i.e. small squares where there is no economic activity are eliminated from the data in the small square selection process, but they still constitute parts of larger squares. In terms of density, smaller squares host much more production per area unit of land compared to large squares, i.e., smaller squares are ‘denser’ on average. The same reasoning applies to the regional density measure. It is noteworthy that within to total variation is quite high for the neighborhood density variables, while it is lower for regional density.

It follows that the density of wage with respect to economic density is not straightforwardly comparable between regressions where density areas of different sizes are used, because “a one percent increase in density” is a relative statement, depending on the absolute scale properties of the area under observation; in terms of actual economic activity, the absolute amount of economic activity

associated with a one percent increase varies quite dramatically across square sizes. Hence, it is important to note that a constant elasticity of wage with respect to economic density across square sizes implies that a vastly larger increase in economic productivity is required to achieve a certain percentage change in an individual's wage in a larger square, compared to a smaller square.

Predictably, mean industry concentration (localization) is decreasing quite sharply with square size, indicating that there indeed is a higher localization present at the smaller neighborhood level. This is further evidence of the concern briefly mentioned in the introduction: the tendency for localization should not necessarily be thought of as exclusively a regional phenomenon.

Since the experience variable uses theoretical years of schooling, it can in a few cases be negative, reflecting the fact that some individuals complete multiple years of schooling in a given year (or start out earlier). This peculiarity illustrates one of the problems with using age as proxy for work experience, but does not pose any significant problems in the empirical part, since the coefficients are determined by within variation only.

The tenure variable is not equal to average number of years that an individual in the dataset has stayed with the same employer since it is only measured since the beginning of the reporting period. It is also pushed downward by the age cutoff values. It should simply be thought of as a control variable, indicating the quality of the employer-employee match. It may also be argued that job switching constitute some form of optimization on the part of the employee, in which case the tenure variable.

4. EMPIRICAL FINDINGS

The results from estimating (1) for each of the three density measures are displayed in table 3. Since the presented coefficients are fixed effect estimations, they are driven by “within” changes; i.e., the elasticity of wage with respect to density should be thought of as the percentage wage change associated with a 1 percent increase in the wage sums⁴ of the square where an individual is working, or alternatively, the percentage wage change associated with moving to a square where density is 1 percent higher, holding other factors constant.

⁴ Basing the analysis on employment density pushes the estimates down somewhat, but does not upset the conclusions. Supplementary regressions available upon request.

Table 3. Neighborhood density regressions on 250 meter, 1 000 meter and 10 000 meter squares (fixed effects) 1991-2008. Dependent variable: yearly wage (log).

Variable	Neighborhood aggregation (square bh)		
	250 m	1 000 m	10 000 m
Neighborhood density (log)	.0194 (.00007)	.0105 (.00006)	.00982 (.00007)
Localization	.0340 (.000338)	.0211 (.000372)	.0378 (.000575)
Years of schooling	.0518 (.00395)	.0520 (.00395)	.0523 (.00395)
Experience	.0287 (.00394)	.0290 (.00395)	.0294 (.00395)
Experience ²	-.0006 (.00000)	-.0006 (.00000)	-.0006 (.00000)
Tenure	.00534 (.00002)	.00529 (.00002)	.00531 (.00002)
Firm size (log)	.0114 (.00008)	.0192 (.00007)	.0210 (.00006)
Establishment schooling	.0166 (.000114)	.0182 (.000114)	.0185 (.000114)
Year dummies	Yes	Yes	Yes
2-digit industry dummies	Yes	Yes	Yes
Within R ²	.39	.38	.38
# Observations	25,538,091	25,538,091	25,538,091
# Individuals	3,478,681	3,478,681	3,478,681
# Neighborhoods	67,779	12,714	1,454

Standard errors in brackets (all variables are significant beyond the 1 percent level). All variables are defined in section 3. One mile is equal to 1 609 meters.

Neighborhood density indeed seems to matter. The elasticity of wage with respect to density is about .02 for 250 meter squares, .01 for 1 000 and 10 000 meter squares, i.e. when observable and unobservable characteristics are controlled for, a doubling of the density of economic activity in a neighborhood is associated with wages increasing by about 1-2 percent, holding observable worker characteristics, firm attributes, and neighborhood characteristics constant. However, the magnitude of the effect depends on the level of aggregation, with the smaller squares representing the upper bound.

One interpretation is then that an increase of economic production in a geographical area is more beneficial to an individual the closer to him it is; i.e. the image that emerges certainly strengthens the case for face to face interaction in learning. Such a story would fit e.g. Jovanovic and Rob (1989) or

other approaches wherein knowledge diffusion is at least in part the product of human interaction. It would also strengthen the assertion that knowledge spillovers are subject to significant distance decay.

The positive industry concentration variable indicates that positive localization economies are present across square sizes, i.e. increases in concentration of an industry will tend to increase wages for workers in that industry; overall, a 10 percentage point increase in industry concentration increases wages in that industry by about 0.2-0.4 percent. The coefficient is increasing slightly with square size, although the lowest coefficient is registered for 1000 meter squares. Further, the tenure coefficient is negative, but the estimate is hardly economically significant.

The schooling and experience coefficients have the expected signs and the magnitudes of the effects are broadly in concordance with previous findings (see e.g. Angrist & Krueger, 1991; Rauch, 1993), but it should be noted that fixed effects regressions can be problematic when analyzing these variables. However, since the schooling variable is theoretical years of schooling, it can tend to jump several years from one year to the next, making it easier to distinguish from the general trend.

4.1 ROBUSTNESS

At least two plausible arguments may be raised against the estimations in table 3. The first is that perhaps a dense square is really a proxy for a dense region and that the coefficient may be driven by the larger agglomeration. The second objection deals with what Combes et al. (2007) refer to as *endogenous quantity of labor*: endogeneity determined by the tendency for some locations in space to be ‘inherently’ more productive and thereby denser. These two arguments are addressed below.

In this empirical framework, the first proposition can be thought of as the tendency for the average square to increase in size as a region experiences in-migration, and hence that neighborhoods and regions tend to move in the same direction in terms of both density and productivity (to the extent that in-migration increases productivity). Wages can be thought of as driven by the overall agglomeration becoming larger, but the effect may still be picked up by the variable measuring an increase in neighborhood size. To test this proposition, it is necessary to consider the possibility of excluded variables in space, by controlling for density at the regional level.

The Swedish labor market is made up of 290 urban regions. These are aggregated into 81 functional regions (local labor markets), each of which is highly integrated in terms of commuting flows. The measure of density used here uses a distance-decay weighted accessibility approach, where the accessibility to economic activity in each urban region is the sum of local, inter-regional, and intra-regional accessibility to wage sums. This approach is part of a line of thought dating back to the market potential measure in the seminal article by Harris (1954). Formally:

$$Density_i = AW_{iL} + AW_{iR} + AW_{iOR}$$

$$AW_{iL} = x_i \exp^{-\lambda_L t_{ii}}, \text{ local accessibility to total wage earnings of municipality } i,$$

$$AW_{iR} = \sum_{j \in R_i} x_j \exp^{-\lambda_R t_{ij}}, \text{ intra-regional accessibility to total wage earnings of municipality } i,$$

$$AW_{iOR} = \sum_{k \notin R_i} x_k \exp^{-\lambda_{OR} t_{ik}}, \text{ inter-regional accessibility to total wage earnings of municipality } i,$$

where internal accessibility is simply the wages accumulated in each urban region, weighted by average commuting time by car. Intra-regional accessibility is accessibility to wages generated in the same functional region. The inter-regional measure picks up accessibility to all other urban-regions in the country. The distance-decay parameter λ takes on three different values for internal, intra-regional and inter-regional accessibility, respectively. These parameter values are based on observed commuting behavior of workers, and are estimated for Swedish municipalities by Johansson et al. (2003).

The summed up accessibility measure represents a continuous view of geography, where all activities are related in space, but where the magnitude of the effect diminishes with distance; first only marginally, but eventually quite sharply, consistent with commuters' non-linear responses to differences in time-travel distances. The integrated labor market approach is important, e.g. in order to be able to assess labor market pooling.

Viewing spatial autocorrelation as an issue of modeling spatial effects in an inclusive manner essentially reduces the problem to an omitted variable problem. The accessibility measure used has been shown to decrease or eliminate spatial autocorrelation, since it captures spatially lagged effects (M. Andersson & Gråsjö, 2009). Hence, apart from capturing the effects of regional density, the measure is also helpful in model specification. When used in previous studies the elasticity of wage with respect to accessibility has been in the 2 percent neighborhood in longitudinal studies (Larsson, Andersson, & Klaesson, 2011).

Table 4 shows the previous estimations, now including the regional density (accessibility) variable.

Table 4. Neighborhood and regional density regressions on 250 meter, 1 000 meter and 10 000 meter squares (fixed effects) 1991-2008. Dependent variable: yearly wage (log)

Variable	Neighborhood aggregation (square bh)		
	250 m	1 000 m	10 000 m
Neighborhood density (log)	.0176 (.00007)	.00774 (.00007)	.00458 (.00009)
Regional density (log)	.0155 (.000148)	.0182 (.000152)	.0178 (.000185)
Localization	.0345 (.000338)	.0210 (.000371)	.0331 (.000577)
Years of schooling	.0515 (.00394)	.0518 (.00395)	.0519 (.00395)
Experience	.0287 (.00394)	.0289 (.00395)	.0291 (.00395)
Experience ²	-.000631 (.00000)	-.000633 (.00000)	-.000633 (.00000)
Tenure	.00541 (.00002)	.00537 (.00002)	.00538 (.00002)
Firm size (log)	.0115 (.00008)	.0191 (.00007)	.0210 (.00006)
Years of schooling (firm mean)	.0155 (.000115)	.0171 (.000114)	.0180 (.000114)
Year dummies	Yes	Yes	Yes
2-digit industry dummies	Yes	Yes	Yes
Within R ²	.39	.38	.38
# Observations	25,538,091	25,538,091	25,538,091
# Individuals	3,478,681	3,478,681	3,478,681
# Neighborhoods	67,779	12,714	1,454
# Urban regions	290	290	290

Standard errors in brackets (all variables are significant beyond the 1 percent level). All variables are defined in section 3. One mile is equal to 1 609 meters.

The coefficient of the accessibility variable performs roughly as expected, even though it is somewhat lower in the 250 meter regression. The neighborhood elasticities of the larger squares are now lower; down by about 20 percent in the 1 000 meter regression cut in half in the 10 000 meter squares. When looking at the 250 meter coefficient, it is only down by about 10 percent.

It seems plausible, then, that some of the effects picked up by the neighborhood density variables in table 3 were indeed the products of regional external economies, particularly in the cases of the larger squares. The most disaggregated measure is however quite resilient. This result underlines the need to

model neighborhood effects at a disaggregated level. A doubling of neighborhood density is associated with wages increasing by roughly 1.8 percent, holding all other variables constant, *including the extent of economic activity at the regional level*. The fact that the accessibility measure is lower in this estimation further suggests that used on its own (or together with the scaled-up neighborhood measures), the accessibility measure successfully picks up part, but not all, of the effects of the neighborhood.

The results hence confirm that there is indeed a regional effect, but that failure to model the neighborhood effect will result in a model where either too much of the external effects are attributed to the regional level or, worse, that the effect is omitted in its entirety, depending on the extent to which dense neighborhoods are correlated with dense regions.

The second objection, referred to as endogenous quantity of labor, deals with direction of causality. The proposition is that if some areas are inherently more productive, then they are possibly made denser in the data generating process. Combes et al. (2008) find that endogeneity is less important than selection, but ideally the problem should be addressed with an instrumental variable approach which, given the empirical framework, does present some challenges.

Instrumental variables commonly used since Ciccone and Hall (1996) include historical populations at the city level. These variables correlate rather poorly with neighborhood density in the smaller squares. Additionally, such variables are determinants of today's region density, which in this empirical framework is another determinant of the dependent variable. This anomaly is an obvious weakness of this approach and essentially means that most available approaches are invalidated by construction.

An additional way forward is noted in Combes et al. (2007), where geological instruments are used in conjunction with a micro-level dataset of French workers. By matching data from the European Soil Database (ESDB) to the square grid using GIS software, several square-level indicators of geology are matched to the dataset. The indicators include water capacity, parent material and other properties of the soil. The argument requires that the geological indicators proxy for historical population distributions, without driving today's productivity. One plausible channel is e.g. the tendency for fertile lands to have determined historical settlement patterns, without being important determinants of modern-day productivity.

There are at least three problems with the geological approach. First, within-Sweden variation is quite low for most available variables. Second, the ESDB is not intended for use at the current level of disaggregation. Third, the time-invariant instruments are hard to combine with the benefits of longitudinal data. Hence, it should be noted that the IV estimations, using a 2008 cross section and

2SLS, in table 5 are intended as robustness checks only. The OLS estimate is reported for comparison, and is clustered at the neighborhood level as recommended by Moulton (1990).

Table 5. Neighborhood density (ln) coefficient (2008 cross-section) from geological instrumental variable regressions.

	Neighborhood aggregation (square b*h)					
	250 m 2SLS	250 m OLS	1 000 m 2SLS	1000 m OLS	10 000 m 2SLS	10 000 m OLS
Neighborhood density (log)	.0442 (.002)	.0282 (.011)	.028 (.001)	.0134 (.001)	.018 (.000)	.0062 (.001)
F (excluded instruments)	2092		3525		11,178	
First-stage R ²	.02		.03		.08	
R ²	.29	.29	.28	.28	.28	.28

Standard errors in brackets (all variables are significant beyond the 1 percent level). The control variables are identical to those reported in table 4. Geological instruments used: parent material (3 dummies), water capacity (3 dummies), depth to rock (4 dummies), top soil type (3 dummies), and sub soil type (2 dummies). Population size: 1,711,078.

Note that unobserved ability is likely to put upward pressure on the estimates in table 5. The size of the coefficient is therefore of lesser importance. What's more interesting is the direction of the effect: across the board, the 2SLS estimates are *higher* than those obtained by OLS. This credulous result indicates that reverse causality is not detrimental to the findings above.

6. CONCLUSION

In this paper I argue that much of the discussion about agglomeration economies over the last decade has in fact implied the existence of a very localized effect of human capital spillovers, in addition to the regional effect from the classic agglomerating factors. Specifically, since human capital has a noted tendency to attenuate with distance and since there is substantial clustering of industries at the sub-city level. The empirical analyses carried out thus far have commonly been constrained by availability of data in analyzing this effect.

Using Swedish register data I estimate the neighborhood effect using squares with bases of 250 meters, 1 000 meters and 10 000 meters to represent the neighborhood level. The effect is found to be slightly below 2 percent in the 250 meter neighborhood squares while controlling for accessibility (market potential) at the regional level. In the 1 000 meter and 10 000 meter squares, the effect is .08 and .05 percent, respectively, suggesting that the results are sensitive to choice of spatial scale.

Broadly, the analysis is in concordance with recent approaches that stress the attenuation of human capital spillovers. The results are consistent with the existence of a neighborhood effect, but they also

underscore that the effect is a very local phenomenon that attenuates sharply with distance. If the neighborhood data is not sufficiently disaggregated, this effect will result in an omitted variable bias that is only perfectly captured at the regional level to the extent that dense neighborhoods are perfectly correlated with dense regions. Even in such a hypothetical case, it would not be possible to untangle the conceptually quite different forces that contribute to agglomeration economies.

REFERENCES

- Andersson, M., & Gråsjö, U. (2009). Spatial dependence and the representation of space in empirical models. *The Annals of Regional Science*, 43(1), 159-180.
- Andersson, M., & Klaesson, J. (2009). Regional interaction and economic diversity: exploring the role of geographically overlapping markets for a municipality's diversity in retail and durables. In C. Karlsson, B. Johansson & R. R. Stough (Eds.), *Innovation, Agglomeration and Regional Competition*. Cheltenham, UK: Edward Elgar.
- Andersson, Å. E., & Beckmann, M. J. (2009). Economics of Knowledge - Theory, Models and Measurements. Cheltenham, UK: Edward Elgar.
- Angrist, J. D., & Krueger, A. B. (1991). Does Compulsory Schooling Attendance Affect Schooling and Earnings? *The Quarterly Journal of Economics*, 106, 976-1014.
- Capello, R. (2002). Entrepreneurship and spatial externalities: Theory and measurement. *The Annals of Regional Science*, 36(3), 387-402.
- Ciccone, A., & Hall, R. E. (1996). Productivity and the Density of Economic Activity. *The American Economic Review*, 86(1), 54-70.
- Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2), 723-742.
- Combes, P.-P., Duranton, G., Gobillon, L., & Roux, S. (2007). Estimating Agglomeration Economies with History, Geology, and Worker Effects In E. L. Glaeser (Ed.), *Agglomeration Economics*: The University of Chicago Press.
- Dicken, P., & Lloyd, P. E. (1990). *Location in Space - Theoretical Perspectives in Economic Geography* (3 ed.). New York: Harper Collins.
- Glaeser, E. L. (1999). Learning in Cities. *Journal of Urban Economics*, 46(2), 254-277.
- Gordon, P., & Ikeda, S. (2011). Does Density Matter? In D. E. Andersson, Å. E. Andersson & C. Mellander (Eds.), *Handbook of Creative Cities*. Cheltenham, UK: Edward Elgar.
- Granovetter, M. S. (1973). The Strength of Weak Ties. *The American Journal of Sociology*, 78(6), 1360-1380.
- Harris, C. D. (1954). The Market as a Factor in the Localization of Industry in the U.S. *Annals of the Association of American Geographers*, 44, 315-348.
- Helsley, R. W., & Strange, W. C. (1990). Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics*, 20(2), 189-212.
- Jacobs, J. (1961). *The Death and Life of Great American Cities*. New York: Random House.

- Jacobs, J. (1969). *The Economy of Cities*. New York: Random House/Vintage Books.
- Johansson, B., Klaesson, J., & Olsson, M. (2003). Commuters' non-linear response to time distances. *Journal of Geographical Systems*, 5(3), 315-329.
- Jovanovic, B., & Rob, R. (1989). The Growth and Diffusion of Knowledge. *The Review of Economic Studies*, 56(4), 569-582.
- Larsson, J. P., Andersson, M., & Klaesson, J. (2011). *How Sensitive is the Urban Wage Premium to Worker and Employer Heterogeneity?* Unpublished manuscript, JIBS working paper.
- Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1), 3-42.
- Marshall, A. (1920). *Principles of Economics* (8 ed.). London: MacMillan.
- Mas, A., & Moretti, E. (2009). Peers at Work. *American Economic Review*, 99(1), 112-145.
- McCloskey, D., & Klammer, A. (1995). One Quarter of GDP is Persuasion. *The American Economic Review*, 85(2), 191-195.
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: National Bureau of Economics Research.
- Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *The Review of Economics and Statistics*, 72(2), 334-338.
- Rauch, J. E. (1993). Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities. *Journal of Urban Economics*, 34(3), 380-400.
- Roback, J. (1982). Wages, Rents, and the Quality of Life. *The Journal of Political Economy*, 90(6), 1257-1278.
- Romer, P. M. (1990). Endogenous Technological Change. *The Journal of Political Economy*, 98(5), S71-S102.
- Rosenthal, S. S., & Strange, W. C. (2001). The Determinants of Agglomeration. *Journal of Urban Economics*, 50(2), 191-229.
- Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, 64(2), 373-389.