13 Dynamic Information Externalities and Employment Growth in the Netherlands[1]

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

13.1. Introduction

Beginning with Romer (1986) and Lucas (1988), the theory of endogenous growth emphasised the role in the growth process of both the stock of knowledge and the (planned or unplanned) transfer of knowledge between economic agents. For example, knowledge spills over between firms via informal contacts between employees, or because employees switch jobs and take their knowledge with them. Indeed, the most important type of knowledge that plays a role in the growth process is not necessarily path-breaking innovations, but may be learning opportunities for everyday people (Glaeser 1999). Empirical tests of this theory have often looked at cities to identify settings in which these external factors most effectively foster growth. Results, however, have been sharply divided. On the one hand, Glaeser *et al.* (1992) and Feldman and Audretsch (1999) find that employment growth is enhanced by diversity of activity across a broad range of sectors. Henderson *et al.* (1995), Black and Henderson (1999a). Beardsell and Henderson (1999), on the other hand, find faster growth when more activity is concentrated in a single sector. While endogenous growth theory is among the most powerful advances in economics in the past quarter-century, the fact that no clear view has emerged regarding situations to which it best applies represents a barrier to its further development and application. In growth models, for example, it is appropriate to treat urban areas as completely specialised as in Black and Henderson (1999b) or to assume that knowledge spills over predominantly between employees within the same industry as Glaeser (1999) does? Or, is industrial diversity such a fundamental component of the growth process that it must be captured in models such as those outlined in Fujita, Krugman, and Venables (1999)? Moreover, the lack of agreement on the relative importance of industrial concentration and diversity sends an ambiguous message regarding policy choices to promote or manage growth in urban areas.

This chapter includes three steps toward a better understanding of the relationship between knowledge spillovers and economic growth using data for the Netherlands. First, we provide insight into potential explanations for

R.A. Boschma and R.C. Kloosterman (eds.), Learning from Clusters: A Critical Assessment, 303–332.

differences in results of two highly influential papers, Glaeser *et al.* (1992) and Henderson *et al.* (1995), using data from outside the United States. Like Combes (2000), who analysed data from France, we draw attention to the importance of differences in the sectoral composition of the two data sets and also focus on the differences in methodologies used. Second, we address several long-standing issues using data at the municipality level from the entire country, as well as data from individual postal zip codes in one of the 12 Dutch provinces, *Zuid-Holland* (South-Holland). This province, which is about the size of the Dallas-Ft. Worth metropolitan area, covers a substantial part of the country's core economic area, the Randstad. Because zip codes average less than 6 km^2 in size, we can analyse employment growth in relatively small (urban) areas within an already small heavily urbanised region. As explained more fully below, this affords better controls for spatial aggregation error and unobserved location attributes than can be found in prior studies (Wallsten 2001). Third, we identify growth determinants among establishments that have remained at one location for a period of years and develop an alternative measure of local competition. Previous studies have not been able to distinguish between employment growth in existing establishments and employment growth attributable to establishment births, deaths, and relocations. We find that local competition may retard employment growth among existing establishments, a result that sets our work apart from earlier studies.

The remainder of the chapter is divided into five sections. Section 13.2 reviews the approaches taken in two prior studies in order to motivate the analysis presented later on. Section 13.3 describes the data. Section 13.4 examines determinants of employment growth in 234 Dutch municipalities. Section 13.5 looks at employment growth in 416 zip (postal) code areas in South- Holland. Section 13.6 concludes.

13.2. Background

Knowledge-based theories of endogenous growth can be tested at the city level. The density of economic activity in cities facilitates face-to-face contact as well as other forms of communication (Lucas 1993). Several hypotheses have been proposed concerning conditions under which knowledge spillovers affect growth. One hypothesis, originally developed by Marshall (1890) and later formalised by Arrow (1962) and Romer (1986) (MAR), contends that knowledge is predominantly sector-specific and hence that regional specialisation will foster growth. Furthermore, (local) market power is also thought to stimulate growth as it allows the innovating firm to internalise a substantial part of the rents. A second hypothesis, proposed by Porter (1990), also states that knowledge is predominantly sector-specific, but argues that its effect on growth is enhanced by local competition rather than market power as firms need to be innovative in order to survive. A third hypothesis, proposed by Jacobs (1969), agrees with Porter that competition fosters growth, but contends that regional diversity in economic activity will result in higher growth rates as many ideas developed by one sector can also be fruitfully applied in others. A fourth hypothesis, of course, could be developed by combining aspects of the other three to emphasise the role of industrial diversity in a non-competitive environment.

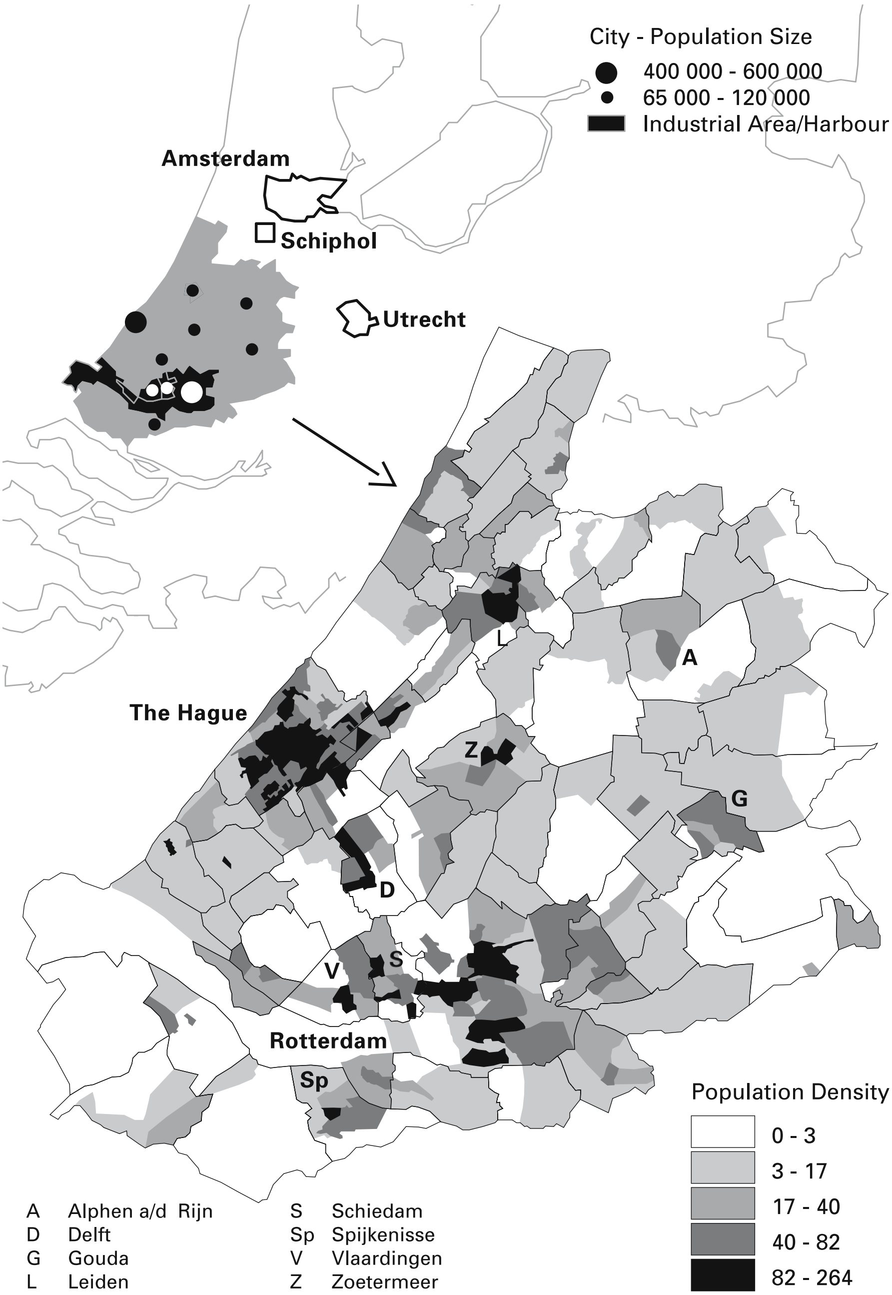
Two important papers that empirically test these hypotheses are by Glaeser *et al.* (1992) and Henderson *et al.* (1995). These papers both use employment data to measure growth[2] but, as indicated above, reach different conclusions, particularly regarding effects of local industrial concentration versus local industrial diversity. The former study finds evidence supporting the Jacobs hypothesis, whereas the latter finds evidence consistent with both the MAR and Jacobs view, depending on whether mature capital goods or high-tech industries are considered. One key difference between these studies rests on whether data from all cities in a given industry are analysed (Henderson *et al.* 1995) or whether only the largest industries in each city are included in the sample (Glaeser *et al.* 1992). Consequently, Glaeser (1998, p.148) suggests that "[a] possible reconciliation of results [on this point] is that scale and concentration may have value for smaller firms; however, diversity has more value for long term growth." Beardsell and Henderson (1999) argue that another important difference lies in the treatment of time invariant firm and/or location attributes. In particular, they state (p.449) that "...rather than the link between the present and the past representing mostly dynamic externalities, an alternative explanation is that there is a location fixed/random effect in estimation that gives rise to the role of history." Glaeser *et al.* (1992, p.1148) counter this view by distinguishing between the role of historical factors, such as natural resource and transport advantages, in location versus the role of these factors in growth. Kim (1999) and Ellison and Glaeser (1999) provide a more complete discussion of issues related to natural advantage and location.

Other explanations for the differences in results from these studies are also possible. For example, in the Henderson *et al.* (1995) study, the strategy of analysing all cities in a given industry turned out to be problematic. Because of disclosure rules, employment data for as many as 30% of cities was censored. This problem led to estimating a Tobit model in which the log of the end-of-period (1987) employment level was regressed on the log of beginning-of-period (1970) employment level. This approach is natural given the circumstances faced. However, controlling for the fixed/random effects of history becomes difficult with only one cross-section of data for each industry. A number of explanatory variables were tried that might be components of a fixed effect, but they performed unevenly.

Glaeser *et al.* (1992), on the other hand, estimated equations to explain growth in, rather than the level of, employment for city industries. As the existing growth models imply that the knowledge externalities are sources of *permanent* growth, they focus on the largest (and hence often mature) industries. Therefore, data was drawn from the six largest industries in each of the U.S. cities studied, and hence censoring did not appear to be as serious as in the Henderson *et al.* (1995) study. Also, they drew a substantial proportion of their observations from non-manufacturing industries (about one-third came from wholesale trade, construction and auto dealers and service stations), whereas Henderson *et al.* (1995) looked only at manufacturing industries. Recent evidence for France shows that indeed the composition of the data set may at least partially explain the difference in findings: Combes (2000) finds that diversity tends to enhance employment growth in services, whereas it tends to retard growth in manufacturing industries. However, specialisation does not seem to foster growth in either type of activity. Finally, in addition to variables measuring agglomeration economies, Glaeser *et al.* (1992) included a control variable in their regressions measuring the national employment growth rate of the industry outside the city.[3] This variable was included to account for national demand shifts and to capture general (industry-wide) technological progress (see Blanchard and Katz 1992).

13.3. Data

Data for this study was drawn from Dutch municipalities and from postal zip code areas in the Province of South-Holland. The Netherlands is a small country with land area of about 41,000 km^2 and population density of 457 persons per km^2. The province of South-Holland is about 1/12th of the country and is heavily urbanised with a population density of 1200 inhabitants per km^2. This province covers a substantial part of the core economic region of the Netherlands, the Randstad and includes the country's second and third largest cities (Rotterdam and The Hague) as well as numerous



Map 13.1 The South-Holland Research Area (416 ZIP Codes)

medium-sized cities such as Leiden, Delft, and Schiedam. Figure 13.1 shows the South-Holland research area and its population build-up. Both the regional (South-Holland) and national (Netherlands) data sets are of particular interest because (i) they include virtually all establishments present in the Netherlands for the period 1991-1997 and in South-Holland for the period 1988-1997, and (ii) the data is available on a fine spatial and industrial scale (see Appendix A). The time intervals are the longest over which employment growth can be measured using this detailed data in both the Netherlands and South-Holland.[4] Establishments were enumerated based on information furnished by the Chamber of Commerce, insurance companies and industrial sector associations, and an annual questionnaire was sent to each. The average response rate to the questionnaire was 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. The Dutch and South-Holland data sets, however, are not identical (Van Oort 2004). Whereas, for the entire Netherlands, only employment *totals* are available by industry and zip code, the South-Holland data set contains information on individual establishments. Thus, in one respect, the South-Holland data resembles the Longitudinal Research Data made available by the U.S. Census Bureau, but contains information on all establishments located there, not just those engaged in manufacturing. A disadvantage of the Dutch and South-Holland data sets, however, is that they do not contain measures of outputs, inputs other than labour, or plant characteristics. Consequently, they are not appropriate for estimating establishment-level production functions, as in Beardsell and Henderson (1999).

Spatial and industrial detail is an obvious advantage. However, the level of detail in both the Dutch and South-Holland data is actually too great for the purposes of this study. When the data is organised into a location-by-activity matrix, most of the cells contain no information. Many of the 6-digit zip code areas, for example, have only residences and individual 5-digit industries are present in only comparatively few 6-digit zip codes. Consequently, the data was first aggregated up to the 4-digit zip code, 2-digit activity code level (roughly the equivalent of 2 digit industries in the U.S. SIC system). In South-Holland, for example, the average size of a 4-digit zip code is about 5.65 km^2, although they tend to be smaller in urban centres where the density of addresses is high and larger in areas that have more open space. In any case, a zip code is quite small, particularly in comparison to U.S. counties or cities.

The South-Holland data was left at the 4 digit zip code level, whereas the data for the Netherlands was further aggregated into 548 municipalities

(69 of which are in South-Holland) in order to conduct analyses similar to those in earlier studies. These municipalities ranged in size from international cities such as Amsterdam to small villages. Because of their relative size to other municipalities, Amsterdam, Rotterdam, Utrecht, and The Hague were further subdivided into 3-digit zip code areas, roughly corresponding to economic areas in the core and periphery, containing 50,000 to 100,000 persons each. This led to a total of 580 geographic units (still referred to as municipalities).

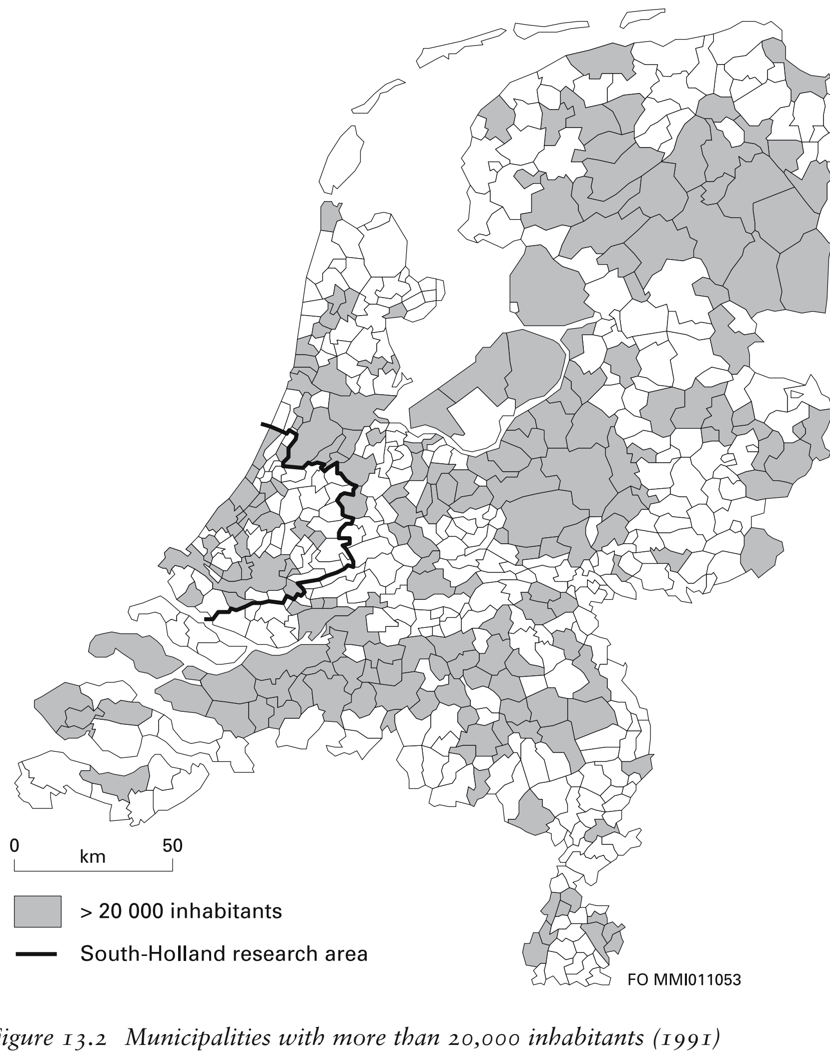


Figure 13.2 Municipalities with more than 20,000 inhabitants (1991)

Netherlands			South-Holland		
Sectors	*Represen-tation (1991)*	*Employ-ment (1991)*	*Sectors*	*Represen-tation (1988)*	*Employ-ment (1988)*
Health care	206	545,343	*Building and construction*	289	67,500
Building and construction	195	244,735	*Remaining Business services*[1]	263	78,610
Retail trade	179	344,665	*Retail trade*	232	70,929
Remaining business services[1]	156	370,842	*Education*	221	45,116
Wholesale trade	109	239,908	*Health care*	202	103,656
Education	100	154,444	*Wholesale trade*	181	67,114
Agriculture and fishery	90	88,620	*Agriculture and fishery*	132	32,406
Government and social insurance	73	143,011			
Distribution by land	105	23,156	*Distribution by land*	33	30,848
Government and social insurance	99	52,475			
Food and beverage processing industry	27	33,419	*Consumer services (non-retail)*	71	1,918

[1]*Remaining business services: juridical, taxes, public relations, consultancy.*

Table 13.1 Largest Sectors Represented in the Netherlands and South-Holland Data

Because previous research dealt with cities, we focus our analysis on employment growth in urbanised areas. For the Netherlands we include in the analysis the 234 municipalities that have 20,000 or more of inhabitants (see Map 13.2). The province of South-Holland is heavily urbanised and hence all zip code areas are included in the data set. However, not all areas in South-Holland have equal population densities. Hence we include an indicator of the degree of urbanisation. This indicator is based on a criterion that distinguishes between the most heavily urbanised and other areas, based on (i) the density of addresses, population and employment (which includes addresses in the own zip code and those in neighbouring zip codes corrected for geographical distance), and (ii) the presence of urban services (such as, for example, the number of hospitals, sports facilities, social-cultural services, public transport, etc. per household). For a more detailed description, see WMD (1999). Applying these criteria, 62% of the zip code areas in South-Holland are classified as urbanised.

Table 13.1 shows the ten sectors that turned up most often among the six largest sectors in either the 234 Dutch municipalities 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, financial institutions and services, health care, education, and wholesale trade. It should be noted that manufacturing industries appear less often in these samples than do non-manufacturing industries. Table 13.2 shows the number of times individual manufacturing industries are among the largest six sectors present in Dutch municipalities. The table also shows total employment in the Netherlands data set for each of these ten industries. In the Dutch municipality data, food and beverages is the most frequently occurring manufacturing sector, but the chemical industry is the largest in terms of employment. Other manufacturing sectors, such as electronics, glass and ceramics, transportation equipment and medical instruments, are also re represented in the province, but in most cases there are too few establishments to permit a meaningful sector-specific analyses.

The Netherlands and South-Holland data sets were used to construct indicators of various types of agglomeration economies that are similar to those used in prior studies (see especially Glaeser *et al.* 1992 and Henderson

Sector	*Representation (1991)*	*Employment (1991)*
Food and beverage processing industry	27	33,419
Furniture industry	25	26,593
Metal products industry	24	12,511
Publishing and reproduction	17	21,177
Chemical industry	17	41,912
Machinery industry	16	16,911
Electrical machinery and instruments	7	16,614
Metal industry (primary)	7	15,089
Glass and ceramic industry	7	8,864

Table 13.2 Largest Manufacturing Industries Represented in the Dutch Municipality Data

	Definition	*South-Holland Average*	*Netherlands Average*
EMPLOYMENT GROWTH	*Change in the natural log of employment*	-0.263	-0.170
CONCENTRATION	*Share of the sector's employment in total employment in the zip code or municipality, divided by the sector's employment share in total employment in South-Holland or the Netherlands*	4.823	4.959
COMPETITION	*Number of establishments per worker in a zip code or municipality divided by the South-Holland or Netherlands ratio of establishments to workers*	1.129	0.788
TURNOVER	*Zip code-specific or municipality-specific sum of establishment births, relocations and deaths over the estimation period divided by the initial stock of establishments*	1.105	-
SHARE	*Employment share of the other 5 largest sectors in total regional employment (i.e., excluding the employment of the sector under consideration)*	0.590	0.484
GINI	*Gini coefficient for the distribution of employment by sector in the zip code or municipality under observation*	0.477	0.292
HHI	*Hirschman-Herfindahl coefficient for the distribution of employment by sector in the zip code or municipality under observation*	-	0.076
URBAN AREA	*Dummy indicating whether the zip code is heavily urbanised*	0.620	-
GROWTH	*Change in the natural log of total (South-Holland or Dutch) employment excluding the zip-code/municipality under consideration*	0.082	0.004
INITIAL WAGE	*Natural log of sectoral wage rates, which are national averages in the Netherlands and regional averages in the South-Holland data sets*	3.881	3.818
WAGE	*Change in natural log of (regional) sectoral wage rates*	0.278	0.301
INITIAL EMPLOYMENT	*Natural log of initial zip code (1988) or municipality (1991) employment*	5.448	6.820
EMPLOYMENT 1997	*Natural log of end-of-period zip code or municipality employment*	5.177	6.584

WORKAREA	*Dummy variable equals 1 if zip code or municipality has more than 500 employees per 100 households*	0.263	8.850
INDUSTRIAL ZONES	*Dummy variable indicating more than South-Holland or Netherlands average of opening up of acres new industrial site relative to total stock of acres industrial site present in base year*	0.151	0.459
DISTANCE ROTTERDAM	*Distance from the zip code's centre to Rotterdam Harbour*	21465.48	-
LACK OF ACCESSIBILITY	*Distance from zip code's or municipality's centre to nearest highway exit or railway station*	6.597	0.639
POPULATION GROWTH	*Change in natural log of the zip code's or municipality's population size*	0.094	0.066
RANDSTAD	*Dummy indicating location within the country's core economic region, the Randstad*	-	0.317
INTERM. ZONE	*Dummy indicating location in area between core and periphery in the Netherlands*	-	0.379

Table 13.3 Definition of the variables used and their average values for South-Holland and the Netherlands

et al. 1995). These indicators and other variables are constructed using data from the base year (1988 and 1991 for respectively South-Holland and the Netherlands) to reduce simultaneity problems. Also, this approach, unlike the one adopted by Feser (2001), facilitates testing as to whether effects of different types of agglomeration economies on growth persist over time. The variable definitions and sample means of the Netherlands and South-Holland data sets are summarised in table 13.3. It should be noted that to economise on notation, the same names are used for certain variables that appear both in the Dutch municipality and South-Holland zip code analyses. For each of the two data sets these variables are defined somewhat differently as emphasised in the paragraphs below.

CONCENTRATION is defined as a location quotient showing the percentage of employment accounted for by an industry in a municipality (or zip code) relative to the percentage of employment accounted for by that industry in the Netherlands (or South-Holland). This variable measures whether an industry is over- or underrepresented in a location compared with its average representation in a larger area. *COMPETITION*, measured as establishments per worker in a municipality (or zip code) industry divided by establishments per worker in that industry in the Netherlands (or South-Holland), indicates whether establishments tend to be larger or smaller in a municipality (zip code) compared to the country (province) as a whole. An alternative measure of local competition was developed for use in the South-Holland analysis and is discussed more fully in Section 13.5.[5] Two variables are used as a measure of industrial diversity to indicate how evenly employment in a municipality is spread across economic sectors. *Gini*, the Gini coefficient for the distribution of employment by sector in a municipality (or zip code), measures the absence of diversity and is similar to the Hirschman-Herfindahl index used by Henderson *et al.* (1995). As Glaeser *et al.* (1992) focus on changes in employment among the six largest sectors in each city, the employment share of the other five largest sectors in total employment in a municipality or zip code can be used as an alternative measure of (the lack of) diversity. Whereas *Gini* varies only across municipalities or zip codes, this index (referred to as *Share*) varies across both locations and industries at a particular location. A positive coefficient of *Concentration* and a negative coefficient of *Competition* support the MAR hypothesis. A positive coefficient of *Concentration* and a positive coefficient of *Competition* support the Porter hypothesis. A negative coefficient of *Gini* or *Share* and a positive coefficient of *Competition* support the Jacobs hypothesis.

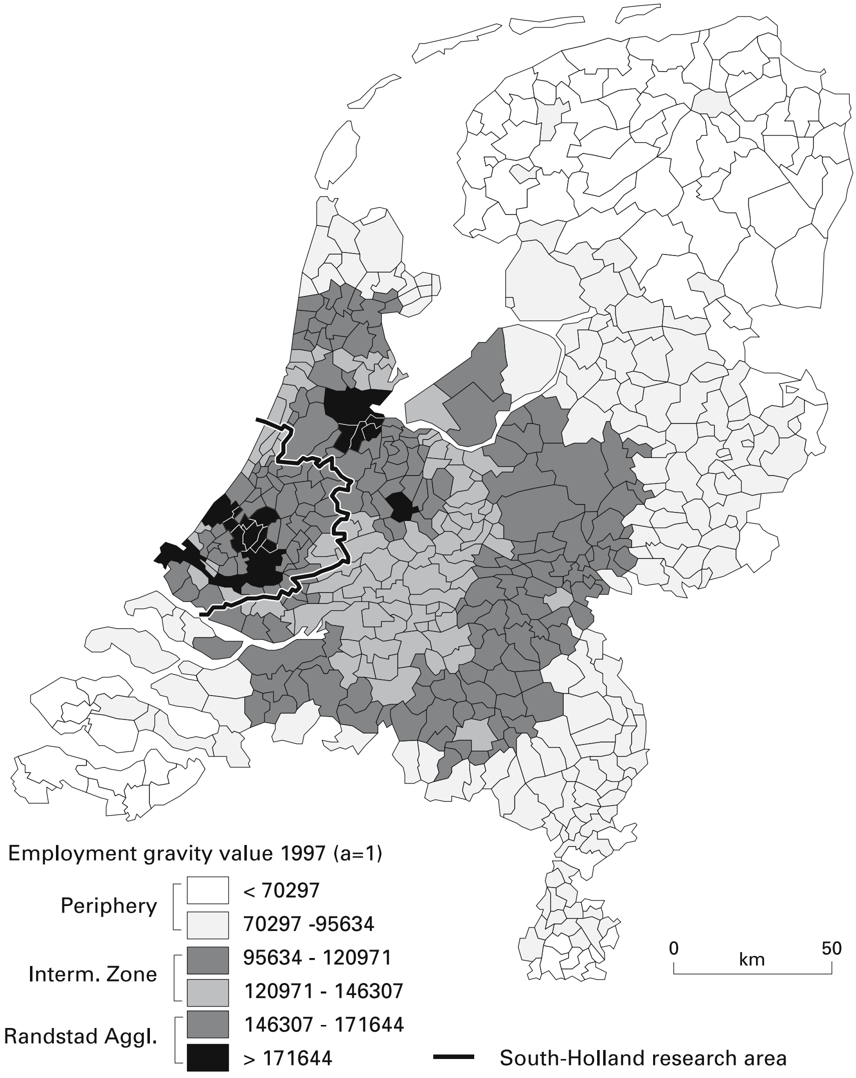
13.4 Analysis of the Dutch municipality data

As previously indicated, the Dutch municipality data is used in this section to compare results from the city-industry (here, municipality-industry) approach applied by Glaeser *et al.* (1992) and the individual-industry approach applied by Henderson *et al*. (1995). The aim here is not to attempt a reconciliation of their results. Rather, this starting point is adopted simply because it is useful to have some idea of how results from the Netherlands compare to those from the U.S. before looking at the South-Holland data.

The individual-industry approach was implemented by running the seven regressions presented in table 13.4 in which the dependent variable was the natural logarithm of 1997 employment (*Employment* 1997).[6] Industries selected represent both traditional manufacturing as well as industries that are more technologically oriented. Three of the explanatory variables (*Competition*, *Concentration*, and *Gini*) have already been discussed. Control variables measuring initial employment in a municipality-industry and region of the country were also included. *Randstad* indicates a location in the core economic region of the country and *Interm.zone* indicates a location in the intermediate zone between the country's core economic region

	Primary metal and metal products	Machinery industry	Electrical machinery and instruments	Transportation equipment	Computers	Audio, video and telecommunications equipment	Medical instruments
CONSTANT	0.596 (1.622)	0.838 (1.804)	-0.196 (-0.283)	1.646 (4.004)	-1.739 (-1.243)	-1.319 (-1.191)	1.774 (3.939)
INITIAL EMPLOYMENT	0.937 (20.456)	0.920 (15.124)	0.874 (9.953)	0.762 (15.514)	0.795 (3.525)	0.810 (5.891)	0.652 (9.574)
COMPETITION	-0.005 (-0.194)	0.009 (0.499)	-0.001 (-0.196)	0.002 (0.310)	0.001 (0.005)	0.013 (1.631)	0.002 (0.114)
CONCENTRATION	-0.059 (-3.168)	-0.092 (-1.663)	-0.158 (-1.840)	0.026 (1.002)	-0.255 (-1.569)	0.025 (0.222)	0.068 (1.144)
GINI	-0.672 (-1.120)	-0.642 (-0.695)	-0.179 (-0.093)	-3.100 (-2.744)	-5.165 (-1.252)	-4.684 (-1.478)	-2.40 (-2.164)
RANDSTAD	-0.189 (-2.003)	-0.542 (-3.532)	0.423 (1.289)	-0.031 (-0.168)	1.695 (2.580)	0.559 (1.083)	0.327 (1.649)
INTERM. ZONE	-0.033 (-0.361)	-0.077 (-0.518)	0.526 (1.680)	0.091 (0.500)	0.506 (0.808)	0.390 (0.814)	-0.166 (-0.886)
Summary Statistics							
Log Likelihood No. of Iterations	-190.430 1	-298.283 4	-412.078 5	-336.139 4	-264.570 8	-325.118 7	-351.006 3
No. of Observations	234	234	234	234	234	234	234
No. of Zero Employment	0	7	48	23	154	122	10

Table 13.4 Explanation of Employment 1997 Levels in 7 Selected Industries (t-values in parenthesis)



Map 13.3 Randstad, Intermediate Zone and National Periphery Spatial Regimes in the Netherlands

and its periphery. The three national zoning regimes used in our analysis are distinguished in Map 13.3 by means of a gravity model of total employment in 1997. In the analysis of urbanisation in the Netherlands, economic activity spreads from the Randstad region towards this so-called Intermediate zone, especially comprising the provinces of Gelderland and Noord-Brabant. This shift of economic activity is explained by increased congestion and increasing land scarcity in the Randstad (Lambooy 1998, Van Oort 2004). The specification shown in Table 13.4 is not as full as that used by Henderson *et al.* (1995) because we have no data on local labour market conditions for Dutch municipalities, such as wage payments or educational attainment. However, we have also included *COMPETITION* as an explanatory variable to achieve consistency with Glaeser *et al.* (1992). Tobit is used as an estimation method for all seven equations. However, this method is equivalent to least squares in the primary metals/metal products sector in which observations on the dependent variable are always positive (Greene 1997, p.965). Estimates converged in eight or fewer iterations.

Similar to findings by Henderson *et al.* (1995), coefficients of *INITIAL EMPLOYMENT*, included to capture persistence of industry employment levels, are positive and highly significant in all seven regressions. In most respects, however, similarities stop there. Findings of Henderson *et al.* (1995) strongly support the idea that the degree of past concentration of an industry positively affects later employment levels (the MAR view) in both traditional capital goods and newer high-tech manufacturing sectors. Furthermore, they report that historical industrial diversity in an area positively affects later employment levels only in high-technology manufacturing. In the results presented in Table 13.4, however, *CONCENTRATION* has either a negative effect or no significant effect on *EMPLOYMENT* 1997 in the seven sectors considered. *GINI* has a negative and significant coefficient with a t-statistic exceeding 2.0 in absolute value in just two sectors, one of which is a newer, technology-oriented manufacturing sector (medical instruments). Coefficients of *COMPETITION*, a variable not used by Henderson *et al.* (1995), have relatively small t-statistics.[7] Additionally, coefficients of dummy variables for location perform unevenly, showing that some industries appear to grow faster inside the Randstad, while others grow more slowly in that region. In any case, results presented in Table 13.4 provide no consistent support for MAR, Porter, or the Jacobs hypotheses.

Reasons why results in table 4 differ from corresponding estimates for the U.S. are not obvious. It is possible to speculate, however, that possible explanations rest on the short time interval (1991-97) for the Dutch municipality data, censoring of the U.S. data and the role of unmeasured establishment and/or municipality characteristics. The issue of unmeasured characteristics is discussed more fully in the next section.

Next, an analysis of municipality-industries (similar to Glaeser *et al.* (1992)) was performed using data from the 234 Dutch municipalities. Results

	EMPLOYMENT GROWTH (All sectors)	EMPLOYMENT GROWTH (Manufacturing sectors)
CONSTANT	-0.260	-3.433
	(-0.585)	(-1.661)
CONCENTRATION	-0.020	-0.016
	(-10.227)	(-4.680)
COMPETITION	0.159	0.363
	(5.490)	(3.801)
GINI	0.018[1]	-1.344[2]
	(0.093)	(-2.223)
GROWTH	0.765	0.244
	(5.160)	(0.555)
INITIAL WAGE	-0.059	0.574
	(-0.666)	(1.150)
INITIAL EMPLOYMENT	0.034	-0.010
	(1.886)	(-0.181)
WAGE	0.223	4.314
	(0.627)	(1.883)
RANDSTAD	-0.023	-0.029
	(-0.639)	(-0.275)
INTERM. ZONE	0.059	0.053
	(1.723)	(0.551)
N	1404	370
R^2	0.173	0.194

[1]The SHARE indicator yields similar results.
[2]The SHARE indicator turns out to be insignificant.

Table 13.5 Determinants of Employment Growth per Municipality (t-values are presented in parenthesis)

from two regressions are presented in Table 13.5. Column (1) shows the outcome from using data on the six industries with largest employment in each municipality and column (2) shows the outcome from using data on just the manufacturing sectors among the six largest sectors in each municipality. In each regression, the dependent variable is the change in the natural logarithm of municipality-industry employment over the period 1991-97. Explanatory variables included *CONCENTRATION* and *COMPETITION* (defined above) but used two alternative indicators for industrial diversity. Although Glaeser et al. (1992) used *SHARE* as an indicator of the absence of diversity, we use the Gini coefficient in both regressions as it was found to perform best. However, in the footnotes to the table we also indicate the results when *SHARE* was used.

Six control variables were also included in each of the Table 13.5 regressions. *INITIAL EMPLOYMENT* measures the number of employees in a municipality-industry at the beginning of the sample period. *GROWTH* is the change in the natural logarithm of employment in an industry outside the municipality. *WAGE* measures the difference in the natural logarithm of wages between industries at the national level (in the Netherlands) in 1991 and (*WAGE* measures the change in the natural logarithm of wages for each industry at the national level over the sample period.[8] *RANDSTAD* and *INTERM. ZONE* were defined previously in the context of the Table 13.3 regressions.

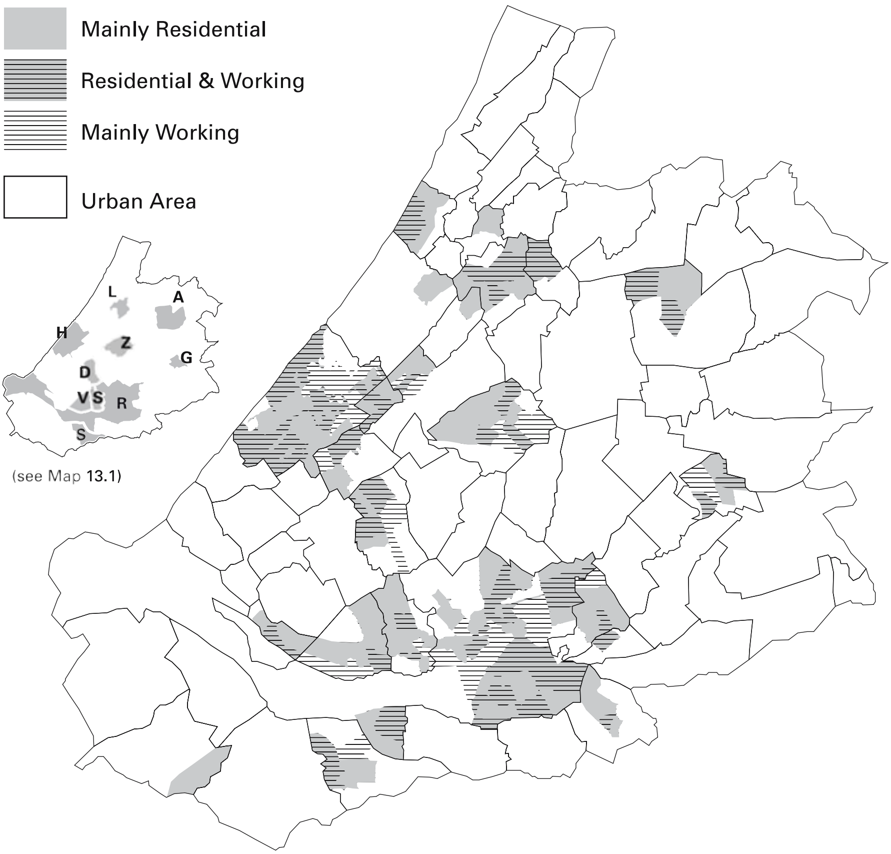
In table 5, both equations are estimated by least squares. Values of R^2 are 0.173 for the all sectors regression and 0.194 for the manufacturing regression. Thus, the explanatory power of both equations is rather low. The small size of many of the municipality industries may be partly responsible for this outcome. In situations where employment is comparatively low in the base year, relatively small absolute employment changes over the sample period can produce relatively large changes in growth rates. Correspondingly, with a small number of establishments operating in some municipalities, there is more room for growth rates to be affected by firm-specific factors (discussed momentarily) that are not controlled.

Results for the agglomeration indicators *CONCENTRATION*, *COMPETITION*, and *GINI* are at least broadly consistent with those obtained in the Glaeser et al. (1992) study. In both regressions, the coefficient of *CONCENTRATION* is negative and significantly different from zero at conventional levels; results that do not support the MAR and Porter hypotheses. The coefficient of *COMPETITION* also goes against MAR as more competition is found to increase growth in both manufacturing and non-manufacturing industries. The various measures of sectoral diversity do not appear to play a role in explaining employment growth for all sectors presented in the first column of Table 13.5. However, there is evidence that industrial diversity matters in determining growth in manufacturing sectors as the coefficient of the *GINI* index is negative and significant with a t-statistic exceeding 2.0 in absolute value (see the second column of Table 13.5). This outcome stands in contrast to the individual-industry analysis presented earlier and supports the notion that the Jacobs hypothesis has greater applicability to sectors in which employment has already reached some minimum threshold size. But if *SHARE* is substituted for *GINI*, its coefficient is not significantly different from zero at conventional levels, thus weakening the conclusion about the role of industrial diversity. Coefficients of control variables performed unevenly. For example, the coefficient of *GROWTH* is positive and highly significant in the column (1) regression, but is not significant at conventional levels in the column (2) regression.

13.5 Analysis of the South-Holland data

Results presented in the previous section are of interest because they highlight the role of agglomeration economies in the Netherlands in analyses similar to those conducted for the U.S. and France. Yet, they can be questioned from at least four perspectives. First, does the Dutch data offer adequate controls for unmeasured municipality and/or establishment specific effects? Second, does *COMPETITION* measure the degree to which establishments in a sector actually are confronted with competition, or does it just measure the relative size of establishments in a sector? Third, do the municipality-industry results apply to establishments present in the base year, or do they merely reflect a tendency for new establishments to start up or move into areas where their sector is underrepresented? Fourth, are the Dutch municipality-industry results misleading because of biases arising from spatial aggregation? These questions, which equally apply to prior empirical studies on the role of knowledge spillovers and agglomeration economies in urban growth, can be addressed more easily with the South-Holland data and are taken up in turn below.

South-Holland's small size and high degree of economic integration offers an important natural control for location-specific attributes. Between locations in South-Holland, there are few differences in resource endowments, political institutions, taxes, culture, environmental amenities (including climate), and environmental regulations. Additionally, the province is small enough for the labour market to be tightly integrated. Workers can live in one zip code and commute to work in any other using either public or private transport modes (and in fact they do!). Thus, wage rates within a sector show little variation between locations[9] and there is no need to control labour force characteristics such as level of education, percentage of workers with particular skills, or percentage of workers who are union members. Moreover, the role of



Map 13. 4 Residential-, Working- and Mixed Spatial Regimes in Urban Areas in South-Holland

history in determining the spatial economic layout of the province can be at least partially controlled using variables measuring distance of a zip code from intercity railway stations or major highway entries or exits (*LACK OF ACCESSIBILITY)*, the harbour in Rotterdam (*DISTANCE ROTTERDAM*), and whether the zip code is in an urbanised area (*URBAN AREA*). Controls for land use patterns can be obtained using variables showing whether a zip code is classified as predominately a work area (*WORKAREA*) and whether it has new industrial sites that can be developed (*INDUSTRIAL ZONES*). Maps 13.4 and 13.5 illustrate the typologies of living work and mixed areas within urban (Map 13.4) and non-urban (Map 13.5) zip codes in South-Holland (see for the selection criteria Section 13.3). Additional control variables include

Map 13. 5 Residential-, Working- and Mixed Spatial Regimes in Non-Urban Areas in South-Holland

region-specific and sector-specific wage rates (both *INITIAL WAGE* and (*WAGE*, defined in the previous section) and zip code *POPULATION GROWTH*.

Unmeasured establishment-specific effects are difficult to control using the South-Holland data set. Because establishments are aggregated into zip code-industries, some of these effects will average out, but other sources of these effects (e.g. clustering of high quality entrepreneurial talent, clustering of older and/or newer plants, and clustering of firms using particular specialised inputs) may remain. This problem can be treated using establishment-level data in a fixed effects framework. However, this approach involves sacrificing information by restricting the sample to establishments that appear in the South-Holland Firm Register in consecutive years. In fact, in their attempt to develop a panel of plants, Black and Henderson (1999a) ended up with sample sizes averaging only 8% of plants in an industry. Using the South-Holland data, an estimation of establishment-specific effects is not a realistic option in any case because employment is the only establishment-specific variable available. In consequence, the South-Holland zip code data is aggregated to zip code totals and analysed as a cross-section for 1988-97.

On the other hand, an advantage of the establishment-level South-Holland data is that it can be used to develop an alternative measure of competition that may be superior to those used in prior studies. More specifically, the relative establishment size variable (*COMPETITION*) used in the previous section and by Glaeser *et al.* (1992) may not be appropriate for two reasons. First, as is also noted by Combes (2000) and Rosenthal and Strange (2003), it is not clear whether this variable measures the extent of competition, internal diseconomies of scale, or broader aspects of industrial organisation. Second, this indicator may be inappropriate in cases where competition is faced from outside the local area and is particularly questionable when the "local area" is as small as a South-Holland zip code. Thus, for South-Holland the individual establishment data is used to develop an alternative measure of competition, *TURNOVER*, defined for each sector in each zip code as the sum of establishment births plus relocations plus deaths over the period 1988-97 divided by the number of establishments in the base year. *TURNOVER* may be a better measure than *COMPETITION* because it is based on establishment dynamics in a zip code (Dumais et al. 2002)..

Additionally, spatial aggregation in the Dutch municipality (and U.S. city) data is a potentially serious problem. Imagine an urban area that can be divided into a number of zones, each of which has the same number of employees and is completely specialised in the output of goods produced by a single (different) industry. Thus, each zone would have a high concentration index and no industrial diversity. From the standpoint of the urban area as a whole, however, concentration in production by a particular industry may or may not exceed its counterpart on a broader geographic scale and a Gini index will reflect maximum possible industrial diversity. Of course, an urban area is unlikely to develop as described in this stylised example. Yet, it is important to recognise that an entirely different view of the contribution of knowledge spillovers to growth could emerge from analysing parts of cities as compared with analysing cities as a whole. In any case, as mentioned previously, the South-Holland data permits the province to be divided into very small spatial units, so possible spatial aggregation error can be better controlled.

A limitation of the South-Holland data is, however, that it is not well suited to individual-industry analyses along the lines of those presented by Henderson *et al.* (1995). Most industries are present only in a small number of zip codes. As a consequence, both beginning-of-period and end-of-period employment would be zero for most observations. This aspect would not be a problem if the aim of the study was to ask why particular industries chose to locate in particular zip codes. However, the primary focus here is on the closely related issue of mechanisms thought to be important to the growth process. This emphasis motivates the decision to look only at employment growth in firms that were present at the beginning of the sample period.

The results from the South-Holland zip code-industry regressions, using the 1988-1997 change in natural logarithms of employment (*EMPLOYMENT GROWTH*) as a dependent variable, are shown in Table 13.6. Explanatory variables have again been constructed using data from the base year (1988 in this case) to minimise simultaneity problems. Similar to the municipality-industry analysis reported in the previous section, attention is restricted to the six largest sectors in each zip code. Because the province contains 416 4-digit zip code areas, a total of 2408 observations were possible. However, in some zip code areas, fewer than six sectors are present and in other zip code areas some of the largest six sectors have little base year employment making growth rate calculations problematic. In consequence, zip code industries with fewer than 50 employees in 1988 were excluded. This yielded a data set with 1797 observations. To gain an insight into the potential differences in the growth process in more and less heavily urbanised areas, we have interacted key variables of interest with *URBAN AREA* (see Map 13.4) to create *URBAN COMPETITION*, *URBAN CONCENTRATION*, and *URBAN SHARE*.

Column (1) presents results from a regression specified similarly to those used in the analysis of the Dutch municipality-industries. The value of R2=0.166 is once again rather low. However most of the estimated coefficients have significant (at 5% under a one-tail test) with plausible signs. Additionally, coefficient estimates obtained are broadly consistent with results presented by Glaeser et al. (1992) and support the Jacobs hypothesis. *CONCENTRATION* and *SHARE* enter with negative and significant coefficients. The effect of *CONCENTRATION* is stronger in more heavily urbanised areas as indicated by the outcome for the variable *URBAN CONCENTRATION*. Furthermore, *COMPETITION* is positively correlated with employment growth at least in urban areas. Thus, these results give additional support for Section 13.4's conclusion (based on municipality data for the entire country) that Jacobs externalities are the dominant type of knowledge spillovers. The fact that

	EMPLOYMENT GROWTH (All Establ.)	EMPLOYMENT GROWTH (All Establ.)	EMPLOYMENT GROWTH (Old Establ.)	EMPLOYMENT GROWTH (Old Establ.)
CONSTANT	2.121	2.933	2.252	2.310
	(3.182)	(4.441)	(2.962)	(3.059)
URBAN AREA	-0.239	-0.129	-0.304	-0.237
	(-1.016)	(-0.538)	(-1.135)	(-0.861)
CONCENTRATION	-0.009	-0.010	-0.009	-0.009
	(-2.434)	(-2.596)	(-2.129)	(-2.141)
COMPETITION	0.080	-	0.086	0.069
	(1.727)		(1.618)	(1.299)
TURNOVER	-	0.058	-	-0.111
		(1.428)		(-2.421)
SHARE	-0.628	-0.728	-0.903	-0.915
	(-2.037)	(-2.339)	(-2.566)	(-2.618)
URBAN CONCENTRATION	-0.009	-0.011	-0.008	-0.009
	(-2.229)	(-2.614)	(-1.777)	(-1.988)
URBAN COMPETITION	0.148	-	0.123	0.143
	(2.698)		(1.964)	(2.289)
URBAN TURNOVER	-	0.011	-	-0.077
		(0.198)		(-1.258)
URBAN SHARE	-0.035	0.054	0.091	0.128
	(-0.096)	(0.147)	(0.217)	(0.306)
GROWTH	1.077	1.043	0.778	0.963
	(6.986)	(6.543)	(4.428)	(5.371)
INITIAL WAGE	-0.452	-0.521	-0.550	-0.499
	(-2.761)	(-3.135)	(-2.946)	(-2.675)
INITIAL EMPLOYMENT	-0.029	-0.106	-0.024	-0.029
	(-0.906)	(-3.551)	(-0.648)	(-0.795)
WAGE	-1.478	-1.590	-1.098	-1.442
	(-3.352)	(-3.546)	(-2.184)	(-2.858)
WORKAREA	0.128	0.085	0.104	0.117
	(2.105)	(1.387)	(1.503)	(1.702)
INDUSTRIAL ZONES	0.172	0.164	0.178	0.215
	(2.513)	(2.355)	(2.282)	(2.754)
DISTANCE ROTTERDAM	3.982E-06	4.015	5.182E-06	4.767E-06
	(1.712)	(1.706)	(1.953)	(1.806)
LACK OF ACCESSIBILITY	0.008	0.008	0.011	0.010
	(1.916)	(1.909)	(2.229)	(2.142)
POPULATION GROWTH	0.131	0.112	0.082	0.084
	(2.254)	(1.908)	(1.238)	(1.282)
N	1797	1797	1797	1797
R^2	0.166	0.148	0.121	0.134

Table 13.6 South-Holland Regression Results (t-values are presented in parenthesis)

the Jacobs hypothesis is supported in this study of very small areas within an urbanised region strengthens the interpretation of our results, suggesting that they are not driven merely by spatial aggregation. Regarding control variables, the coefficient of *GROWTH* suggests that a 10% increase in the growth rate of an industry in South-Holland is associated with an increase in the growth rate of that industry in a zip code by 10.8%. This outcome indicates a tendency for industries to grow at about the same rate in zip codes where they are among the largest employers. Moreover, results from column (1) indicate that industries with comparatively high wage levels and wage increases tend to grow more slowly than other industries. Furthermore, employment growth is faster (i) if over the estimation period industrial zones expanded by more than the South-Holland average (*INDUSTRIAL ZONES*), (ii) if the area is a work area rather than a residential area (*WORKAREA*), and (iii) the faster the zip code's population growth (*POPULATION GROWTH*. Coefficients of *URBAN AREA*, *DISTANCE ROTTERDAM* (which also measures proximity to Amsterdam and Utrecht), *LACK OF ACCESSIBILITY* and *INITIAL EMPLOYMENT* in a zip code-industry are not found to be significant at conventional levels.

The specification shown in column (2) of Table 13.6 is the same as for the regression in column (1) except that *TURNOVER* is substituted for *COMPETITION*. The positive coefficients of *TURNOVER* and *URBAN TURNOVER* do not differ significantly from zero at the 5% level under a one-tail test. Thus, the alternative measure of competition indicates that greater numbers of establishment births, deaths and relocations in a zip code-industry do not lead to higher growth rates. This outcome weakens support for the Jacobs hypothesis found in the column (1) regression. Other coefficient estimates in the column (2) regression are similar to those presented in column (1) as multicollinearity between the various explanatory variables is very low, the only exception being that initial employment now becomes significant.

Also, as previously described, an advantage of the South-Holland data is the ability to distinguish establishments present at the beginning of the sample period from others that either moved in or started up after that time. Consequently, a regression was estimated (see column (3)) to look at the growth of zip code-industry employment only by the original (old) establishments present in 1988. In 1997, these establishments accounted for 64% of all South-Holland establishments as well as 83% of total South-Holland employment. Results from this regression again support the Jacobs hypothesis. Coefficients of *CONCENTRATION* and *SHARE* are negative and significantly different from zero and effects are equally strong in more and in less heavily urbanised areas. *COMPETITION* is found to be positively correlated with employment growth only in more heavily urbanised areas. This outcome is important because it suggests that the results that focus exclusively on existing firms reflect more than just a tendency for new firms to move into an area where their line of business is underrepresented. With respect to the other explanatory variables, three differences are worth mentioning. First, when analysing growth in existing firms, proximity to Amsterdam and the region's hinterland (as measured by the distance to Rotterdam) is positively related to employment growth in zip codes. Second, the coefficient on local population growth is no longer significant. Thus, local population growth may be a factor for attracting new establishments, but not a factor in the growth of old ones. Third, the larger the distance to intercity railway stations and highway entries and exits, the faster the employment growth. Thus, a mildly surprising result is that congestion appears to hamper growth in existing firms in the province of South-Holland.

Because *COMPETITION* and *TURNOVER* may not measure the same phenomenon, we ran a regression using employment growth in old establishments as the dependent variable with both indicators included as explanatory variables. Results are presented in the fourth column of Table 13.6. The coefficient of *TURNOVER* is lower (actually, it is negative and significant) than that reported in column (2) of Table 13.6. This outcome would be expected because in the all establishments regression, establishment births and relocations contribute to both *TURNOVER* and employment growth, whereas in the old establishments regression, births and relocations contribute only to *TURNOVER*. In contrast, *COMPETITION* and *URBAN COMPETITION* have positive coefficients, although only the coefficient for *URBAN COMPETITION* is significantly different from zero. Hence, if the Combes (2000) and/or the Rosenthal and Strange (2003) interpretation of this variable is accepted (that is, the variable measures internal diseconomies of scale or broader aspects of industrial organisation), the conclusions concerning the dominant type of externalities are substantially altered. Whereas regional diversity and lack of specialisation still foster employment growth (as predicted by Jacobs), regional competition (as appropriately measured by *TURNOVER*) is found to hamper rather than to foster growth in existing firms. In other words, although Jacobs' ideas concerning the regional composition is found to be supported by the South-Holland data, the fact that lack of competition is found to foster growth gives partial support to the views of MAR and Porter. The negative and significant coefficient of *TURNOVER* also emerges when *COMPETITION* and *URBAN COMPETITION* are excluded from the model.

13.6. Conclusions

The theory of endogenous growth emphasises the importance of knowledge and knowledge spillovers in the growth process. Considering the alternative hypotheses concerning the circumstances under which knowledge externalities are most likely to foster growth, the question arises as to whether knowledge spills over primarily between firms in the same sector, or whether growth is determined predominantly by knowledge spillovers between industries? In other words, is knowledge sector-specific or can ideas conceived in one sector be fruitfully applied in other sectors as well?

This chapter addresses this question, providing empirical evidence from the Netherlands. The regression results using data on Dutch municipalities give at least some support for Jane Jacobs (1969)'s hypothesis that knowledge spills over between sectors and that competition fosters growth because of the necessity to innovate. In this respect, the results are similar to Glaeser *et al.* (1992)'s analysis of employment growth in U.S. cities, and are in conflict with Henderson *et al.* (1995)'s findings that industrial concentration is more important than industrial diversity.

However, the data set for the province of South-Holland, which covers a substantial part of the core economic region of the country, enables us to correct several flaws in the analysis of Dutch municipalities. The most important of these is that it permits a sources-of-growth analysis in that changes in regional sectoral employment can be broken down to identify the separate contributions of growth by existing establishments as well as growth contributed by establishment births, deaths and relocations. As the theory of knowledge spillovers and growth focuses on dynamic externalities rather than at location choice, the appropriate dependent variable in the analysis is employment changes in existing firms. The results are markedly different from the results mentioned above. The results for regional composition still support Jacobs' theory that knowledge is not necessarily sector-specific and that ideas conceived in one sector can fruitfully be applied in others. However, the fact that lack of regional competition is found to foster growth gives support to the ideas of Marshall, Arrow, Romer and Porter that knowledge creation is stimulated by the possibility of rent capture. Hence, this outcome does not give full support to any of the existing hypotheses concerning the circumstances that foster growth.

Notes

1 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 chapter was written. He would also like to thank the Netherlands Organization for Scientific Research (NWO) for financial support (visiting grant B46-386). This chapter has benefited from careful comments by Erwin Bulte, Arjen Gielen, Henri de Groot, Jan Lambooy, John List, Bart Los and Willem van Groenedaal.

2 Because of the lack of data on sectoral output and the capital stock at the city level an appropriate measure of total factor productivity cannot be constructed. Glaeser *et al.* (1992) built a small model in which output is produced with only one input, labour, under conditions of decreasing returns to scale. Then, technological progress enhances the marginal value product of labour and hence the demand for labour increases. In that model, assuming constant prices for inputs and outputs, employment growth is an appropriate indicator of output growth.

3 Note that the variable measuring sectoral national growth rates outside the city would be virtually the same for each observation in the Henderson *et al.* (1995) analysis.

4 Henderson (1997) finds that effects of agglomeration economies on employment growth peak after about 5 years and die out after 6-7 years. Thus, for both data sets, the time interval over which employment growth was measured appears to be long enough to allow measurable data to emerge. See also Combes (2000).

5 Combes (2000) does not agree that the *COMPETITION* variable as constructed by Glaeser *et al.* (1992) is a proper measure of the degree of competition an industry faces. However, given that this variable measures the impact of relative firm size on employment growth, he argues that it can be used as a test for the importance of internal economies of scale; he proposes measuring competition by the inverse of a local Herfindahl index of productive concentration.

6 A related analysis was also performed using the 580 municipalities data set (i.e., after including the smallest municipalities) with similar results to those presented in table 3. These and all other results that are described, but not explicitly reported in the text, are available from the authors on request.

7 These results differ from those obtained by Combes (2000) in his analysis of (regional) employment growth in France. For manufacturing industries, he finds that (i) diversity slows down employment growth, (ii) specialisation hardly matters and (iii) smaller firms grow faster (where size is measured in terms of the number of employees per firm, which coincides with Glaeser *et al.*'s (1992) *COMPETITION* measure).

8 Note that the two wage variables could not be used in the individual industry analysis as they have no variation within a sector.

9 Although no zip code-specific sectoral wage data is available, the Dutch Central Bureau of Statistics distinguishes five regions in this province (so-called COROP regions) for which it calculates average sectoral wages. Pearson correlations of sectoral wages between regions range from 0.76 to 0.86.

References

Arrow, K.J. (1962), 'The economic implications of learning by doing'. Review of Economic Studies 29, pp.155-73.

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

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

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

Blanchard, O.J. and Katz, L.F. (1992), 'Regional evolutions'. Brookings Papers on Economic Activity, pp.1-76.

CBS (various years), 'Annual Regional Economic Data'. Voorburg: Statistics Netherlands.

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

Dumais, G., Ellison, G. and Glaeser, E.L. (2002), 'Geographic concentration as a dynamic process'. The Review of Economics and Statistics 84, pp.193-204.

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

Feldman, M.P. and Audretsch, D.B. (1999), 'Innovation in cities: Science-based diversity, specialisation, and localized competition'. European Economic Review 43, pp.409-29.

Feser, E.J. (2001), 'A flexible test for agglomeration economies in two US manufacturing industries'. Regional Science and Urban Economics 31, pp.1-20.

Fujita, M., Krugman, P. and Venables, A.J. (1999), The spatial economy: Cities, regions and international trade. Cambridge: MIT Press.

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

Glaeser, E.L. (1998), 'Are cities dying?' Journal of Economic Perspectives 12, pp.139-60.

Glaeser, E.L. (1999), 'Learning in Cities'. Journal of Urban Economics 46, pp.254-277.

Greene, W.H. (1997), Econometric Analysis. Upper Saddle River, New Jersey: Prentice-Hall.

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

Henderson, J.V. (1997), 'Externalities and industrial development'. Journal of Urban Economics 42, pp.449-71.

Jacobs, J. (1969), The economy of cities. New York: Vintage.

Kim, S. (1999), 'Regions, resources, and economic geography: Sources of U.S. regional comparative advantage, 1880-1987'. Regional Science and Urban Economics 29, pp.1-32.

Lambooy, J.G. (1998), 'Polynucleation and economic development: the Randstad'. European Planning Studies 6, pp.457-66.

Lucas, R.E. (1988), 'On the mechanics of economic development'. Journal of Monetary Economics 22, pp.3-42.

Lucas, R.E. (1993), 'Making a miracle'. Econometrica 61, pp.251-72.

Marshall, A. (1890), Principles of economics. London: Macmillan.

Van Oort, F.G. (2004), Urban growth and innovation. Spatially bounded externalities in the Netherlands. Alderhot: Ashgate

Porter, M. (1990), The competitive advantage of nations. New York: Free Press.

Romer, P.M. (1986), 'Increasing returns and long run growth'. Journal of Political Economy 94, pp.1002-37.

Rosenthal, S. and Strange, W. (2003), 'Geography, industrial organisation and agglomeration', Review of Economics and Statistics 85, pp.377-393.

Wallsten, S.J. (2001), 'An empirical test of geographic knowledge spillovers using geographic information systems and firm-level data', Regional Science and Urban Economics 31, pp.571-99.

WMD (1999), Woonmilieudatabase 1998 Toelichting [Explanatory Note to the 1988 Living Environment Data Base]. Delft: ABF-Research.

Appendix A: Description of the data

The data used in this paper is derived from various sources. The most important sources are the longitudinal datasets of the Firm Register South-Holland (BRZ) 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). However, actual firm level data is only available for South-Holland; the Netherlands dataset only gives information on sectoral employment (i.e. aggregate employment of all firms in a specific sector) in each 6-digit zip code area. The variables *EMPLOYMENT GROWTH*, *CONCENTRATION*, *COMPETITION*, *SHARE*, *INITIAL EMPLOYMENT* and *TURNOVER* on location-industry level are calculated from these data (*TURNOVER* could only be calculated for the South-Holland analysis). The data concerning agricultural employment was derived from the Agricultural Statistics of the Dutch Central Bureau of Statistics (CBS) at 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 Netherlands wage data were obtained from CBS Labour Statistics whereas the regional South-Holland wage rates were obtained from the CBS's Annual Regional Economic Dataset (various years). The variable measuring the distances to Rotterdam was constructed using an ArcGIS geographical information system. A detailed description of the data and the verifications applied can be found in Van Oort (2004).