

Growth within a Dutch City: The Spatial Impact of Agglomeration Externalities^{*}

Daan P. van Soest¹, Shelby Gerking² and Frank G. van Oort³

¹Department of Economics and CentER, Tilburg University, Tilburg, The Netherlands

²Department of Economics, University of Central Florida, Orlando, United States

³Department of International Economics and Economic Geography, University of Utrecht, Utrecht, The Netherlands & Institute for Spatial Research (RPB), The Hague, The Netherlands

October 4, 2005

Abstract

This paper examines the extent to which agglomeration economies in one location affect employment growth and establishment births using data from the Dutch province of South-Holland. The data are of particular interest because they represent a census, rather than a sample, of all establishments and the location of establishments can be pinpointed to within 416 (postal) zip code areas averaging less than 6 km² in size. Results suggest that agglomeration economies positively affect employment growth and the location of new establishments, but with the possible exception of manufacturing, this effect dies out quickly with distance. Thus, the main finding is that for many industries, agglomerative forces may well operate at a geographic scale that is smaller than a city.

Key words: endogenous growth, agglomeration economies, spatial analysis

JEL codes: O12, O18, R11, R30.

1. INTRODUCTION

Endogenous growth theory emphasizes the role of agglomeration economies as a crucial factor leading to external economies of scale in production (Romer 1986, Lucas 1988). An important, although unresolved, aspect of this theory is the extent to which economic effects of agglomeration economies decay over distance. For example, while the growth literature assumes that agglomeration economies operate at the aggregate or national level, urban economists generally regard effects of these externalities as more localized. As discussed more fully by Rosenthal and Strange (2003a), empirical findings on this issue are mixed as well. Using data from Japanese prefectures, Dekle and Eaton (1999), find that in manufacturing, effects of agglomeration economies in a single location spread nationwide. Ciccone and Hall (1996) find that increased population density enhances labor productivity, but by using U.S. states as their geographic unit of analysis, implicitly assume that this effect spreads beyond the boundaries of cities. Rosenthal and Strange (2003b), on the other hand, use data from U.S. zip codes, finding that agglomeration economies have much more localized effects. This finding of localized effects of agglomeration economies comports with related studies of employment (e.g., Henderson 2003) as well as with results from studies of geographic spillovers of knowledge that occur through innovations and patents (Jaffe 1989, Jaffe, Trajtenberg, and Henderson 1993, Audretsch and Feldman 1996, Anselin, Varga, and Acs 1997). In any case, the geographic effects of agglomeration economies on economic growth has broad implications ranging from the development of endogenous growth

models to practical conclusions that might be drawn by policy makers regarding urban planning.

This paper measures the extent to which agglomeration economies in one location contribute to growth at other locations by examining detailed data from South-Holland (*Zuid-Holland*), the most densely populated province of the Netherlands (see Figures 1 and 2). South-Holland is approximately 2350 km², covers a large part of the country's core economic area (the 'Randstad'), has a high population density (about 1190 persons/km²), and can be viewed as a single urbanized area. Available data represent a census, not a sample, of all establishments located there and are used to analyze both overall employment growth and establishment births. The data are not subject to disclosure restrictions that have hampered other studies (i.e., Henderson, Kuncoro, and Turner 1995) and location of establishments can be pinpointed within 416 contiguous (postal) zip code areas averaging less than 5.65 km² in size. Thus, the extent of geographic spillovers of agglomeration economies can be more readily identified here than in prior studies. In fact, the average size of the zip code areas studied by Rosenthal and Strange (2003b) is more than 100 times the average size of the spatial units considered here. Results suggest that with the possible exception of manufacturing, spatial effects of agglomeration economies fade quickly with distance and that the geographic scope of these externalities is much smaller than a city.

2. CONCEPTUAL FRAMEWORK

The conceptual framework used to investigate effects of agglomeration economies on economic activity is summarized in equation (1),

$$Y_{jk}=f(A_{jk},A_{jo},Z_k)+\varepsilon_{jk}, \quad (1)$$

where Y_{jk} denotes a measure of growth of economic activity over the period 1988–1997 in industry j , zip code area k , A_{jk} denotes a vector of variables reflecting agglomeration economies in industry j , zip code area k measured in 1988, A_{jo} denotes a vector of variables reflecting agglomeration economies at other locations in 1988, and Z_k denotes a vector of controls. In the analysis presented in Section 4, Y_{jk} is measured in two ways: as employment growth, or as new establishment births. The purpose of estimating equation (1) is to determine the extent to which agglomeration economies within a zip code and outside a zip code affect these two measures of economic activity.

When employment growth is used, methods applied follow Glaeser *et al.* (1992), in that the largest zip code–industries are analyzed. The alternative approach of looking at the growth of industries one–at–a–time (Henderson *et al.* 1995) is less restrictive in that it allows effects of agglomerative forces to differ by industry and does not limit consideration to only the largest industries in a zip code. This approach was not pursued using detailed industries, however, because zip code–industry employment totals frequently are quite small, making calculations of growth rates problematic. However, results of a similar analysis for three broad sectors (producer services, consumer services, and manufacturing) are reported. These three sectors also are analyzed when Y_{jk} is measured as the number of births of new establishments. Studying new establishment births follows on the work of Rosenthal and Strange (2003a), and has the advantage that entrepreneurs take location characteristics as given when deciding where to do business.

In equation (1), the vector A_{jk} contains measures of widely studied possible sources of external economies of scale pertaining to zip code j , industry k . These measures are similar to those used by Glaeser *et al.* (1992) and can be derived from a formal economic

model (see de Lucio *et al.* (2002)). Concentration of an industry in a zip code is measured using the location quotient shown in equation (2). The location quotient is the percentage of a zip code's employment in an industry in 1988 divided by the corresponding South-Holland percentage in 1988.

Concentration of industry k in location j

$$= \frac{\text{employment in industry } k \text{ in location } j / \text{total employment in location } j}{\text{employment in industry } k \text{ in the region} / \text{total employment in the region}}. \quad (2)$$

This variable is expected to contribute positively to employment growth and to births of new establishments if, for example, local inter-establishment movements of knowledge and personnel are important.

Local competition is measured by the 1988 ratio of establishments per worker in a zip code–industry to establishments per worker in that industry in South-Holland shown in equation (3).

Competition faced by industry k in location j

$$= \frac{\text{firms in industry } k \text{ in location } j / \text{employees in industry } k \text{ in location } j}{\text{firms in industry } k \text{ in the region} / \text{employees in industry } k \text{ in the region}}. \quad (3)$$

Larger values of this measure mean that establishments in a zip code–industry are relatively small as compared to the size of establishments in that industry in the Province. 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 (2003a), who view it as a broader measure of industrial organization. In any case, this variable is included to test whether the presence of small establishments in an industry contributes to faster employment growth and to fostering establishment births.

The lack of industrial diversity in a zip code is measured by the percentage of employment in 1988 in the largest five industries, excluding the one under observation as shown in equation (4).

Lack of industrial diversity for industry k in location j

$$= \frac{\text{employment in the five other biggest industries in location } j}{\text{total employment in location } j}. \quad (4)$$

Larger values of this percentage suggest lower levels of industrial diversity, and this variable is included to reflect spillovers between industries in a region. If these spillovers are important, there should be a negative relationship between lack of diversity and both employment growth and new establishment births in an industry.

In equation (1), A_{jo} denotes distance-weighted versions of the three agglomeration indicators just discussed. These were calculated according to equations (5)–(7) so as to allow for spillovers of agglomeration economies to zip code j from all other zip codes. In these equations, distance (in kilometers) between two locations i and j is denoted by d_{ij} . Thus, agglomeration in nearby zip codes is weighted more heavily than agglomeration in zip codes that are farther away.

Distance – Weighted Concentration of industry k in location j

$$= \frac{\sum_{i \neq j} \frac{1}{d_{ij}} (\text{industry } k \text{'s employment in location } i) / \sum_{i \neq j} \frac{1}{d_i} (\text{total employment in location } i)}{\text{industry } k \text{'s employment in the region} / \text{total employment in the region}} \quad (5)$$

Distance – Weighted Competition faced by industry k in location j

$$= \frac{\sum_{i \neq j} \frac{1}{d_{ij}} (\text{firms in the industry } k \text{ in location } i) / \sum_{i \neq j} \frac{1}{d_{ij}} (\text{employees in industry } k \text{ in location } i)}{\text{firms in industry } k \text{ in the region} / \text{employees in industry } k \text{ in the region}} \quad (6)$$

Distance – Weighted Lack of industrial diversity of industry k in location j

$$= \frac{\sum_{i \neq j} \frac{1}{d_{ij}} (\text{employment in the six biggest industries in location } i)}{\sum_{i \neq j} \frac{1}{d_{ij}} (\text{total employment in location } i)}.$$
(7)

Nine variables are included as controls (Z_k): the 1988 industry wage rate, total employment in the zip code–industry in 1988 and its distance-weighted counterpart, total employment in the zip code outside the industry under consideration in 1988 and its distance-weighted counterpart, a dummy variable indicating whether the zip code was primarily industrial as opposed to residential in 1988, and variables indicating distance in kilometers to Amsterdam, the Rotterdam harbor, and Utrecht. These variables were included to capture the scale of activity in a zip code–industry, the overall scale of economic activity in a zip code, and the spatial layout of economic activity in the province.

In estimating equation (1), heterogeneity among locations and possible endogeneity of explanatory variables are potential problems. Regarding heterogeneity, the small size of South–Holland provides a natural control for some, but not all, important location–specific factors that affect the growth or attraction of economic activity. 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. However, unmeasured land use patterns, zoning regulations, access to raw materials, and prior input choices of existing establishments do vary between zip codes and may bias the resulting estimates if they remain uncontrolled.

In the employment growth analysis, one approach to dealing with remaining unmeasured heterogeneity problem is study long-term growth so as to minimize the importance of location-specific factors and so that all establishments are effectively new. This approach taken in the Glaeser *et al.* (1992) study, however, is not available here because South-Holland data comparable to those used here did not exist prior to 1988. Instead, dummy variables (fixed effects) are included for each of the 69 municipalities and 31 industries represented in the data set. In the establishment birth analysis, estimation of equation (1) in a negative binomial framework is used to account for unmeasured heterogeneity between locations. This estimator is a standard choice when analyzing count data and generalizes the Poisson model by introducing (in this analysis) a multiplicative random, zip code specific, unobserved effect into the conditional mean function for establishment births (Cameron and Trivedi (1998 pp. 96–102)). Thus, the negative binomial approach allows zip code areas to differ in a manner not fully accounted for by the observed covariates, has a natural interpretation as a random effects estimator, and offers a statistical control for unobserved heterogeneity among zip code areas.

Regarding endogeneity, an unobserved characteristic of a zip code area may affect the future pattern of economic activity, which feeds back through establishment behavior to affect the level of agglomeration. This problem is particularly troublesome when Y_{jk} is

measured as employment growth. One option for dealing with endogeneity in this situation is to instrument for the agglomeration variables. This approach was not pursued because it is unclear how these instruments are appropriately constructed and alternative methods of constructing instruments may result in widely differing estimates. An alternative and much simpler approach (see Dekle and Eaton 1999) rests on assuming that the municipality and industry fixed effects described above capture all unmeasured characteristics of a location that might make it a more or less productive place to do business. On the other hand, endogeneity appears to be a less serious problem when Y_{jk} is measured as establishment births because profit-maximizing entrepreneurs take an area's characteristics as exogenous when making location decisions. As discussed more fully by Rosenthal and Strange (2003b), this aspect represents an advantage of analyzing agglomeration economies using establishment births.

3. DATA

Data for this study come from annual censuses of establishments in all lines of business in South-Holland over the period 1988–1997. Establishments are enumerated based on information furnished by the Chamber of Commerce, insurance companies, and industrial sector associations, and a 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, they permit identification of industries and locations of establishments at a 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 '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, they may be bankrupt enterprises that had not yet been eliminated from the register. Enterprises not actually doing business are excluded from the data set.

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 (1999), Beardsell and Henderson (1999), and Henderson (2003). These studies are important because the contribution of agglomeration economies to productivity is estimated from plant—level production functions in a fixed effects framework; nevertheless, 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. In any case, estimation of establishment—level production functions is not a realistic option using the South—Holland zip code data. Employment is the only establishment—specific variable available; hence, the data are not well suited for panel analyses. Our approach of aggregating data for each industry into geographic units may average out some of the establishment—specific effects. A portion of these effects will remain uncontrolled, particularly 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.

Additionally, the level of detail available in the data is actually 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. Thus, they were aggregated up to the 4-digit zip code, 2-digit activity code level. The empirical analysis described in the next section proceeded at this level of industry and spatial detail. Further details regarding the data and construction of variables are discussed in Appendix A.

4. EMPIRICAL RESULTS

This section presents estimates of equation (1) with (Y_{jk}) measured both as employment growth by zip code–industry and the number of establishment births in a zip code–industry. The employment growth analysis is presented first, followed by an examination of the establishment birth data.

4.1 ANALYSIS OF EMPLOYMENT GROWTH

The employment growth analysis looks at the six industries in each of 416 zip codes with largest employment in 1988 (see Glaeser *et al.* (1992)). The dependent variable measures employment growth of a zip code–industry over the period 1988–97. Because zip code areas are small, some of the six largest industries 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. Additional estimates were obtained by taking different numbers of industries in a zip code and by using different minimum employment cut-offs, but without much change in the results described below.¹

Table 1 lists the ten industries that turned up most often in the 1997 observation

data set together with the number of employees in each. The most well represented sectors are building and construction, retail trade, business services (predominantly financial institutions and services), health care, education, and wholesale trade. Manufacturing industries are less frequently represented in the data than are non-manufacturing industries. Among the ten industries listed in Table 1, health care was the largest employer in South-Holland in 1988 with 103,166 employees and the metal products industry was the smallest employer with 8,026 employees. Overall, employment in South-Holland grew by 8.4% over the period 1988-1997. Employment in consumer services, producer services, and manufacturing sectors grew by 15.9%, 14.0%, and -9.8%, respectively, over this period.

Results from the employment growth analysis are presented in Table 2. Covariates are listed in the first column, their means are listed in the second column (with standard deviations in parentheses), and ordinary least squares regression results of two different specifications are reported in the third and fourth columns. Specification (i), shown in the third column, explains percentage growth in employment of a zip code-industry using just within-zip code agglomeration economy measures, and specification (ii), shown in the fourth column, uses these same measures and also includes distance-weighted agglomeration economy measures. Both regressions include 31 industry dummies and 69 municipality dummies to control for unobserved heterogeneity, although the outcome for the agglomeration economy variables is about the same whether or not the fixed effects are included. Thus, the small size and economic integration of South-Holland may help to ameliorate possible unobserved

¹ We have checked the robustness of our results to changes in this arbitrary cut-off level, exploring the impact of using cut-off levels of 75, 100, 125 and 250 employees. All conclusions reported using the 50

heterogeneity.

In both regressions, the adjusted R^2 equals 0.21. The comparatively low value for the adjusted R^2 may occur because of the small size of the South–Holland zip codes. 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 coefficient estimates. The Lagrange multiplier test for spatial error dependence (denoted LM(SE)) suggests that spatial autocorrelation is not a problem.

Regression results presented in Table 2 are broadly consistent with those of Glaeser *et al.* (1992) and go a step further to suggest that spatial effects of agglomeration economies are small. As shown in both specifications (i) and (ii), within-zip code industrial diversity and within-zip code competition foster employment growth, but within-zip code industrial concentration is not an important determinant of employment growth. The industrial specialization variable may have performed poorly because zip code-industries analyzed are dominated by different types of services (see Table 1) whose growth prospects are tied less to each other than to demands of the local population (but see the industry-specific establishment birth analyses below). Also, specification (ii) shows that coefficients of the three distance-weighted agglomeration economy measures are not significant determinants of employment growth at the 5% level. In fact, the only distance-weighted control variable that shows up significantly is the one with respect to industry employment outside the zip code; the more employment in the same industry outside the zip code, the lower the rate of growth of the zip-code industry under consideration.

employees cut-off level remain unaffected.

Together, these outcomes suggest that agglomeration economies affect local employment growth, but that effects of agglomeration economies in one location on employment growth in other locations die out quickly with distance. In one respect, this outcome is surprising. The small size of South-Holland zip codes leads to an expectation of spatial interconnectedness and Dutch zip code data frequently are grouped or spatially smoothed prior to analysis.² To measure the extent to which growth of zip code industries depends on the composition of economic activity in neighboring zip codes, we just ‘spatially smooth’ the agglomeration externality indicators on the right-hand side of the regression equation, but not the dependent variable. And, our results suggest that, given the small size of South-Holland zip code areas, agglomeration economies appear to operate on an even smaller geographic scale than implied by the estimates of Rosenthal and Strange (2003b) who looked at zip code areas in the U.S. In fact, estimates presented here suggest that agglomeration economies may well operate on a spatial scale that is much smaller than a city. Other results from control variables entered in the Table 2 regressions are that employment growth is higher in work areas, higher in areas with higher initial wages, and higher in zip codes that are further from the Rotterdam harbor. Of these three, the one with respect to initial wage rates is surprising, and suggests that the municipality fixed effects have not been able to fully solve the potential endogeneity problem we are confronted with. Finally, the industry’s 1988 employment in a zip code, total zip code employment, and proximity to Amsterdam and Utrecht have no discernable effect on employment growth.

The results presented in Table 2 regarding spatial effects of agglomeration economies are robust to three types of changes in specification. First, to check the extent

² We are indebted to a referee for pointing this out.

to which employment growth in different industries responds differently to agglomeration economies, an alternative data set was constructed to measure the growth rate of employment in each zip code in consumer services, producer services, and manufacturing.³ As previously noted, aggregating the detailed industry data into broad categories is necessary because zip code-industry employment totals frequently are small. Separate regressions were run for each of these three industries using the 416 zip codes as units of observation (results available from the authors on request). Explanatory variables used were identical to those shown in specification (ii) of Table 2.

In the producer services and manufacturing regressions, coefficients of the distance-weighted agglomeration economy measures were not significant at the 5% level. On the other hand, in the regression for consumer services, the coefficient of distance-weighted lack of industrial diversity was positive and significant at the 1% level and the coefficient of distance-weighted concentration was negative and significant at the 5% level; distance-weighted competition was found not to significantly affect growth rates. These results suggest that firms in the consumer services industry tend to grow slower if they are locally concentrated, but grow faster if the region itself is specialized in just a few other industries. Also, the coefficient of own-zip code competition was positive and significant at the 5% level in all three regressions, and the coefficient of own-zip code industrial concentration was negative and significant (at the 5% level) in the consumer services regression, but not in the other two.

³ Consumer services consist of retail, wholesale trade, health care, governmental services, environmental services, unions, cultural, sports and recreational services, and personal services. Manufacturing consists of food and beverages, tobacco, textiles, clothing, leather goods, paper, timber, oil processing, chemicals, synthetic and rubber, glass and ceramics, primary metal, metal products, machinery, computer manufacturing, audio and telecom equipment, cars, furniture and recycling. Producer services contain publishing and reproduction, telecom and postal services, banks, financial services, real estate intermediaries, computer services, research and development and other business services.

As a second robustness check, the same general pattern of results as shown in Table 2 are obtained when including the same control variables, but using alternative measures of own-zip code and distance weighted agglomeration economies. For industrial concentration, alternative measures included the fraction of a zip code's employment in an industry (which is just the numerator of equation (2)) and employment in a zip code industry per square kilometer. Alternative measures of competition included the number of establishments per employee in an industry (the numerator of equation (3)) and the number of establishments in a zip code-industry divided by total zip code employment. Lack of industrial diversity was alternatively measured with a Gini coefficient to reflect inequality in the distribution of employment across industries in a zip code. Additional regressions using these alternative agglomeration economy measures instead of those used in Table 2 continue to show that within-zip code competition and industrial diversity, but not within-zip code industrial concentration, foster employment growth. When distance-weighted measures of these variables are entered, however, their coefficients are not significant at the 5% level. These same conclusions about spatial effects of agglomeration economies are drawn when entering, in addition to own-zip code agglomeration variables, measures of agglomeration at the municipality level or with 10-kilometer bands.

As a third robustness check, the two specifications shown in Table 2 were re-estimated using a spatial lag model (Anselin 1988). The spatial weight matrix was developed using the reciprocal of distance between zip code centroids. In both regressions, coefficients of the spatially lagged dependent variable were positive, but not significant at the 5% level. Also, coefficient estimates again suggest that both own-zip code industrial diversity and competition positively affect employment growth, but that

coefficients of distance-weighted agglomeration economy measures are not significantly different from zero at the 5% level.

4.2 ANALYSIS OF ESTABLISHMENT LOCATION CHOICE

As discussed in Section 2, results of the analysis of employment growth may be affected by possible endogeneity; unobserved characteristics of a zip code area may affect the future pattern of economic activity, which feeds back through establishment behavior to affect the level of agglomeration. In the Table 2 regressions, this issue was dealt with by including municipality fixed effects, but this may not have solved all endogeneity problems, as suggested by the positive sign of the coefficient on the wage rate. An alternative approach would be to examine the pattern of new establishment births within the province. For new establishments, which can in principle be anywhere, a location's characteristics can be treated as exogenous.

Table 3 summarizes available information regarding establishment births in South-Holland for consumer services, producer services, and manufacturing. In total, over the period 1988–97, 57,249 establishment births occurred in all sectors, or 13.8 births per zip code per year. Only two out of the 416 zip codes had no establishment births at all. More births occurred in the consumer services sector than in other sectors. All but six zip codes had at least one establishment birth in this sector. Province-wide, there were about two-thirds as many establishment births in producer services as in consumer services. Compared to producer and consumer services, relatively few births occurred in manufacturing and 33 zip codes had no establishment births in this sector.

Table 4 presents results from estimating negative binomial count models to explain the number of new establishments "born" in consumer services, producer services, and

manufacturing. Poisson estimates were initially obtained, but tests (see Greene 2003, p. 743) indicated overdispersion. The negative binomial regressions treat zip codes as the unit of observation; thus, there are 416 observations in each. Including fixed effects for the 69 South-Holland municipalities turned out to be infeasible, but as discussed in Section 2, the negative binomial model controls for heterogeneity among locations by introducing a multiplicative random, zip code specific, unobserved effect into the conditional mean function for establishment births. The goodness of fit of the regressions is indicated by the LR-index (which is a pseudo- R^2), which suggests that the explanatory variables are able to explain 5%-7% of the variation in the dependent variable. The relatively poor fit is due to the small size of the zip codes in combination with the fact that location decisions are also influenced by unobservable and idiosyncratic factors.

Coefficient estimates presented in Table 4 indicate both similarities and differences between determinants of establishment births and determinants of employment growth. Consider first the performance of the agglomeration variables and their distance-weighted counterparts that are the major focus of this study. In Table 4, coefficients of lack of industrial diversity are negative and significantly different from zero in all three equations. This outcome is consistent with the employment growth regressions as well as with recent results of Holl (2004) for Portugal. It suggests that industrial diversity fosters growth. On the other hand, in Table 4 coefficients of competition never are significantly different from zero at conventional levels of significance (as compared with the positive and significant result presented in Table 2). Also, industrial concentration is attractive to new establishments in producer services and manufacturing (as compared with having no effect on employment growth in the Table 2 regressions).

Similar to results shown in Table 2, coefficients of the distance-weighted agglomeration variables generally are not significant at 5%, suggesting that agglomeration effects decay rapidly with distance. The chief exception here is that distance-weighted industrial concentration significantly affects establishment births in manufacturing. This outcome is at least broadly consistent with Dekle and Eaton (1999) who found that agglomeration economies in Japan operate on a national scale in manufacturing, but on a more localized scale in financial industries.

Regarding the outcomes for control variables: (1) the coefficient of the 1988 industry wage rate is negative and significant at the 1% level in the producer services regression, but not significant in the regressions for the other sectors, (2) the industry's employment in a zip code is attractive to new establishments in consumer services, but not in the other two sectors, (3) total employment outside the industry but within the zip code is attractive to new establishments in producer services and manufacturing, (4) industrial areas receive fewer new establishments in consumer and producer services, on average, than other zip codes, and (5) all new establishments appear to prefer locations close to Utrecht, which reflects proximity to the Province's economic hinterland, and manufacturing has a tendency to locate close to Rotterdam harbor.

Similar conclusions to those presented in Table 4 can be drawn when using the number of employees in new establishments as the dependent variable, rather than the number of establishments themselves. Indeed, the results (available from the authors on request) explaining employment creation in producer services, consumer services, and manufacturing show that coefficients of the spatially lagged agglomeration economy variables generally are not significantly different from zero at the 5% level. The most

important difference in the results lies in the role of local competition. Whereas the coefficients of the own-zip code competition variables are negative and consistently insignificant in the Table 4 regressions, they are negative and significant at the 5% level or better in the regressions explaining employment creation. This outcome is similar to the results reported in Table 2.

5. DISCUSSION AND CONCLUSIONS

This paper presents empirical evidence on the role of agglomeration economies on economic growth using data from the Netherlands. These data support analysis of industry growth in very small postal zip code areas in one of 12 Dutch provinces, the heavily urbanized province of South-Holland. Because municipalities in South-Holland cities are made up of many zip codes, these data allow us to analyze the impact of agglomeration economies on a spatial scale that is much smaller than a city. Therefore, to explain growth in industries in a specific zip code, we are able to distinguish between the economic characteristics of the zip code under consideration as well as those in neighboring zip codes. Economic growth is measured both by employment growth in a zip code industry and by the number of new establishments locating in that zip code.

Two main conclusions emerge from the analysis. First, agglomeration economies do contribute to growth. Industrial diversity is found to consistently promote both employment growth and new establishment creation. The presence of many small establishments (interpreted as a measure of local competition) promotes employment growth in both old and new establishments. Local industrial concentration, on the other hand, is a weak determinant of employment growth in both existing and new establishments, although it does appear to positively influence the birth of new

establishments irrespective of their initial employment levels. These results are roughly consistent with those of Glaeser et al. (1992). Second, we find that agglomeration economies are highly localized: the own-zip code agglomeration indicators have a much stronger effect on growth than indicators of agglomeration in nearby zip codes. This conclusion holds in the face of several types of changes in how agglomeration indicators are defined, including whether spatial effects are measured using distance-weighted indicators, municipality indicators, or indicators based on the aggregating the information within bands of 10 kilometers around the zip code under consideration. In other words, agglomeration economies in a zip code appear to have little effect on employment growth elsewhere, despite the fact that these zip codes average only about 6 km² in size. An implication of this analysis of growth both within and between cities 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 heretofore received little attention.

REFERENCES

- Anselin, L. 1988. *Spatial econometrics: Methods and models*. Kluwer Academic Publishers, Dordrecht.
- Anselin, L., A. Varga, and Z. Acs. 1997. "Local geographic spillovers between university research and high technology innovations," *Journal of Urban Economics* 42, 422–48.
- Audretsch, D.B. and M.P. Feldman. 1996. "R&D spillovers and the geography of innovation and production," *American Economic Review* 86, 630–40.
- Beardsell, M. and J.V. Henderson. 1999. "Spatial evolution and the computer industry in the USA," *European Economic Review* 43, 431–56.
- Black, D. and J.V. Henderson. 1999. "Spatial evolution of population and industry in the United States," *American Economic Review* 89, 321–27.
- Cameron, A.C. and Trivedi, P., 1998. *Regression analysis of count data*. Cambridge University Press, Cambridge.
- Ciccone, A. and Hall, R.E., 1996. Productivity and the density of economic activity, *American Economic Review* 86, 54–70.
- Combes, P. 2000. "Economic structure and local growth: France, 1984–1993," *Journal of Urban Economics* 47, 329–355.
- Dekle, R. and Eaton, J. 1999. "Agglomeration and land rents. Evidence from the prefectures," *Journal of Urban Economics* 46, 200–214.
- de Lucio, J.J. , J.A. Herce, and A. Goicolea. 2002. "The effects of externalities on productivity growth in Spanish industry," *Regional Science and Urban Economics* 32, 241–58.
- Glaeser, E.L., H.D. Kallal, J.A. Scheinkman, and A. Shleifer. 1992. "Growth in cities,"

Journal of Political Economy 100, 1126–1152.

Greene, W.H. 2003. *Econometric analysis*. Upper Saddle River, New Jersey, Prentice-Hall.

Henderson, J.V. 2003. “Marshall’s scale economies,” *Journal of Urban Economics* 53, 1–28.

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

Holl, A. 2004. “Transport infrastructure, agglomeration economies, and firm birth: empirical evidence from Portugal,” *Journal of Regional Science* 44, 693–712.

Jaffe, A.B. 1989. “Real effects of academic research,” *American Economic Review* 79, 957–970.

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

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

Romer, P.M. 1986. “Increasing Returns and Long Run Growth,” *Journal of Political Economy* 94, 1002–1037.

Rosenthal, S.S. and W.C. Strange. 2003a. “Geography, industrial organization and agglomeration,” *Review of Economics and Statistics* 85, 377–393.

Rosenthal, S.S. and W.C. Strange. 2003b. “Evidence on the nature and sources of agglomeration economies,” in Henderson, J.V. and J-F Thisse, eds. *Handbook of Urban and Regional Economics*. Amsterdam: North-Holland.

RPD, 1998. *Werklocaties 1998*. The Hague: Rijksplanologische Dienst, Rapportage
bedrijventerreinen en voorradige kantoorlocaties (IBIS).

TABLE 1: RANKING OF INDUSTRIES BY FREQUENCY OF OCCURRENCE
(WITH EMPLOYMENT LEVELS) IN THE SOUTH-HOLLAND TOP 6 DATA SET

Industries	Number of times an industry is among the six largest industries in all zip code areas	Total sector employment in South-Holland in 1988
Building and construction	231	66,016
Retail trade	194	69,997
Business services	185	76,913
Health care	174	103,166
Education	165	43,893
Wholesale trade	140	66,063
Agriculture and fishery	89	31,393
Government and social insurance	82	51,945
Distribution by land	76	22,534
Metal products industry	35	8,026

TABLE 2: 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		
	Means	(i)	(ii)
Constant		-0.945 [*] (0.448)	0.114 (0.568)
Annual regional industry wage rate in 1988 (in thousands of 1988 Dutch guilders)	46.256 (8.434)	0.008 [*] (0.003)	0.009 ^{**} (0.003)
The industry's employment in the zip code in 1988	402.132 (582.569)	5.44E-05 (1.99E-04)	9.78E-05 (2.05E-04)
Distance-weighted employment within the industry <i>outside</i> the zip code in 1988	19053.036 (5086.486)		-3.18E-05 ^{**} (9.24E-06)
Employment in the zip code in 1988 <i>outside</i> the industry under consideration	2739.529 (2821.126)	1.64E-06 [*] (8.38E-07)	3.25E-06 ^{**} (1.06E-06)
Distance-weighted employment <i>outside</i> the zip code <i>outside</i> the industry under consideration in 1988	13182.728 (5408.328)		-9.45E-06 (5.74E-06)
Concentration in 1988	4.823 (13.582)	-0.009 (0.007)	-0.007 (0.007)
Distance-weighted concentration in 1988	0.518 (0.259)		0.1858 (0.306)
Competition in 1988	1.129 (0.924)	0.131 ^{**} (0.031)	0.144 ^{**} (0.034)
Distance-weighted competition in 1988	0.986 (0.147)		-0.437 (0.308)
Lack of industrial diversity in 1988	0.590 (0.137)	-0.689 ^{**} (0.202)	-0.658 ^{**} (0.202)
Distance-weighted lack of industrial diversity in 1988	0.333 (0.142)		-0.677 (0.519)
Dummy variable indicating whether the zip code's function in 1988 is predominantly industrial as opposed to residential	0.263 (0.440)	0.2845 [*] (0.065)	0.275 ^{**} (0.065)
Variable indicating the distance to Amsterdam	51393.010 (11850.717)	2.54E-06 (4.66E-06)	-2.68E-06 (5.14E-06)
Variable indicating the distance to Rotterdam harbor	21465.484 (10532.938)	1.06E-05 [*] (5.24E-06)	-8.26E-08 (6.69E-06)
Variable indicating the distance to Utrecht	52615.179 (10004.027)	3.06E-06 (3.05E-06)	4.82E-06 (3.31E-06)
Industry fixed effects		YES	YES
Municipality fixed effects		YES	YES
SUMMARY STATISTICS			
N		1797	1797
Adjusted R ²		0.213	0.219
LM (SE)		0.016	1.029

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

Table 3: Establishment Births in South-Holland

Sector	Total Births	Mean Births per Zip Code	Standard Deviation of Births	Zip Codes with Zero Births
Consumer services	27,919	59.90	64.70	6
Producer Services	18,578	44.68	44.65	11
Manufacturing	3720	8.94	8.41	33
TOTAL (All Sectors)	57,249	137.61	62.54	2

TABLE 4: ANALYSIS OF ESTABLISHMENT BIRTHS IN THREE INDUSTRIES IN SOUTH-HOLLAND

EXPLANATORY VARIABLE	Consumer services	Producer services	Manufacturing
Constant	5.044** (1.678)	6.419** (1.291)	5.671** (1.782)
Annual regional industry wage rate in 1988 (in thousands of 1988 Dutch guilders)	-0.010 (0.019)	-0.040** (0.015)	-0.004 (0.016)
The industry's employment in the zip code in 1988	3.63E-04** (7.79E-05)	-2.50E-04 (1.29E-04)	-9.03E-05 (1.67E-04)
Distance-weighted employment within the industry <i>outside</i> the zip code in 1988	6.63E-06 (1.11E-05)	-0.017 (0.013)	-2.10E-04 (2.11E-04)
Employment in the zip code in 1988 <i>outside</i> the industry under consideration	3.30E-05 (4.74E-05)	2.12E-04** (4.12E-05)	1.50E-04** (3.23E-05)
Distance-weighted employment <i>outside</i> the zip code <i>outside</i> the industry under consideration in 1988	-1.04E-05 (1.18E-05)	7.28E-06 (1.23E-05)	1.88E-05 (1.53E-05)
Concentration in 1988	-0.281 (0.156)	0.398* (0.177)	0.262** (0.082)
Distance-weighted concentration in 1988	0.042 (0.067)	0.014 (0.016)	0.186** (0.091)
Competition in 1988	-0.042 (0.076)	-0.007 (0.037)	-0.014 (0.010)
Distance-weighted competition in 1988	-0.023 (0.045)	-0.003 (0.013)	0.018 (0.017)
Lack of industrial diversity in 1988	-2.859** (0.483)	-2.488** (0.505)	-3.543** (0.408)
Distance-weighted lack of industrial diversity in 1988	0.165 (0.235)	0.042 (0.104)	-0.507* (0.305)
Dummy variable indicating whether the zip code's function in 1988 is predominantly industrial as opposed to residential	-0.546** (0.097)	-0.254* (0.010)	-0.112 (0.138)
Variable indicating the distance to Amsterdam	0.011 (0.010)	0.004 (0.010)	-0.005 (0.014)
Variable indicating the distance to Rotterdam harbor	0.001 (0.014)	1.21E-04 (0.012)	-0.037 (0.018)
Variable indicating the distance to Utrecht	-0.012 (0.007)	-0.018** (0.005)	-0.017* (0.007)
SUMMARY STATISTICS			
N	416	416	416
Log likelihood	-1996.654	-1901.561	-1262.926
LR index	0.059	0.050	0.065

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

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 NUTS3 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 the Arc GIS geographical information system. A detailed description of the data and the verifications

applied can be obtained from the authors.

FIGURE 1 RANDSTAD, INTERMEDIATE ZONE AND NATIONAL PERIPHERY
SPATIAL REGIMES IN THE NETHERLANDS

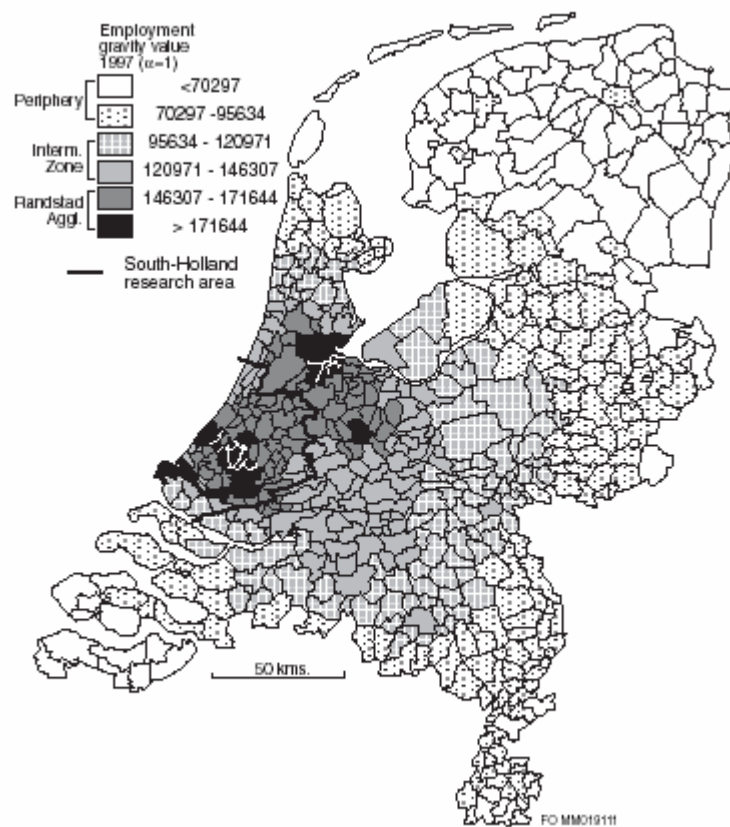


FIGURE 2 THE SOUTH-HOLLAND RESEARCH AREA (416 ZIP CODES)

