CentER

Discussion paper

No. 2002-41

KNOWLEDGE EXTERNALITIES, AGGLOMERATION ECONOMIES, AND EMPLOYMENT GROWTH IN DUTCH CITIES

By Daan P. van Soest, Shelby D. Gerking and
Frank G. van Oort

April 2002

ISSN 0924-7815

Knowledge Externalities, Agglomeration Economies, and Employment Growth in Dutch Cities*

Daan P. van Soest[1], Shelby D. Gerking[2] and Frank G. van Oort[3]

[1]Department of Economics and CentER, Tilburg University, Tilburg, The Netherlands
[2]Department of Economics, University of Central Florida, Orlando, United States
[3]Department of International Economics and Economic Geography, University of Utrecht, Utrecht, The Netherlands & Institute for Spatial Policy Analysis, The Hague, The Netherlands

Abstract:

This paper extends the work of Glaeser *et al.* (1992) by looking at effects of agglomeration economies on employment growth in Dutch city-industries and in very small (postal) zip code-industries in the Dutch province of South-Holland. At both levels of geographic detail, findings are broadly consistent with results from the earlier study in that employment growth is enhanced by industrial diversity and local competition, but retarded by industrial specialization. Also, a novel feature of the analysis presented here is that we examine the extent to which agglomeration economies in one location affect employment growth in other locations.

Key words: endogenous growth, knowledge spillovers, agglomeration economies, spatial lag models;
JEL codes: O12, O18, R11, R30.

Knowledge Externalities, Agglomeration Economies, and Employment Growth in Dutch Cities

1. Introduction

Endogenous growth theory emphasizes the role of knowledge possessed by economic agents and identifies knowledge spillovers between them as a crucial factor leading to external economies of scale in production (Romer 1986, Lucas 1988). Two important aspects of this theory, however, are largely unresolved. First, under what circumstances does knowledge spill over most easily between economic agents? Whereas Glaeser *et al*. (1992) and Feldman and Audretsch (1999), find that local employment growth is enhanced by diversity of activity across a broad range of sectors, Henderson *et al*. (1995), Black and Henderson (1999a), and Beardsell and Henderson (1999), find evidence that employment grows faster when more activity is concentrated in a single industry. Second, how does knowledge generated in one location affect economic growth in another location? Using data on innovations and patents, several papers find that knowledge is geographically bounded within the region where it was generated (Jaffe 1989, Jaffe, Trajtenberg, and Henderson 1993, Audretsch and Feldman 1996, Anselin, Varga, and Acs 1997, Rosenthal and Strange 2000). Black and Henderson (1999b, p. 255) recognize the importance of links between these two questions and refer to modeling human capital spillovers over space as an exciting extension of the literature (see also Hanson 2001). Additionally, the relative importance of various types of externalities in fostering both localized economic growth and growth among more geographically dispersed areas has broad implications ranging from the formulation and interpretation of endogenous growth models to practical conclusions that might be drawn by policy makers regarding urban planning and development.

The purpose of this paper is to measure the contribution of knowledge externalities to both local growth and growth at other locations using unique data from the Netherlands. We follow existing literature by focusing on relationships between employment growth and types of agglomeration economies thought to reflect these externalities and obtain four main results. First, we look at Dutch city-industries, and find that the relationship between agglomeration economies and employment growth identified by Glaeser *et al.* (1992) carries over to the Dutch case. Sensitivity analyses, however, suggest that this outcome rests to some extent on how key agglomerative factors are measured as well as on other basic issues of model specification. Second, we conduct a closely related analysis of employment growth in (postal) zip code-industries in the Dutch province of *Zuid-Holland* (South-Holland), finding that the same types of agglomeration economies that are important to growth at the city-industry level also promote growth within a city. Third, we look at the extent to which agglomeration economies in one location affect employment growth at other locations using both the Dutch city and South-Holland zip code data. We find that spatial effects of agglomeration economies appear to fade quickly with distance, even for South-Holland where geographic units analyzed are small areas within a single city. Taken together, an important implication of these results is that agglomerative forces associated with knowledge externalities may well operate at a geographic scale that is much smaller than a city. Fourth, because the South-Holland data distinguish between employment growth in establishments present in a base year and employment growth due to later establishment births and/or relocations, we briefly look at effects of agglomeration economies on sources of employment growth.[1]

2. Background and Data

Cities are fertile grounds for testing knowledge-based theories of endogenous growth because dense urban agglomerations provide opportunities for learning through the many new contacts they offer (Lucas 1993, Glaeser 1999). Prior studies have tested three (in some respects competing) hypotheses concerning the conditions under which knowledge spillovers affect growth. The first hypothesis, originally developed by Marshall (1890) and later formalized by Arrow (1962) and Romer (1986) (MAR), emphasizes spillovers between firms in the same industry that arise, for example, through inter-firm movements of skilled labor. These spillovers are thought to be most important when little prevailing local competition is present so that rents associated with sector-specific knowledge can be internalized. The second hypothesis, as developed by Porter (1990), agrees that knowledge spillovers within a localized industry are most important, but argues that their effects on growth are enhanced when fierce local competition forces firms to innovate in order to survive. The third (Jacobs 1969) emphasizes spillovers between (rather than within) industries, playing up the notion of cross-fertilization of ideas, and contends that they promote growth most effectively in a competitive environment.

As indicated in the introduction, several papers test these hypotheses and reach conflicting conclusions. An important difference in the methodologies used in these studies is whether data are analyzed one manufacturing industry at a time as in Henderson *et al.* (1995) or whether the sample consists of the (six) largest industries of all types in each city (city-industries) as in Glaeser *et al*. (1992). The "one-industry-at-a-time" approach is less restrictive than looking at city-industries in that it allows effects of agglomerative forces to differ by industry and does not limit consideration only to the largest six industries in a city. The Netherlands, however, may not be the best setting for analyzing growth of detailed manufacturing industries because of the relatively small number of locations where they can flourish.[2] Consequently, we adopt the city-industry approach, probe the robustness of results obtained to a variety of changes in specification, such as increasing the number of industries taken from each city, changing the way in which key variables are measured, and considering the extent to which effects of agglomeration economies spill over between geographic locations.

Data for this study come from annual surveys of establishments in all lines of business in the Netherlands. Establishments are enumerated based on information furnished by the Chamber of Commerce, insurance companies, and industrial sector associations, and an annual questionnaire is sent to each. Annual response rates to the questionnaire averaged 96%. Questionnaire results identify each establishment's 6-digit zip code (a small area containing about 100 different mailing addresses), and 5-digit activity code. Thus, these data represent a census of all business establishments in the Netherlands. They are not subject to disclosure rules that apply to publicly available employment data in the U.S. While Dutch 5-digit activity codes have no direct parallel to the industrial classification system used in the U.S., they still permit identification of industries defined at a very detailed level.

The data are extensively checked for accuracy with special attention given to businesses reporting that they have no employees. Many such businesses turn out to be (1) 'mailbox' enterprises that may be established for purposes other than making or selling goods and services (i.e., tax advantages for the owner, access to businesses–only retail outlets, or provide a postal address for a firm doing business at another location), or (2) bankrupt enterprises that had not yet been eliminated from the register. Enterprises not actually doing business are excluded from the data set. Also, data available from each of the 12 Dutch provinces are not exactly the same. For example, the survey was first conducted in South-Holland in 1988 with all other provinces participating by 1991. Also, in most provinces, only employment totals are available by industry and zip code, whereas the South-Holland data set contains information about individual establishments.

A disadvantage of the data is that they do not contain measures of outputs, inputs other than labor, or plant characteristics. Consequently, they are not appropriate for estimating establishment-level production functions, as in Black and Henderson (1999a) and Beardsell and Henderson (1999). While these two studies are important because the contribution of agglomeration economies to productivity is estimated from plant level production functions in a fixed effects framework, construction of the panel sacrificed considerable information. Black and Henderson, for example, ended up with sample sizes averaging 8% of plants in an industry; thus better control of establishment-specific effects was obtained by accepting a possible selectivity problem. Using either the Dutch municipality or the South-Holland zip code data, on the other hand, estimation of establishment-level production functions is not a realistic option in any case because employment is the only establishment-specific variable available, and hence the data are not well suited for panel analyses. It is possible that aggregation of the data for each industry into geographic units may average out a portion of the establishment-specific effects. A portion of these effects will remain uncontrolled, however, if there is geographic clustering of high quality entrepreneurial talent, clustering of older and/or newer establishments, or clustering of establishments using particular specialized inputs.

Also, the level of detail available in the data actually is too great for meaningful analysis. Many of the 6-digit zip code areas, for example, have only residences and individual 5-digit industries are present in relatively few 6-digit zip codes. Additionally, postal zip codes are arbitrary administrative units, not functional economic areas. In consequence, the data were used in two ways. First, for the Netherlands as a whole we identified 57 cities, thought to approximate functional economic areas, and then built up employment totals and other variables from the 5-digit activity code, 6-digit zip code level. These cites included Amsterdam, Rotterdam, and The Hague, and all cities considered had a population of at least 50,000 persons.[3] Then, data for industries at the 2-digit activity code level (roughly the same level of detail as 2-digit industries in the U.S.) were built up from the 5-digit activity code information. Industries were aggregated in this way mainly to maintain comparability with other studies and because many of the more detailed sectors have only a very limited presence in the Netherlands.

Second, we constructed a data set for the province of South-Holland by aggregating the 6-digit zip code, 5-digit activity code data up to the 4-digit zip code, 2-digit activity code level. South-Holland is approximately 2350 km^2, covers a large part of the core economic area of the Netherlands (the 'Randstad'), and has a high population density (about 1190 persons/km^2). The province has 416 4-digit zip code areas and the average size of each is about 5.65 km^2, although they tend to be smaller in urban centers where the density of addresses is high and larger in areas that have more open space. Because the province can be viewed as a single urbanized area, the South-Holland data provide an opportunity to: (1) determine whether the same set of factors that promotes growth at the city level also operate at a more localized level and (2) look for evidence of the extent of knowledge transfer between locations within a city. Also, in comparison to the Dutch city data, the South-Holland data cover a longer time period (1988-97), offer better controls for unmeasured heterogeneity, and, as mentioned earlier, are more detailed than the data available for other provinces.

Table 1 shows the ten sectors that turned up most often among the six largest sectors in the 57 Dutch cities and the 416 zip code areas in South-Holland, and the number of employees in each. The most well represented sectors in each of the two samples are building and construction, retail trade, business services (predominantly financial institutions and services), health care, education, and wholesale trade. Notice that manufacturing industries appear less often in these samples than do non-manufacturing industries. Further details regarding the data sets as well as information on the construction of the explanatory variables are discussed in Appendix A.

3. *Employment Growth in Dutch Cities*

Our analysis begins by using data from Dutch city-industries to explain employment growth using an equation (Table 2, Column (1)) that comes as close as possible to the specification used by Glaeser *et al*. (1992). The dependent variable is the natural logarithm of the ratio of 1997 employment to 1991 employment in city-industries. All explanatory variables measure initial (1991) conditions both to determine their effects on employment growth and to reduce complications from simultaneity. Three explanatory variables measured aspects of agglomeration economies emphasized by the previously described theories of knowledge spillovers.[4] Local industrial concentration is measured by taking the ratio of the percentage of a city's employment in an industry in 1991 to the corresponding national percentage. This ratio, sometimes referred to as a location quotient, tells whether an industry is relatively over or underrepresented in a city compared to the nation as a whole. Industrial diversity in a city is measured by the percentage of employment in 1991 in the largest five industries, excluding the one under observation. Larger values of this percentage suggest lower levels of industrial diversity. Local competition is measured by the 1991 ratio of establishments per worker in a city-industry to establishments per worker in that industry nationally. Higher values of this measure means that establishments in a city-industry are relatively small as compared to the size of establishments in that industry nationally. While Glaeser *et al.* argue that smaller establishment size implies greater local competitive pressure, this interpretation has been called into question by Combes (2000), who contends that it may instead measure internal diseconomies of scale, and by Rosenthal and Strange (2000), who view it as a broader measure of industrial organization. In any case, a positive coefficient of industrial concentration and a negative coefficient of competition lend support to the MAR hypothesis. Positive coefficients of industrial concentration and competition support the Porter hypothesis. A negative coefficient of industrial diversity and a positive coefficient of competition supports the Jacobs hypothesis.

Additionally, other variables included as controls are: (1) the growth rate of the industry outside the city measured as the natural logarithm of 1997 employment to 1991 employment (to capture general changes in industry productivity or shifts in demand for the industry's output), (2) the 1991 industry wage rate, (3) employment in the city-industry in 1991, (4) whether the city is located in the Randstad, and (5) whether a city is located in the most rural areas of the Netherlands, referred to here as the periphery. To replicate the Glaeser *et al.* approach, we first focus on analyzing growth in the largest six city-industries, so we have 57x6=342 observations. The means and standard deviation of the dependent and explanatory variables are presented in Table A1 in Appendix A.

Regression results based on the Glaeser *et al.* specification, presented in Table 2, Column (1), support the Jacobs hypothesis. Standard errors, computed using White's correction for heteroskedasticity, are reported beneath estimated coefficients. Summary statistics show that the adjusted R^2 equals 0.227, which is lower than the corresponding value of 0.450 obtained by Glaeser *et al.* (see their Table 3, Column (4), p. 1143). Also, the likelihood ratio statistic LR(SL) tests the null hypothesis of no spatial lag in the dependent variable and the Lagrange multiplier statistic LM(SE) tests the null hypothesis of no spatial dependence in the error term. Under the null hypotheses, both statistics are Chi-square distributed with one degree of freedom and show no evidence of spatial autocorrelation.

Column (1) results suggest that local concentration in an industry retards employment growth while local industrial diversity and local competition promote growth.[5] As noted in the table, coefficients of these variables are significant at the 5% level, or lower. Also, estimates show that a one standard deviation increase in the share of a city's employment in the five largest industries lowers employment growth in a city-industry by 5.3% over the seven-year period studied. One standard deviation increases in the location quotient measure of industrial specialization and in the firms per worker measure of competition alter city-industry growth rates by –11% and 7.5%, respectively. Regarding performance of the other control variables, initial employment in a city-industry, the initial industry wage, and location in the Randstad or periphery are unimportant determinants of employment growth. Glaeser *et al.* found that both high initial wages and high initial employment levels lowered the rate of employment growth, but in that study, data on wages were available by both city and industry whereas in the Netherlands, data on wages are available by industry but not by city. Additionally, similar to the Glaeser *et al.* results, the coefficient of industry growth outside the city is significantly greater than unity, indicating that employment growth rates are higher in cities than in rural areas.[6]

Because of the previously cited differences in results obtained in related studies of employment growth in cities, it is worthwhile to check whether these results are robust against several types of specification changes. First, the variables measuring industry growth outside the city and the industry wage rate were replaced with 24 industry dummy variables 56 city dummy variables to better control for industry-specific and city-specific effects on employment growth.[7] Coefficients of these dummies were jointly significant at the 1% level, but because the overall performance of the agglomeration economy measures was broadly similar to that reported in Column (1) (the coefficient of the industrial diversity variable, however, more than doubled in absolute value), results of this regression are reported in Table A2 in Appendix A rather than in Table 2. In any case, it appears that the measures of agglomerative factors used here are not simply proxies for factors such as labor market characteristics or natural location advantages that might be expected to vary systematically across cities or for factors that might vary systematically across industries.

Second, the largest 12 (instead of 6) industries were taken from each city (thus increasing the number of observations to 57x12=684; for the means and standard deviations of the various variables, see Column (2) in Appendix A Table A1). Because including the industry- and city-specific dummy variables did not dramatically alter results here either, we report the outcome based on using the Table 2, Column (1) with the exception that industrial diversity is measured as the share of a city's employment in the largest eleven (rather than five) industries in 1991, not counting the one under observation. As shown in Table 2, Column (2), increasing the number of (smaller) industries taken per city results in a smaller coefficient of industrial diversity that does not differ significantly from zero at the 5% level. This outcome suggests that the Jacobs hypothesis may be more applicable to larger city-industries and is consistent with one of the findings of Henderson *et al.* That study included many small city-industries in the analysis and found little evidence that initial industrial diversity mattered for later employment growth.[8]

Third, using the "top 12" sample, we tested whether the competition variable matters more for non-manufacturing industries than manufacturing industries.[9] A possible conjecture in this regard is that the competition variable is an indicator of both product market and labor market competition for non-manufacturing establishments that sell goods and services only locally, but an indicator of just labor market competition for manufacturing establishments that are more likely to sell in national or worldwide markets (see Feldman and Audretsch 1999). Results presented in Appendix A, Table A2 show that the coefficient of competition is positive and significantly different from zero in a regression using 507 non-manufacturing city-industries, and is positive with a t-statistic of less than one in a regression using 177 manufacturing city-industries. This outcome might be interpreted as a specification test because a finding that local competition matters for industries that produce traded goods might cast doubt on the method used here for identifying effects of competition on employment growth. Furthermore, our results for manufacturing industries stand in contrast to those obtained in the Henderson *et al.* study, as we find that growth is harmed by industrial specialization. With respect to the role of space, we find that manufacturing industries grow slower in the country's periphery, and a Lagrangian Multiplier test (LM(SE)) detects spatial error dependence.

Fourth, we checked whether the lack of support found above for the MAR and Porter hypotheses was due to the way in which industrial specialization and competition were measured. Regarding specialization, the location quotient measures relative concentration of an industry in a city as compared to its Netherlands' average (see Appendix B). What might matter for spillovers, however, could be the absolute scale of an industry in a city, rather than its relative scale in a city compared to the country. In consequence, we tried measuring industrial specialization with the fraction of a city's employment in an industry (which is just the numerator of the location quotient) and with employment in an industry per square kilometer. A similar argument can be made regarding the competition variable in that knowledge spillovers may be more fully internalized when an industry consists of a large number of small establishments rather than of a small number of large establishments *independently* of the industry's average firm size in the rest of the country. In any case, we tried measuring competition as the number of establishments per employee in an industry (the numerator of the competition variable used in Columns (1) and (2)), and the number of establishments in a city-industry divided by total city employment.

Overall, these alternative measures of specialization and competition performed inconsistently as illustrated by the example regression presented in Table 2, Column (3). This regression is specified identically to the one in Column (1), except that specialization is measured by city-industry employment divided by total city employment and competition is measured by the number of city-industry establishments per employee. In this regression, measures of industrial diversity and specialization have coefficients that do not differ significantly from zero at the 5% level, and the coefficient of the competition measure is negative with a t-statistic of –2.09. One interpretation of this outcome is that, in contrast to regressions presented in Columns (1) and (2), knowledge spillovers are most easily internalized when establishments present in a city have market power. Little confidence, however, can be placed in this result. When the regression was rerun with the industry dummies, none of the three agglomeration economy variables had coefficients with t-statistics that exceeded unity. Thus, the relative magnitude of a city's agglomeration economies compared to the Netherlands as a whole appears to matter more for growth than their absolute scale. A possible explanation here is that the relative measures of industrial specialization and competition control for the size of industries at the national level, whereas the absolute measures do not.

Fifth, even though diagnostics from the three regressions reported in Columns (1)-(3) in Table 2 indicate little evidence of spatial autocorrelation (but see the top-12 manufacturing regression in Column 2 of Table A2), we nevertheless looked at the extent to which effects of agglomeration economies spill over between cities. This analysis is pursued because it bears on the question of how far knowledge is transmitted over space as well as on the closely related issue of whether cities are the most appropriate geographic unit in which to study knowledge spillovers (see below). We investigated this issue by: (i) adding distance-weighted (gravity) measures of industrial diversity, competition, and industrial specialization to the Table 2, Column (1) regression and (ii) estimating the Table 2, Column (1) regression in a spatial lag framework (Anselin 1988), both with and without gravity variables. For a given city, these gravity variables are defined as distance-weighted sums of agglomeration economies in all other cities. These variables capture direct effects of agglomeration economies elsewhere on city-industry employment growth in a particular city. On the other hand, the spatial lag model posits an indirect relationship between the agglomeration economies elsewhere and employment growth at home that may occur, for example, through input-output linkages. Calculation of the distance-weighted agglomeration economy variables is more fully explained in Appendix B and a brief overview of the spatial lag model is presented in Appendix C. Both ordinary least squares and maximum likelihood spatial lag estimates consistently indicate that coefficients of the gravity variables for industrial specialization and diversity are not significant at the 5% level and that the coefficient of the gravity competition variable is positive and significantly different from zero at 1%. Also, the spatial lag coefficients are never significantly different from zero at 5% using two alternative spatial weight matrices (see Appendix C), no matter whether the gravity variables are included. These outcomes are illustrated by the example spatial lag estimates reported in Table 2, Column (4), which includes distance-weighted agglomeration economy measures as explanatory variables. Notice that in this regression the three "own-city" agglomeration variables still perform much as they did in the Table 2, Column (1) regression. Thus, agglomeration in one city appears to have only limited effects on employment growth in other cities, at least in comparison to their effects on growth at home. This outcome is consistent with prior studies (Jaffe 1989, Jaffe *et al.* 1993, Audretsch and Feldman 1996, Anselin *et al.* 1997, Rosenthal and Strange 2000) that found that knowledge stays close to the location where it was generated.

4. Employment Growth in South-Holland

This section analyzes the South-Holland zip code data in two ways that tie in directly with results just presented for Dutch cities. First, we look at relationships between agglomeration economies and employment growth using the South-Holland zip code data. City and sub-city analyses turn out to be complementary in that, together, they shed light on the extent to which employment growth within cities is localized, as well as for choosing the appropriate geographic level at which to study effects of agglomeration. Second, because the South-Holland data are more detailed than are corresponding data for other provinces, we look at the extent to which agglomeration economies contribute to employment growth in establishments originally present in a zip code as contrasted with their effects on stimulating establishment births and/or relocation of establishments from elsewhere.

Because of South-Holland's small size, the zip code data provide natural control for important location-specific factors that affect growth. In fact, several variables enumerated in prior studies (Henderson *et al.* 1995, Henderson 1997, Kim 1999, Ellison and Glaeser 1999) as potentially important location-specific factors are roughly constant between locations in South-Holland. For example, the province is small enough that workers can live in one zip code area and commute to work in almost any other (as well as to areas in other provinces) using either public or private transportation modes, and in fact they do. Thus, wage rates within a sector would be uniform and there is little need to control for labor market characteristics such as job search efficiency, level of education, percent of workers with particular skills, or percent of workers who are union members. Additionally, differences between locations in energy prices, taxes, environmental amenities (such as climate), environmental regulations, and cultural aspects are quite small. Land use patterns and zoning regulations, however, do vary between zip codes. Controls for these and other location-specific factors within the province are discussed later on.

Does employment growth associated with agglomeration economies occur in localized areas within a city? One perspective on this question can be obtained by looking at whether the same set of factors determining employment growth at the city level also are at work at the zip code level in a single urbanized area. If so, this outcome would suggest that agglomeration economies foster growth in a relatively small area nearby to where they are generated. On the other hand, a finding that employment growth in a zip code is unrelated to agglomerative factors there may, in light of findings in the previous section, indicate that we need to look across a larger area (perhaps a whole city) to see their effects. Additionally, an outcome that different agglomeration economies are associated with growth at the zip code level, as compared to the city level, could suggest a problem with spatial aggregation bias. In particular, suppose that we divide a city into a number of areas (perhaps, zip codes), each of which with the same number of employees and each completely specialized in the output of goods produced by a different industry. Thus, each area would have high industrial concentration and no industrial diversity, but from the standpoint of the city as a whole, indicators of industrial concentration and diversity would tell a different story.

These ideas are investigated in two steps: (i) we estimate a regression shown in Table 3, Column (1) that is specified as closely as possible to both the work of Glaeser *et al.* and to the regression for Dutch cities shown in Table 2, Column (1) and (ii) we extend this model to allow for the possibility that agglomeration economies in one zip code affect employment growth in neighboring zip codes. Establishment data used in these regressions were aggregated into zip code industries and the six with largest employment initially were selected. However, because zip code areas are small, some of the six largest sectors had little employment making employment growth rate calculations problematic. Therefore, those with less than 50 employees in the base year were (arbitrarily) excluded from the analysis. This minimum employment cut-off reduced the number of zip code industries in the data set from 2496 (416x6) to 1797. Also, the dependent variable measured employment growth of a zip code-industry over the period 1988-97 and explanatory variables measure characteristics present in the base year (1988). The agglomeration economy indicators are defined in the same way as for Dutch cities, except that they now are measured at the zip code level and the relative specialization and competition variables compare a zip code to South-Holland, rather than to the Netherlands. Additional controls measure initial employment in a zip code-industry, growth of the industry in South-Holland, but outside the zip code, and initial (regional) wages in an industry. Also, whether a zip code was predominantly a work area in 1988 (rather than a residential area) measures initial land use patterns, and distance (in kilometers) of a zip code from the Rotterdam harbor, from Amsterdam, and from Utrecht captures the spatial layout of the province. Column (4) in Table A1 presents means and standard deviations of all variables used in the Table 3 regressions.

This regression is estimated by ordinary least squares and uses the employment growth rate for all establishments in a zip code-industry as the dependent variable. The R^2=0.128, a lower value than that obtained when running the corresponding regression for Dutch municipalities. The small size of the zip code areas may be partly responsible here. Many zip code industries have fewer than 100 employees, so relatively small absolute employment changes over the sample period can produce relatively large changes in growth rates. Standard errors, obtained using White's correction for heteroskedasticity, are shown beneath coefficients estimates. Diagnostics for spatial autocorrelation (the Lagrange Multiplier test for spatial error dependence) suggests that spatial autocorrelation is not a problem.

Coefficient estimates from this regression are similar in many respects to those found earlier for Dutch municipalities and show that industry specialization retards growth, while industrial diversity and competition foster growth. Effects of these three agglomeration measures are somewhat smaller in absolute value than those for Dutch city-industries shown in Table 2, Column (1). Also, in the Table 3, Column (1) regression, we controlled for land use patterns, zoning, and the general spatial layout of the province by including a dummy variable indicating whether a zip code was classified as a work area in 1988 and variables indicating distance of a zip code from the Rotterdam harbor, from Amsterdam, and from Utrecht. Employment growth tended to be greater in work areas and away from the Rotterdam harbor, but distance from Amsterdam and Utrecht did not matter. Regarding performance of other controls used in Table 3, Column (1), coefficients of growth in a zip code-industry outside the zip code was positive and significantly different from zero at conventional levels. High initial wages, on the other hand, retard growth and initial employment in a zip code-industry was an unimportant determinant of zip code-industry employment growth. These results again support the idea that Jacobs-type externalities foster growth, while MAR externalities tend to slow it down and suggest that the same types of agglomeration economies found to be important at the city level also are important at the zip code level.

Just as in analyzing the Dutch city-industry data, we altered the Table 3, Column (1) regression in a number of ways to see whether the above outcome is sensitive to changes in specification. Results are largely unaffected by adding 31 industry dummy variables, by raising the zip code-industry employment cut-offs for inclusion in the data set to 75, 100, 125 and 250 employees, or by excluding those industries for which the theory of knowledge spillovers are less likely to apply; see Table A3 Columns (1)-(3).[10] The same measures of absolute scale for industrial specialization and competition performed unevenly as occurred with the Dutch city data. Also, as an alternative strategy to control for land use patterns, we included dummy variables for each of the 69 municipalities in South-Holland.[11] As shown in Appendix A, Table A2, Column (1), when municipality dummies are included together with the industry dummies, coefficient estimates for industrial specialization, industrial diversity, and competition are statistically significant and similar in magnitude to those reported in Table 3, Column (1). This outcome parallels findings for Dutch cities discussed in the previous section.

The regression in Table 3, Column (2) allows for agglomerative factors in one zip code to affect growth in other zip codes. In addition to the explanatory variables included in the Table 3, Column (1) regression, we also included gravity variables for the three measures of agglomerative factors and for initial employment and then estimated the equation in a spatial lag framework. This approach allows for agglomeration economies to have both direct and indirect effects on employment growth in other locations as previously discussed in Section 3. As shown in Column (2), coefficients of "own-zip code" agglomeration economy variables are virtually unchanged as compared to those presented in Column (1), an outcome that again supports the Jacobs hypothesis.[12] Also, the spatial lag coefficient and coefficients of the gravity variables are not significantly different from zero at conventional levels, an outcome that is remarkable in light of the fact that South-Holland zip codes average less than 6 km^2 in size. This result suggests that effects of agglomeration economies in one location on employment growth at another location die out quickly with distance. An important implication of this result together with the finding that effects of agglomeration economies on employment growth are similar at the city and sub-city levels is that these effects appear to be highly localized and may well occur on a geographic scale that is smaller than a city.

Finally, because previous studies were unable to track movements of individual establishments over time, attention generally has been focused on overall industry employment growth. This variable, however, includes both growth of existing establishments as well as employment changes due to establishment births and relocations. These three sources of employment growth may be driven by different sets of explanatory factors. To better isolate effects on employment growth arising from a location's composition of economic activity, the regression presented in Table 3, Column (3) is specified like the one presented in Table 3, Column (1), however, it uses the employment growth rate of only those establishments originally present in a zip code-industry in 1988 as the dependent variable. In 1997, these "old" establishments represented 64% of all establishments and their employment represented 83% of employment in all establishments. Thus, establishments arising from births and relocations are relatively small compared to old establishments and old establishments account for most of the employment at the end of the period. As shown in Table 3, results for old establishments are quite similar to those for all establishments and again support the Jacobs hypothesis. Although this result might be expected because old establishments account for most of the employment at the end of the period, it still suggests that results presented in Table 3, Column (1) are not driven by the inclusion of new and relocating establishments.[13]

5. Discussion and conclusions

This paper presents empirical evidence on the role of agglomeration economies on employment growth by extending the work of Glaeser *et al*. (1992) and using data from the Netherlands. The agglomeration economies studied are thought to affect growth through the planned or unplanned transfer of knowledge within or between industries. Our data permit us to analyze industry growth in Dutch cities as well as in very small postal zip code areas in one of 12 Dutch provinces, the heavily urbanized province of South-Holland. At both of these levels of geographic aggregation, we find that local industrial diversity and the presence of many small establishments (interpreted as a measure of local competition), but not local industrial specialization, tend to promote growth. Our results, which are consistent with those of Glaeser *et al*., suggest that (1) the same set of factors that promote growth at the city level also promote growth in small areas with cities and (2) knowledge is not necessarily industry-specific and that ideas generated in one sector may also be fruitfully applied in others. These results turn out to be robust against many, but not all, changes in model specification that we investigated.

We also look at the possibility that agglomeration economies in one location might affect employment growth in other locations. As might be expected from prior analyses of patents and innovations, spillover effects between cities appear to be unimportant. A remarkable result, however, is that this same result emerges in the analysis of South-Holland zip codes. Despite the fact that these zip codes average only about 6 km^2 in size, there is little evidence that agglomeration economies in any one of them has much effect on employment growth elsewhere. An implication of our analyses of growth both within and between cities and zip codes is that agglomeration economies may well operate on a geographic scale that is much smaller than a city. This possibility might usefully be the subject of additional empirical studies for other countries because the appropriate geographic scale at which to study effects of agglomeration economies has seldom been directly investigated. Finally, the South-Holland data permit us to look at employment growth in *existing* firms (as opposed to industry growth which is also determined by firm entry and exit). In this case, we again find that local industrial diversity and local competition are important to their growth.

ENDNOTES

[*]Daan van Soest is grateful to the Netherlands Organization for Scientific Research (NWO) for financial support of the PRET research program. Shelby Gerking acknowledges the hospitality of CentER at Tilburg University where this paper was written. He would also like to thank NWO for financial support (visiting grant B46-386). This paper has benefited from presentations at conferences of the Regional Science Association, the European Economic Association, and from comments from seminar participants at Colorado State University, and the Universities of Groningen, Tilburg, and Wyoming, and from a large number of constructive comments by Bert van der Knaap, Raymond Florax, Henri de Groot, Jan Lambooy, John List, Bart Los, Richard Shearmur, and Will Strange. Please send all comments to Daan van Soest, Tilburg University, Department of Economics, P.O. Box 90153, 5000 LE Tilburg, The Netherlands; e-mail: d.p.vansoest@kub.nl.

[1]Identification of new versus old establishments also is possible using Dun and Bradstreet Marketplace data for the U.S. at the zip code level (see Rosenthal and Strange 2000).

[2]In their study of industry growth in U.S. cities, Henderson *et al.* included all cities even though many of them had relatively few establishments present. Therefore, they faced the additional problem that because of federal disclosure rules, industry employment values were censored in as many as 30% of the cities studied. Censoring is not a problem in our data sets; see below.

[3]This population cut-off value was chosen because it roughly corresponded to the size of the smallest cities included in the Glaeser *et al.* analysis.

[4]For exact specifications of the agglomeration economy indicators, see Appendix B.

[5]We also tried two alternative measures of industrial diversity in this regression; Gini and Hirschman-Herfindahl indices of inequality of industry employment shares. Coefficients of these variables were negative and differed significantly from zero at conventional levels,

again indicating that greater industrial diversity in a municipality leads to faster employment growth. These results as well as others discussed later that are not presented in tables are available from the authors on request.

[6]It may be interesting to note that had we constrained coefficients of the national industry growth rate to unity, resulting estimates would be interpreted as an explanation of the differential shift term in shift-share analysis (see Dunn 1960 and Perloff *et al.* 1960). The differential shift term measures the extent to which an industry in a region (or city) grows faster or slower than it does on average in a broader geographic area. In an earlier day, there was considerable debate among regional scientists as to what determines the value of the differential shift term (see Houston 1967). The Glaeser *et al.* results suggest that Jacobs-type externalities are important in this regard.

[7]In the presence of industry dummy variables, inclusion of the wage rate and industry growth outside the city causes a multicollinearity problem.

[8]We also tried lowering the value of population needed for inclusion in the sample from 50,000 persons to 20,000 persons in order to include a larger number of smaller municipality-industries in the analysis. This alteration also weakened the performance of industrial diversity measures.

[9]The comparison of manufacturing industries with non-manufacturing industries could not be carried out using the "top 6" sample because it contained too few manufacturing industries to warrant separate analysis. Also, we tried running a regression specified identically to the one in Table 2, Column (1) that included manufacturing and business services and excluded many local service sectors such as hotels, wholesale and retail trade, education, and health care. Results here were similar to those described for manufacturing industries and are reported in Table A2, Column (4).

[10]The results are also invariant with respect to using a Hirschman-Herfindahl index as a measure for industrial diversity. However, when using the Gini coefficient as a measure of diversity, no statistically significant relationship was found.

[11]In principle, we could include dummies for each of the 416 zip code areas in place of dummy variables for the 69 municipalities. This step was not taken because of the large number of explanatory variables that would have to be included in the regressions.

[12]When estimating the Table 3, Column (2) specification by ordinary least squares, coefficients of the gravity variables are not significantly different from zero. Also, when estimating the Table 3, Column (1) equation in a spatial lag format (without the gravity variables), the spatial lag coefficient is not significantly different from zero at conventional levels.

[13]These conclusions also hold when including industry and municipality dummies, as can be seen from the fourth column in Table A3.

REFERENCES

Anselin, L. 1988. *Spatial econometrics: Methods and models*. Kluwer Academic Publishers, Dordrecht.

Anselin, L. 1995. *SpaceStat. A software program for the analysis of spatial data (version 1.80).* Regional Research Institute, West Virginia University, Morgantown WV.

Anselin, L., Varga, A. and Acs, Z., 1997. Local geographic spillovers between university research and high technology innovations, *Journal of Urban Economics* 42, 422-48.

Arrow, K.J., 1962. The economic implications of learning by doing, *Review of Economic Studies* 29, 155-173.

Audretsch, D.B. and Feldman, M.P., 1996. R&D spillovers and the geography of innovation and production, *American Economic Review* 86, 630-40.

Beardsell, M. and Henderson, J.V., 1999. Spatial evolution and the computer industry in the USA, *European Economic Review* 43, 431-56.

Black, D. and Henderson, J.V., 1999a. Spatial evolution of population and industry in the United States, *American Economic Review* 89, 321-27.

Black, D. and Henderson, J.V., 1999b. A theory of urban growth, *Journal of Political Economy* 107, 252-284.

Combes, P., 2000. Economic structure and local growth: France, 1984-1993, *Journal of Urban Economics* 47, 329-355.

Dunn, E.S., 1960. A statistical and analytical technique for regional analysis, *Papers and Proceedings, Regional Science Association* 6, 97-112.

Ellison, G. and Glaeser E.L., 1999. The geographic concentration of industry: Does natural advantage explain agglomeration, *American Economic Review* 89, 311-16.

Feldman, M.P. and Audretsch, D.B., 1999. Innovation in cities: Science-based diversity, specialization, and localized competition, *European Economic Review* 43, 409-29.

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

Glaeser, E.L., 1999. Learning in cities, *Journal of Urban Economics* 46, 254–277.

Hanson, G.H., 2001, Scale economies and the geographic concentration of industry, *Journal of Economic Geography* 1, 255-276.

Henderson, J.V., Kuncoro, A., and Turner, M., 1995. Industrial development in cities, *Journal of Political Economy* 103, 1067-85.

Henderson, J.V., 1997. Externalities and industrial development, *Journal of Urban Economics* 42, 449-71.

Houston, D.B., 1967. The shift-share analysis of regional growth: a critique, *Southern Economic Journal* 33, 577-81.

Jacobs, J., 1969. *The economy of cities*. Vintage, New York.

Jaffe, A.B., 1989. Real effects of academic research, *American Economic Review* 79, 957-70.

Jaffe, A.B., Trajtenberg, M., and Henderson R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics* 63, 577-98.

Kim, S., 1999. Regions, resources, and economic geography: Sources of U.S. regional comparative advantage, 1880-87, *Regional Science and Urban Economics* 29, 1-32.

Lucas, R.E., 1988. On the mechanics of economic development, *Journal of Monetary Economics* 22, 3-42.

Lucas, R.E., 1993. Making a miracle, *Econometrica* 61, 251-72.

Marshall, A., 1890. *Principles of economics*. Macmillan, London.

Perloff, H.S., Dunn E.S., Lampard E.E., and Muth, R.F., 1960. *Regions, Resources and Economic Growth*. Johns Hopkins Press, Baltimore.

Porter, M., 1990. *The competitive advantage of nations*. Free Press, New York.

Romer, P.M., 1986. Increasing returns and long run growth, *Journal of Political Economy* 94, 1002-1037.

Rosenthal, S. and Strange, W., 2000. Geography, industrial organization and agglomeration, Syracuse University, Syracuse, NY.

RPD, 1998. *Werklocaties 1998*. Rijksplanologische Dienst, Rapportage bedrijventerreinen en voorradige kantoorlocaties (IBIS), Den Haag.

APPENDIX A: DESCRIPTION OF THE DATA

The data that are used in this paper are derived from various sources. The most important ones are the longitudinal datasets of the Firm Register South-Holland (BZH) and the National Information System on Employment (LISA, the nationwide firm register in which the BZH is embedded). Registration is at the level of individual firms, including detailed information on location (6-digit zip code) and activity (5-digit SBI93-code, completely consistent with NACE and ISIC industrial classifications). The data concerning agricultural employment were derived from the Agricultural Statistics of the Dutch Central Bureau of Statistics (CBS) on the municipality-level, and localized to 4-digit zip codes on the basis of the Land Use Statistics (Bodemstatistiek CBS, function agriculture). Various other sources have been consulted to construct and verify the remaining variables, like data from the Chamber of Commerce in 1990 and CBS statistics on (aggregate) employment development. The national (and, in case of South-Holland, regional) industry-specific wage rates were calculated from CBS Regional Economic Data. Whereas for the Netherlands only the nation's average wage rate is available for each individual industry, information on industry-specific wage rates is available for each of the five COROP regions that together make up the province of South-Holland. Information on each zip-code's economic function (whether it is predominantly a residential area or a working area) was obtained from RPD (1998).

All variables measuring physical distances (such as the distances between the zip codes necessary for the potential model and the spatial lag models, as well as the distance to Rotterdam harbor, Utrecht and Schiphol) were constructed using Atlas*GIS, ArcInfo and ArcView geographical information systems. Additional calculations (of distance- and weight matrices) were carried out within the statistical package SpaceStat (Anselin 1995). A detailed description of the data and the verifications applied can be obtained from the authors.

A summary of the data sets is given in Table A1, which presents the means and standard deviations for the dependent and explanatory variables used in this paper. Six data sets have been constructed. For the Netherlands, the following four have been used: (1) a data set consisting of the 6 largest industries in the 57 cities, (2) a data set consisting of the 12 largest industries in the 57 cities, (3) a data set consisting of the manufacturing industries (if they are among the largest 12 in their city) and (4) a similar one for the non-manufacturing industries. Two additional data sets have been used for the South-Holland analyses, which contain information largest six industries in each (4-digit) zip code with initial employment levels at least 50 (Column 5) or 250 employees (Column 6). In the main text, some regression results are discussed that are obtained from alternative data sets which are not presented here; additional information on these are available upon request.

TABLE A1: MEANS AND STANDARD DEVIATIONS (in parenthesis) OF THE VARIABLES IN THE VARIOUS DATA SETS

	NETHERLANDS				SOUTH-HOLLAND	
VARIABLE	All industries Top-6 (1)	All industries Top-12 (2)	Manufacturing industries Top-12 (3)	Non-manufacturing industries Top-12 (4)	All industries (1988 employment≥50) (5)	All industries (1988 employment≥250) (6)
Log(Employment 1997 divided by base-year employment) in the city/zip code-industry	0.044 (0.501)	0.010 (0.576)	-0.295 (0.974)	0.116 (0.272)	-0.264 (1.067)	-0.330 (1.077)
Log(Employment 1997 divided by employment 1988) in firms in the zip code-industry that already existed in 1988					-0.524 (1.186)	-0.562 (1.243)
Log(Netherlands/South-Holland employment in 1997, divided by initial Netherlands/South-Holland employment) in the industry *outside* the city/zip code	0.127 (0.098)	0.100 (0.116)	-0.025 (0.125)	0.144 (0.073)	0.082 (0.166)	0.087 (0.167)
Annual industry wage rate in the base year (in thousands of base year Dutch guilders)	46.671 (8.800)	47.426 (9.154)	47.256 (5.621)	47.486 (10.104)	46.256 (8.434)	46.078 (8.417)
Employment in the city/zip code-industry in the base year	4836.225 (5811.472)	3553.544 (4906.419)	2038.655 (2162.388)	4082.41 (5457.499)	402.132 (582.569)	757.389 (742.563)
Employment in the municipality industry in 1988					6609.251 (9504.878)	8614.596 (10629.29)
Distance-weighted sum of industry employment in cities/zip codes in the base year *outside* the location under observation	6988.816 (3862.437)				5870.308 (4257.261)	6747.502 (4748.918)
Dummy variable indicating presence in the Randstad	0.386 (0.487)	0.386 (0.487)	0.328 (0.471)	0.406 (0.492)		
Dummy variable indicating presence in the country's periphery	0.228 (0.420)	0.228 (0.420)	0.266 (0.443)	0.215 (0.411)		
Dummy variable indicating whether the zip code's function in 1988 is predominantly industrial as opposed to residential					0.263 (0.440)	0.352 (0.478)
Variable indicating the distance to Amsterdam					51393.010 (11850.717)	51380.661 (11060.560)
Variable indicating the distance to Rotterdam harbor					21465.484 (10532.938)	21034.532 (10290.864)
Variable indicating the distance to Utrecht					52615.179 (10004.027)	52963.928 (9124.551)
City/zip code-industry's share of city/zip code employment relative to industry's share of Netherlands/South-Holland employment in the base year	2.135 (4.760)	1.845 (3.571)	2.887 (4.552)	1.481 (3.097)	4.823 (13.582)	5.365 (9.381)
City/zip code-industry's share of city/zip code employment	0.092 (0.040)	0.066 (0.040)	0.048 (0.024)	0.073 (0.043)	0.147 (0.128)	0.196 (0.144)
Distance-weighted sum of the industry's employment in all cities/zip codes divided by the distance-weighted sum of total employment in all cities/zip codes *outside* the location under observation, relative to industry's share of the Netherlands' or South-Holland's employment in the base year	0.543 (0.159)				0.518 (0.259)	0.524 (0.254)
Number of establishments per employee in the city/zip code-industry relative to the number of establishments per employee in the Netherlands/South-Holland industry in the base year	0.806 (0.358)	0.826 (0.393)	0.678 (0.425)	0.878 (0.368)	1.129 (0.924)	0.705 (0.557)
Number of establishments per employee in the city/zip code-industry	0.164 (0.187)	0.092 (0.084)	0.048 (0.045)	0.107 (0.088)	0.115 (0.118)	0.196 (0.144)
Distance-weighted sum of the industry's number of establishments in all cities/zip codes divided by distance-weighted sum of the industry's employees in all cities/zip codes outside the city/zip code under observation, relative to the number of establishments per employee in the Netherlands/South-Holland industry in the base year	1.005 (0.276)				0.986 (0.147)	0.982 (0.171)
City/zip code's other top five (or eleven) industries' share of total city/zip code employment in the base year	0.460 (0.060)	0.730 (0.053)	0.559 (0.191)	0.725 (0.055)	0.590 (0.137)	0.537 (0.132)
Distance-weighted sum of the top six industries' employment in all cities/zip codes divided by the distance-weighted sum of total employment in all cities/zip codes *outside* the location under observation in the base year	0.185 (0.034)				0.333 (0.142)	0.330 (0.130)
Number of observations	342	684	177	507	1797	783

TABLE A2: ANALYSIS OF EMPLOYMENT GROWTH IN DUTCH CITIES (standard errors are presented in parenthesis)

EXPLANATORY VARIABLE	Log(Employment in 1997/Employment in 1991) in the city-industry			
	All industries Top-6 (1)	Manufacturing industries Top-12 (2)	Non-manufacturing industries Top-12 (3)	Selection of industries[a] Top-12 (4)
Constant	0.543 (0.339)	0.831 (1.319)	-0.115 (1.779)	0.189 (0.592)
Log(Netherlands employment in 1997/Netherlands employment in 1991) in the industry outside the city		3.267** (0.692)	0.954** (0.157)	1.997** (0.250)
Annual industry wage rate in 1991		0.029 (0.015)	-0.001 (0.001)	-5.64E-04 (0.250)
Employment in the city-industry in 1991	8.91E-07 (8.95E-06)	-1.60E-05 (3.05E-05)	-7.87E-07 (2.32E-06)	-1.04E-06 (8.58E-06)
Dummy variable indicating presence in the Randstad	0.071 (0.297)	-0.173 (0.152)	-0.050 (0.027)	-0.103 (0.077)
Dummy variable indicating presence in the country's periphery	0.205 (0.187)	-0.358* (0.158)	-0.081** (0.030)	-0.241** (0.083)
City-industry's share of city employment relative to the industry's share of Netherlands employment in 1991	-0.032** (0.007)	-0.070** (0.016)	-6.75E-05 (0.004)	-0.026** (0.007)
Number of establishments per employee in the city-industry relative to the number of establishments per employee in the industry in the Netherlands in 1991	0.252** (0.086)	0.150 (0.164)	0.156** (0.032)	0.215** (0.079)
City's other top five industries' share of total city employment in 1991	-1.851* (0.796)			
City's other top eleven industries' share of total city employment in 1991		-2.889 (1.553)	0.098 (0.233)	-0.528 (0.736)
Industry fixed effects	YES**	NO	NO	NO
Municipality fixed effects	YES**	NO	NO	NO
SUMMARY STATISTICS				
N	342	177	507	395
Adjusted R^2	0.651	0.276	0.127	0.247
LR(SL)	0.145	0.528	0.268	0.465
LM(SE)	1.100	4.596*	1.524	1.951

[a]: All industries except wholesale trade, retail trade, hotels and restaurants, government and social insurance, education, health care, waste disposal services, trade unions, sports and culture, personal services (hair dressers, sauna's etc).

*: Significant at 5%, **: Significant at 1%.

TABLE A3: ANALYSIS OF EMPLOYMENT GROWTH IN ZIP CODE INDUSTRIES IN SOUTH-HOLLAND (standard errors are presented in parenthesis)

EXPLANATORY VARIABLE	Log(Employment in 1997/Employment in 1988) in the zip code industry			
	All firms (1)	All firms, selection of industries[a] (2)	All firms, 250 employees or more (3)	Old firms (4)
Constant	-4.065	-0.349	-0.983	-4.954
	(1.212)	(0.655)	(0.759)	(1.302)
Log(South-Holland employment in 1997/ South-Holland employment in 1988) in the industry outside the zip code		0.757*	0.657	
		(0.299)	(0.351)	
Annual regional industry wage rate in 1988 (in thousands of 1988 Dutch guilders)	0.014	-0.002	-0.004	0.022
	(0.020)	(0.006)	(0.004)	(0.021)
The industry's employment in the zip code in 1988	-1.66E-05	-1.58E-04*	3.19E-05	1.25E-05
	(4.89E-05)	(8.11E-05)	(5.78E-05)	(5.36E-05)
Industry's share of zip code employment relative to industry's share of Netherlands employment in 1988	-0.011**	-0.014**	-0.023*	-0.011**
	(0.003)	(0.003)	(0.011)	(0.002)
Number of establishments per employee in the zip code industry relative to establishments per employee in the industry in South-Holland in 1988	0.181**	0.271**	0.358**	0.189**
	(0.029)	(0.038)	(0.079)	(0.033)
Zip code's other top five industries' share of total zip code employment in 1988	-0.589**	-1.477**	-0.651*	-0.798**
	(0.207)	(0.337)	(0.320)	(0.229)
Dummy variable indicating whether the zip code's function in 1988 is predominantly industrial as opposed to residential	0.327**	0.458**	0.161	0.321**
	(0.072)	(0.088)	(0.096)	(0.082)
Variable indicating the distance to Amsterdam	2.18E-05*	2.15E-06	7.54E-06	2.43E-05*
	(9.91E-06)	(7.02E-06)	(9.36E-06)	(1.10E-05)
Variable indicating the distance to Rotterdam harbor	3.44E-05**	1.12E-05	1.41E-05	4.28E-05**
	(1.03E-05)	(8.10E-06)	(1.01E-05)	(1.18E-05)
Variable indicating the distance to Utrecht	7.13E-06	4.86E-06	4.53E-06	1.24E-05
	(5.87E-06)	(5.03E-06)	(5.38E-06)	(6.92E-06)
Municipality fixed effects	YES**	NO	NO	YES**
Industry fixed effects	YES**	NO	NO	YES**
SUMMARY STATISTICS				
N	1797	968	783	1797
Adjusted R^2	0.217	0.136	0.114	0.205

[a]: All industries except wholesale trade, retail trade, hotels and restaurants, government and social insurance, education, health care, waste disposal services, trade unions, sports and culture, personal services (hair dressers, sauna's etc).

*: Significant at 5%, **: Significant at 1%.

APPENDIX B: CONSTRUCTION OF THE KNOWLEDGE INDICATORS

In line with Glaeser *et al.* (1992), industry concentration, which is interpreted as indicating the extent to which knowledge is industry-specific, is calculated as follows:

$$\textit{Concentration of industry k in location j} \\ = \frac{\textit{employment in industry k in location j/total employment in location j}}{\textit{employment in industry k in the region/total employment in the region}},$$

where "location" is either the zip code (in the South-Holland analyses) or the city (in the Netherlands analyses), and the "region" is respectively South-Holland or the Netherlands. All information used refers to the base year (1991 for the Netherlands, 1988 for South-Holland). With respect to the three hypotheses cited in Section 2, MAR and Porter would predict that concentration is positively correlated with (employment) growth, whereas Jacobs would predict a negative relationship. Similarly, competition is calculated as follows:

$$\textit{Competition faced by industry k in location j} \\ = \frac{\textit{firms in industry k in location j/employees in industry k in location j}}{\textit{firms in industry k in the region/employees in industry k in the region}},$$

and a positive relationship of this variable with employment growth would give support to the hypotheses as formulated by Porter and Jacobs, whereas a negative one would give support to MAR. Finally, the role of economic diversity can be analyzed using the following variable:

$$\textit{Industrial Diversity for industry k in location j} \\ = \frac{\textit{employment in the five other biggest industries in location j}}{\textit{total employment in location j}};$$

As Jacobs argues that knowledge is not industry-specific and that ideas generated in one industry can fruitfully be applied in another, she would predict a negative relationship between lack of diversity and employment growth; if the relationship is positive, this would give additional support for the claim of MAR and Porter that knowledge is predominantly industry-specific.

To allow for knowledge directly spilling over between locations, we calculate distance-weighted versions of the knowledge indicators. They were constructed by replacing the arguments in the standard knowledge indicators by distance-weighted ones. Denoting the distance between two locations *i* and *j* by d_{ij} (as measured in kilometers), the distance-weighted variables are calculated as:

$$\textit{Distance}-\textit{Weighted Concentration of industry } k \textit{ in location } j$$
$$= \frac{\sum_{i\neq j}\frac{1}{d_{ij}}\left(\textit{industry } k\textit{'s employment in location } i\right) / \sum_{i\neq j}\frac{1}{d_{i}}\left(\textit{total employment in location } i\right)}{\textit{industry } k\textit{'s employment in the region/total employment in the region}},$$

$$\textit{Distance}-\textit{Weighted Competition faced by industry } k \textit{ in location } j$$
$$= \frac{\sum_{i\neq j}\frac{1}{d_{ij}}\left(\textit{firms in the industry } k \textit{ in location } i\right) / \sum_{i\neq j}\frac{1}{d_{ij}}\left(\textit{employees in industry } k \textit{ in location } i\right)}{\textit{firms in industry } k \textit{ in the region/employees in industry } k \textit{ in the region}}, \text{ and}$$

$$\textit{Distance}-\textit{Weighted Industrial Diversity of industry } k \textit{ in location } j$$
$$= \frac{\sum_{i\neq j}\frac{1}{d_{ij}}\left(\textit{employment in the six biggest industries in location } i\right)}{\sum_{i\neq j}\frac{1}{d_{ij}}\left(\textit{total employment in location } i\right)}.$$

APPENDIX C: SPATIAL LAG AND SPATIAL ERROR MODELS

In the literature on knowledge externalities and growth (Glaeser *et al.*, Combes 2000), standard OLS regression specifications have been used of the form:

$$y=X\beta+\varepsilon, \tag{C1}$$

in which y denotes an Nx1 vector of location-industry growth rates, X denotes a matrix of observations on a set of location- or industry-specific explanatory variables, β is the coefficient vector and ε denotes a vector of disturbances. In this specification, distance does not play a role; the impact of spillovers is assumed to be fully localized.

One way to account for the role of distance would be to reformulate equation (C1) as a spatial lag model (Anselin 1988), which is specified as follows:

$$y=\rho Wy+X\beta+\varepsilon, \tag{C2}$$

in which matrix W would reflect (inverse) distances between locations (see Appendix B; industries in the same zip code are assumed to be less than one kilometer apart) and the spatial coefficient ρ would index the strength of employment growth linkages over space. An alternative specification to the spatial lag model is the spatial error specification shown in equation (C3):

$$y=X\beta+u; \qquad u=\lambda Wu+\varepsilon. \tag{C3}$$

In equation (C3), u is interpreted as the outcome of a spatial autoregressive process involving a weight matrix (W) and a spatial autoregressive coefficient (λ). The distinction between the spatial lag and spatial error specifications is important because in the former case, growth in one location is linked to growth in other locations, whereas in the latter case, linkages between locations occur via the error generation process.

Note that whereas the spatial error term simply corrects for spatial correlation in the error term, the spatial lag model has a clear economic interpretation. By rewriting (C2) as follows,

$$y=AX\beta+A\varepsilon, \tag{C4}$$

the elements of the matrix $A=(I-\rho W)^{-1}$ show how a change in X in one location affects employment growth rates in other locations after taking all such spatial linkages into account. Column sums of A have the interpretation of spatial multipliers. Of course, A depends on an estimate of ρ and the specification of W. With respect to the ways in which weight matrix W can be constructed, two alternative specifications of have been applied, one that allows for *all* spatial linkages that depend on distance and does not distinguish between industries, and one that allows for links only between a given industry in one location and that same industry in other locations. Both specifications were tried for all regressions mentioned in this paper, but space was never found to be very important. Therefore, we only present the regression results based on the specification that includes all spatial relationships between zip-codes and/or cities.

TABLE 1:

RANKING OF INDUSTRIES BY FREQUENCY OF OCCURRENCE (WITH EMPLOYMENT LEVELS) IN THE NETHERLANDS AND SOUTH-HOLLAND DATA SETS

	Netherlands (1991)[a]			South-Holland (1988)	
Industries	Representation	Employment	Industries	Representation	Employment
Health care	56	388,297	Building and construction	231	66,016
Business services	48	298,897	Retail trade	194	69,997
Retail trade	46	214,531	Business services	185	76,913
Building and construction	45	131,336	Health care	174	103,166
Wholesale trade	37	170,314	Education	165	43,893
Education	24	102,727	Wholesale trade	140	66,063
Government and social insurance	18	104,670	Agriculture and fishery	89	31,393
Furniture industry	6	11,927	Government and social insurance	82	51,945
Publishing and reproduction	5	19,192	Distribution by land	76	22,534
Chemical industry	5	14,397	Metal products industry	35	8,026

[a]Figures shown pertain to the Dutch city-industry "top 6" data set.

TABLE 2: ANALYSIS OF EMPLOYMENT GROWTH IN DUTCH CITIES (standard errors are presented in parenthesis)

EXPLANATORY VARIABLE	Log(Employment in 1997/Employment in 1991) in the city-industry			
	Top 6 (1)	Top 12 (2)	Top 6 Absolute (3)	Top 6 Spatial lag model (4)
Constant	0.147 (0.249)	0.075 (0.308)	0.201 (0.298)	-0.856 (0.428)
Log(Netherlands employment in 1997/Netherlands employment in 1991) in the industry outside the city	1.558** (0.267)	1.924** (0.178)	2.083** (0.266)	1.349** (0.293)
Annual industry wage rate in 1991 (in thousands of 1991 Dutch guilders)	0.001 (0.003)	-1.56E-04 (0.002)	-0.003 (0.003)	0.002 (0.003)
Employment in the city-industry in 1991	-1.91E-06 (4.52E-06)	-1.36E-06 (4.31E-06)	-4.05E-06 (4.73E-06)	-2.69E-06 (4.43E-06)
Distance-weighted sum of the city-industry's employment outside the city under observation in 1991				2.73E-06 (1.08E-05)
Dummy variable indicating presence in the Randstad	-0.043 (0.057)	-0.072 (0.046)	-0.060 (0.060)	-0.165 (0.087)
Dummy variable indicating presence in the country's periphery	-0.097 (0.063)	-0.169** (0.050)	-0.118 (0.065)	0.046 (0.110)
City-industry's share of city employment relative to the industry's share of Netherlands employment in 1991	-0.023** (0.005)	-0.027** (0.006)		-0.039** (0.007)
City-industry's share of city employment in 1991			0.252 (0.764)	
Distance-weighted sum of the industry's employment in all cities divided by the distance-weighted sum of total employment in all cities outside the city under observation, relative to industry's share of Netherlands' employment in 1991				0.085 (0.162)
Number of establishments per employee in the city-industry relative to the number of establishments per employee in the Netherlands industry in 1991	0.208** (0.074)	0.171** (0.052)		0.153* (0.076)
Number of establishments per employee in the city-industry in 1991			-0.298* (0.142)	
Distance-weighted sum of the industry's number of establishments in all cities divided by distance-weighted sum of the industry's employment in all cities outside the city under observation relative to the number of establishments per employee in the Netherlands industry in 1991				0.399** (0.120)
City's other top five industries' share of total city employment in 1991	-0.881* (0.431)		-0.483 (0.479)	-0.891* (0.438)
City's other top eleven industries' share of total city employment in 1991		-0.370 (0.405)		
Distance-weighted sum of the top six industries' employment in all cities divided by the distance-weighted sum of total employment in all cities outside the city under observation in 1988				3.304 (1.950)
Spatially lagged dependent variable				-0.088 (0.181)
SUMMARY STATISTICS				
N	342	684	342	342
Adjusted R^2	0.227	0.266	0.169	0.276
LR(SL)	0.053	1.397	1.635	0.220
LM(SE)	1.482	0.001	0.007	0.423

*: Significant at 5%, **: Significant at 1%.

TABLE 3: ANALYSIS OF EMPLOYMENT GROWTH IN ZIP CODE INDUSTRIES IN SOUTH-HOLLAND (standard errors are presented in parenthesis)

EXPLANATORY VARIABLE	Log(Employment in 1997/Employment in 1988) in the zip code industry		
	All establishments (1)	All establishments Spatial lag model (2)	Old establishments (3)
Constant	-0.519 (0.404)	-0.459 (0.556)	-0.944* (0.464)
Log(South-Holland employment in 1997/ South-Holland employment in 1988) in the industry outside the zip code	1.037** (0.220)	0.960** (0.158)	0.726** (0.241)
Annual regional industry wage rate in 1988 (in thousands of 1988 Dutch guilders)	-0.007** (0.002)	-0.005* (0.0029)	-0.008** (0.003)
The industry's employment in the zip code in 1988	-6.69E-05 (4.71E-05)	-9.16E-05 (5.10E-05)	-5.77E-05 (5.34E-05)
Distance-weighted sum of the industry's zip code employment in 1988 *outside* the zip-code under observation		1.19E-05 (7.35E-06)	
Industry's share of zip code employment relative to industry's share of Netherlands employment in 1988	-0.015** (0.003)	-0.014** (0.002)	-0.015** (0.003)
Distance-weighted sum of the industry's employment in all zip codes divided by the distance-weighted sum of total employment in all zip codes *outside* the zip-code under observation, relative to industry's share of South-Holland's employment in 1988		0.066 (0.214)	
Number of establishments per employee in the zip code industry relative to establishments per employee in the industry in South-Holland in 1988	0.197** (0.030)	0.183** (0.029)	0.187** (0.032)
Distance-weighted sum of the industry's number of establishments in all zip codes divided by distance-weighted sum of the industry's employment in all zip codes *outside* the zip-code under observation, relative to the number of establishments per employee in the industry in South-Holland in 1988		-0.255 (0.212)	
Zip code's other top five industries' share of total zip code employment in 1988	-0.817** (0.205)	-0.804** (0.197)	-1.042** (0.229)
Distance-weighted sum of the top six industries' employment in all zip codes *outside* the zip-code under observation divided by the distance-weighted sum of total employment in all zip codes in 1988		-0.370 (0.417)	
Dummy variable indicating whether the zip code's function in 1988 is predominantly industrial as opposed to residential	0.267** (0.065)	0.269** (0.059)	0.240** (0.074)
Variable indicating the distance to Amsterdam	6.20E-06 (4.72E-06)	9.93E-06 (5.89E-06)	7.85E-06 (5.34E-06)
Variable indicating the distance to Rotterdam harbor	1.49E-05** (5.25E-06)	1.75E-05** (6.98E-06)	2.06E-05** (6.00E-06)
Variable indicating the distance to Utrecht	2.61E-06 (3.18E-06)	3.34E-06 (3.56E-06)	5.87E-06 (3.59E-06)
Spatially lagged dependent variable		0.348 (0.318)	
SUMMARY STATISTICS			
N	1797	1797	1797
Adjusted R^2	0.128	0.137	0.092
LR(SL)	0.432	0.783	0.555
LM(SE)	0.016	1.086	0.640

*: Significant at 5%, **: Significant at 1%.