Ann Reg Sci (2004) 38:291–310
DOI: 10.1007/s00168-004-0196-7

The Annals of
Regional Science

Localization effects of firm startups and closures in the Netherlands

Corina Huisman[1], **Leo van Wissen**[1,2]

[1]Netherlands Interdisciplinary Demographic Institute, P.O. Box 11650, 2502 AR The Hague, The Netherlands (e-mail: {huisman; wissen}@nidi.nl)
[2]Faculty of Spatial Sciences, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands

Abstract. In this article localization trends as a result of startups and closures are investigated in the Netherlands, using a distance-based approach. A major advantage of this method is that it does not suffer from aggregation bias that is inherent in area-based methods. This method controls for the existing spatial clustering of the industry. Plant openings and closures can either reinforce or weaken the existing localization pattern. We studied these localization tendencies for industries at the one-digit level. The major finding is that plant closures have a strong deconcentration effect, at the local as well as the regional level. Startups have a concentration effect at smaller spatial scales, but beyond 18 km this component also contributes to deconcentration. This result is in line with the spatial process of sprawl for most economic activities. However, results are different for economic sectors, and manufacturing clearly deviates from this general pattern, because it shows a localization trend except at the very small spatial level. Based on these results we do not find much support, except in the manufacturing industry, that industry dynamics in terms of new firm formation and closures, leads to stronger spatial agglomeration tendencies in the Netherlands.

JEL classification: R120, R340

1. Introduction

Spatial clustering of industries is an important topic in economic geography, regional economics, and regional science. Under certain conditions, which were already described by Marshall (1890), and later refined and extended by Arrow (1962) and Romer (1986), firms tend to co-locate in specific areas because spatial proximity to each other allows them to profit from positive externalities, particularly information spillovers, the availability of a large and specialized labor market, and local intra-industry specialization. These positive externalities are called *agglomeration economies*. Agglomeration economies may occur between firms within the same industry or in different industries. The spillovers that occur between firms within the same industry are called *localization economies*. In addition, between-industry positive externalities may also be important, which are called *urbanization economies*. These effects lead to the spatial clustering of different types of industries, or local diversification. The presence or absence of localization or urbanization economies is an important explanation for regional differences in economic growth. The process is also recursive and may lead to sustained and increasing growth once an initial localized advantage is established.

While agglomeration economies may lead to spatial localization, spatial *competition* may lead to spatial dispersion. If firms are substitutes, there exists rivalry among firms, which increases with higher firm densities. If competition, instead of agglomeration economies prevails, then dispersion is the outcome.

Despite the prominent role of *agglomeration economies* in the literature, empirical evidence of these recursive processes remains somewhat ambiguous. This pertains to many aspects of empirical verification. First, there is the problem of measuring if there is spatial clustering or not, and if present, the degree of spatial clustering. Second, there are problems in identifying if this spatial clustering is the result of agglomerative forces, or of natural advantages, or simply chance. A number of authors deal with the empirical verification of agglomeration effects (e.g., Adams and Jaffe 1996; Ciccone 2002; Ellison and Glaeser 1999; Henderson 1986, 2003; Hoogstra and van Dijk 2004; Jaffe et al. 1993; Nakamura 1985; van Oort 2004). Although these studies verify the existence of agglomeration effects, the exact nature and magnitude of these effects remains somewhat unclear. For instance, Ellison and Glaeser (1999) conjecture that at least half of the observed geographical concentration at the state level in the U.S. is due to natural advantages, and the rest is due to agglomeration effects and chance. At the same time they hope that other researchers may provide better estimates in the future.

Third, and related to the first two problems is the issue of the appropriate spatial scale of the process. Van Oort (2004) looks at the question to what extent the results change with varying spatial scales. His main result is that agglomeration effects work predominantly at the local level. Van Oort and Atzema (this issue) not only look at spatial proximity but also at other spatial configurations, such as urban networks. However, the present paper is based on distance-related statistics. Duranton and Overman (2002) conclude that localization (which may be due to any source of geographical clustering, including agglomeration effects) takes mostly place at distances below 50 kilometres.

This paper is focused on the first problem: the measurement of spatial clustering among industries in the Netherlands. While most contributions in this field deal with the analysis of localization in the spatial pattern of existing industries (see e.g., Ellison and Glaeser 1997; Feser and Sweeney 2000; Duranton and Overman 2002), here we take a firm demographic approach (Van Wissen 2002), and analyze the localization effects of the demographic processes of firm startups and closures. Do firm startups and closures have a concentration or dispersion effect on the clustering of industries? Moreover, what is the combined effect of startups and closures on the localization of industries? This approach views spatial clustering as a dynamic concept, where the spatial pattern is subject to change as a result of new firms added to, and closing firm subtracted from the stock. This view is consistent with evolutionary approaches to the emergence of industry clusters, such as the model of Klepper (2002) of the Detroit automobile industry, or Arthur (1990), about the evolution of Silicon Valley. Dumais et al. (2002) focused on the spatial clustering aspects of new firms and closures, in the US industry. If the spatial configuration of new firms is more dispersed than that of incumbents, there is a trend toward deconcentration, at least on account of this demographic component. On the other hand, if entrants are more clustered than existing firms, there is a tendency towards concentration. Concentration or dispersion effects may also result from closures. Dumais et al. found that the locational choices of startups lead to deconcentration, whereas closures lead to higher concentration of the industry.

This article investigates these processes for different industries at the firm level in the Netherlands. In the second section the methodology, derived from the method of Diggle's K-statistic is presented (Diggle 1983; Feser and Sweeney 2000). Section 3 describes the data. Sections 4 and 5 present the results: First, in Sect. 4, we illustrate the method using all business plants in the Netherlands. Next, in Sect. 5 we present results at the one-digit level of economic sectors. Section 6 concludes.

2. Measuring localization using Diggle's K-statistic

Do starting firms have a tendency to co-locate close to existing establishments? Are closing firms relatively more clustered among incumbent firms in the industry than others? In sum, do the demographic events of startups and closures have a concentration or a deconcentration effect on the population of firms? Various authors have used spatial statistical methods to study the question of geographical concentration. Some are based on the number of counts of firms in a geographical area (e.g., Ellison and Glaeser 1997; Dumais et al. 2002). In this category we also find the well-known spatial autocorrelation functions, such as Moran's I and Geary's c statistic. Other methods are based on distances between firms (Duranton and Overman 2002; Feser and Sweeney 2000). The area-based approach suffers from problems of aggregation (Duranton and Overman 2002). In particular, there are three main problems to this approach, viz.(1) results are difficult to compare across spatial scales; (2) there may be spurious correlation across aggregated variables (the so called Modifiable Areal Unit Problem MAUP); and (3) the spatial configuration of areas is not appropriately taken into account. Distance based statistics do not suffer from these problems. It is therefore not surprising that results based on both types of methods may give different answers. For instance, Duranton and Overman (2002) find that using the method of Ellison and Glaeser gives the result that 94% of industries in the UK show a tendency of localization, whereas using a distance based approach the figure drops to 43%. Moreover, the rank correlation coefficient between the overall industry rankings that results from both methods is as low as 0.04 and not significant. This is strong evidence that area-based approaches for detecting spatial clustering are severely biased.

In order to test for localisation economies we use a distance-based approach, which rests on the use of so-called K-functions. Diggle (1983) describes the K-functions as a model of constructing processes for spatial point patterns in biology. A spatial point process is any stochastic mechanism that generates a countable set of events x_i in the plane. The K function is defined as (Diggle 1983, p. 47):

$$K(t) = \lambda^{-1}E \text{ [number of further events within distance } t \text{ of an arbitary event]} \tag{1}$$

where λ is the intensity of the process, i.e. the mean number of events per unit area. Diggle and Chetwynd (1991) use K-functions in an epidemiological setting. They compare the distribution of persons diagnosed with a certain disease (the cases) with the distribution of persons selected at random from entries in the birth register (the controls).

In our case we compare the distribution of firms experiencing a specific demographic event (the cases), to the distribution of existing firms (the controls). Comparisons are made for all firms (Sect. 4), but also sector-specific (Sect. 5).

The method of K-statistics basically involves a cumulative counting process (Fig. 1). You take a firm that started during the year, and count the number of starting firms within a circle with a specific radius (s). Repeat this counting for several distances (in our case for each kilometre, until the maximum possible distance is reached). Repeat the counting of starting firms within certain distances for all other starting firms, and finally calculate the average for all these starting firms.

This results in a $K(s)$-function for starting firms. This function is later referred to as $K_{11,}$ where the index 1 denotes starting firms. Thus, K_{11} is based on distances of starting firms relative to other starting firms. In general, the K_{ij} function measures the location of members of group i vis-à-vis members of group j. If this process is repeated but now by taking an incumbent firm, and count the number of existing firms surrounding this firm, results in a second $K(s)$-function (referred to as K_{22}). The difference between these two K-functions ($D_{1122}(s) = K_{11}(s) - K_{22}(s)$) allows us to state whether starting firms are more or less clustered to each other (for a given radius s) than existing firms are to each other. A third K-function, which is most appropriate for our purpose, may be calculated by taking a starting firm and counting the number of surrounding existing firms. This function is referred to as K_{12}. The difference between K_{12} and K_{22} ($D_{1222}(s) = K_{12}(s) - K_{22}(s)$) now gives information whether starting firms are more or less clustered to existing firms than existing firms are to each other. If this is the case, starting firms tend to locate

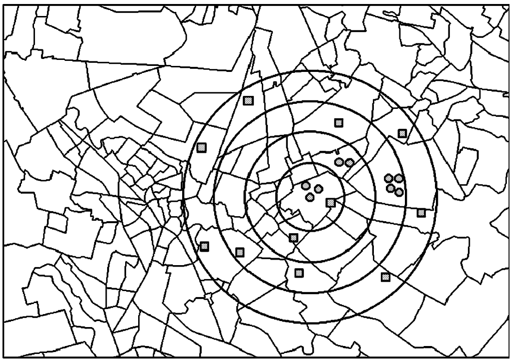

Fig. 1. Basics of K-statistics

relatively close to existing firms. Hence, these locational choices of starting firms lead to an increased spatial clustering of the industry. If starting firms are relatively more distanced from incumbents than incumbents are among themselves, the reverse is true and locational choices of new firms lead to de-localization of the industry. This process is explained in more detail in section four, using the example of all starting and existing firms in the Netherlands.

More formally, the standard *K*-function is equal to:

$$K_{ij}(s) = \lambda_j^{-1} E\left[N_{ij}(s)\right] \tag{2}$$

where λ_j is the intensity of type *j* events (controls), or mean number of type *j* events per unit area. *E*[.] is the expected value operator, and N_{ij}(s) is the number of further type *j* events within distance *s* of an arbitrary type *i* event (cases). If cases are on average neither more, nor less clustered than the set of controls the following equality will hold:

$$K_{11}(s) = K_{12}(s) = K_{22}(s) \tag{3}$$

The next function will hold positive values when the incidence of cases is more spatially clustered than the incidence of controls:

$$D_{1122}(s) = K_{11}(s) - K_{22}(s) \tag{4}$$

So, significantly positive values $D_{1122}(s)$ represent spatial clustering of type 1 events (cases) over and above the degree of spatial clustering of type 2 events (controls) attributable to heterogeneity.

If the incidence of the cases is more clustered to the incidence of controls, than the incidence of controls is clustered to itself, the following function will hold positive values:

$$D_{1222}(s) = K_{12}(s) - K_{22}(s) \tag{5}$$

In our example of starting firms, positive values indicate that new firms are located closer to existing firms, than existing firms are among themselves. Our empirical analysis is predominantly based on this statistic.

One advantage of $D_{1222}(s)$ is its simple interpretation: $\lambda_2 \mathrm{D}_{1222}(s)$ represents the expected number of excess starting firms within distance *s* of an incumbent, by comparison with the number expected in the absence of spatial clustering. We denote this quantity $\lambda_2 \mathrm{D}_{1222}(s)$ as the *Absolute Concentration Index* (ACI(*s*)). Likewise we can define the *Relative Concentration Index* RCI(*s*) as:

$$\mathrm{RCI}(s) = \frac{K_{12}(s)}{K_{22}(s)} * 100 \tag{6}$$

If starting firms choose locations relatively closer to incumbents than incumbents are located among themselves, RCI(*s*) > 0, which indicates a spatial clustering trend. For instance, a value of RCI(*s*) = 50 means that on average we find 50% more starters within a circle with radius *s* around an existing firm than expected on the basis of the density of existing firms. This is the most useful index for our purpose. Similarly, a value smaller than zero indicates the opposite process: dispersion as a result of the relatively larger distances between startups and existing firms.

Conceptually, a spatial cluster is often envisioned as a tight grouping of events within a given study area. This implies both a distinct range for *s* and a particular point of reference (e.g., a metropolitan centre). The *D*-function and corresponding ACI generalises beyond both of these conceptual bounds: the clustering hypothesis may be evaluated at any distance and the value of the statistic at each distance is based on the cumulative frequency of counts using every establishment in the region (or sample) as a point of reference (Feser and Sweeney 2000).

For estimating the *K*-functions we follow Diggle and Chetwynd (1991). For data $x_i \in A$: $i,\ldots,n$, where $n = n_1 + n_2$, with n_1 events of type 1 (cases) and the remainder n_2 of type 2 (controls), unbiased estimators for the $K_{ij}(s)$ can be obtained as follows:

$$K_{11}(s) = |A|(n_1(n_1-1))^{-1}\sum_{i=1}^{n_1}\sum_{\substack{j=1\\ j\neq i}}^{n_1} w_{ij}\delta_{ij}(s) \tag{7}$$

$$K_{22}(s) = |A|(n_2(n_2-1))^{-1}\sum_{i=n_1+1}^{n}\sum_{\substack{j=n_1+1\\ j\neq i}}^{n} w_{ij}\delta_{ij}(s) \tag{8}$$

$$K_{12}(s) = |A|(n(n_1-1)(n_2-1))^{-1}\sum_{i=1}^{n_1}\sum_{j=n_1+1}^{n} (n_2 w_{ij} + n_1 w_{ji})\delta_{ij}(s) \tag{9}$$

where $|A|$ is the area of the study region, n_1 is the total number of cases and n_2 the total number of controls. $\delta_{ij}(s)$ is an indicator variable taking the value 1 if $d_{ij} < s$ and 0 otherwise. w_{ij} is a correction factor for events within distance s of the edge of the study area. Diggle and Chetwynd defined the edge correction weights (w_{ij}) as the inverse of the proportion of the circumference of the circle with centre x and radius s, which lies within A. Let $d_{ij} = ||x_i\text{-}x_j||$ be the Euclidean distance between x_i and x_j. Then write $w_{ij} = w(x_i,d_{ij})$. w_{ij} is the inverse of the conditional probability that an event is observed, given only that it is a distance d_{ij} away from the ith event, x_i. See Fig. 2, and note that in general $w_{ij} \neq w_{ji}$ (Diggle 1983).[1]

In order to test for agglomeration effects in this paper we compare two *K*-functions: one for establishments experiencing a demographic event (cases) and one for already existing establishments (controls). The distances were calculated by using the x- and y-coordinates of the four-digit postcodes of the firms. Using these two *K*-functions we will calculate the relative localization indices to investigate localization effects of startups and closures at different distances s in the Netherlands. Before presenting these results, the next section describes the data that were used in the analysis.

[1] Instead of using the proportion of the circumference of a circle, we used the observed proportion of the area of the circle with centre x and radius d_{ij} inside the study area. This calculation was performed using postcode areas whose centres of gravity lie within the circle. In order to decide whether a postcode region lies within a circle, the x- and y-co-ordinates of the centres of postcode regions were compared. If the distance between x- and y-co-ordinates of two regions is smaller than d_{ij}, then the area of this postcode region is added to the numerator part of the proportion. The denominator of the proportion is simply $\pi \cdot d_{ij}^2$. Although this method is not exact, the errors in calculating the proportion of the circle within the study area were small. In some cases proportions were calculated with values larger than one. In these cases the proportions were adjusted and manually set to 1.

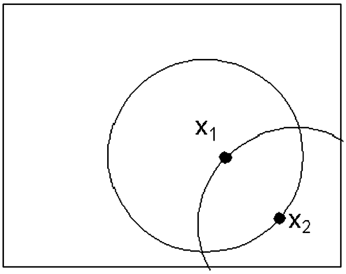

Fig. 2. $w_{ij} \neq w_{ji}$

3. Data: The LISA register of business establishments

The data used for analysing the localisation economies were obtained from the LISA register of business establishments (plants), which was provided by the Ministry of Housing, Spatial Planning and the Environment (VROM). The LISA register (National Information System Labour Markets) holds information on all business plants in the Netherlands, where paid work is being performed. Besides firm plant information the LISA register also holds information on governmental, educational, and public health service plants, as well as free professions (although the latter category is underrepresented in the register due to underreporting). LISA information is gathered by the so-called Regional Co-operative Bodies (RSV's), of which there exist 18 in the Netherlands. The basic unit in the LISA register is a business establishment or plant[2]. For the present research we used LISA-data from 1996 up to 1999[3]. The LISA register contains information per establishment on location (postcode (6-digit)), economic activity (5 digit NACE code), and the number of employed. By comparing different years, it is possible to make statements on starting and closing firms. Unfortunately, LISA-registration numbers for the province of Noord-Holland were not consistently used over the years, so we were not able to link individual firms over time. Therefore we had to exclude this region from the data set (Ekamper et al. 2001). Firms that relocate from one LISA RSV-region to the other are categorized as a closure in the region of departure, and a new firm in the region of destination. Hence, startups and closures include firms that have relocated at the interregional level. However, the number of firms in this category is less than five percent of the total number of starting or closing firms, and therefore they do not dominate the results.

We chose the year 1997 for the spatial clustering analysis. The stock of firms pertains to the situation at the first of May of this year, and startups and closures relate to events that occurred between the first of May of 1997 and one year later. We had geographical coordinates of postcodes at the 4-digit level, so straight-line distances could be calculated between all pairs of 4-digit postcodes in the Netherlands. The total number of 4-digit postcodes in the

[2] Plant and establishment are used here interchangeably. A business plant is defined as a location of a firm, institute, or free profession (i.e., any factory, workplace, shop or other working accommodation, or a complex of these) in which or from where an economic activity or independent profession is performed by one or more employed persons (at least one person for 12 hours per week)

[3] The data were kindly provided by the National Planning Agency as part of a larger research project on constructing a spatial micro-simulation model of firm demography

Netherlands is about 4,000. Establishments were grouped into 16 one-digit economic sectors.

Aggregation to four-digit postcode regions reduces the total spatial information available: all establishments in a given postcode region are now located at a single geographic location (the centroid). The most pronounced influence of this on D is probably over short distances, since the statistic represents a counting process over increasingly large concentric zones. At shorter distances, the lumpiness imposed by the postcode reduction increases the number of counted establishments, effectively turning each individual K function into a step function. In general this will tend to inflate D over shorter ranges. At larger distances, the effects of aggregation are washed out since events at the spatial margin constitute a smaller share of total events (Feser and Sweeney 2000). On average our postcode regions are 9.0 square kilometres (equivalent to the area of a circle with a radius of 1.6 kilometres), but especially in urban regions where most establishments are located (see also Fig. 3) they are smaller (on average in Den Haag 1.38 km, in Rotterdam 2.93 km and in Utrecht 2.16 km). For this reason we believe that the bias introduced by taking postcode centroids as establishment locations is small.

For the analysis of 1997 we have information on 517,655 establishments at our disposal. The total study area equals 31,928 square kilometres, and we find on average 16.2 establishments per square kilometre. When comparing the years 1997 and 1998 we found 56,458 starting firms and 34,358 closures during 1997, which corresponds to a startup rate of 109, and a closure rate of 66.4 per 1000 incumbents.

Figure 3 shows the number of establishments per square kilometre per postcode region. The numbers between brackets in the legend refer to the number of regions that fall in each category. Not surprisingly, the highest density of firms is found in the larger cities. This is obvious for the large cities of the Randstad (Den Haag, Rotterdam and Utrecht), but also for the province capitals. The highest firm density is observed in the centre of Rotterdam, with 1,666 establishments per square kilometre. The actual number of establishments here is 1,014.

The spatial pattern of business establishments is therefore strongly clustered around the major urban centres. When analysing localization patterns of specific industries, we have to control for his overall spatial clustering. An industry is said to be localized if it shows a higher degree of spatial clustering than the overall spatial clustering as depicted in Fig. 3.

In Table 1 the number of establishments by 1-digit economic activity codes are shown. The 16 different sectors are very diverse in size, ranging from 104 establishments in fishery (B) to almost 167 thousand repair of consumer goods and trade establishments (G). The table also shows the number of starting and closing firms, and the startup and closure rate.

4 Results: All business establishments

In order to understand and interpret the K- and D-statistics, we start with a detailed example based on data for all firms in the Netherlands. The question is whether the demographic events birth and death have a concentrating or a de-concentrating effect on the location of firms in the Netherlands. Afterwards analyses are made sector specific at the SBI'93 1-digit level.

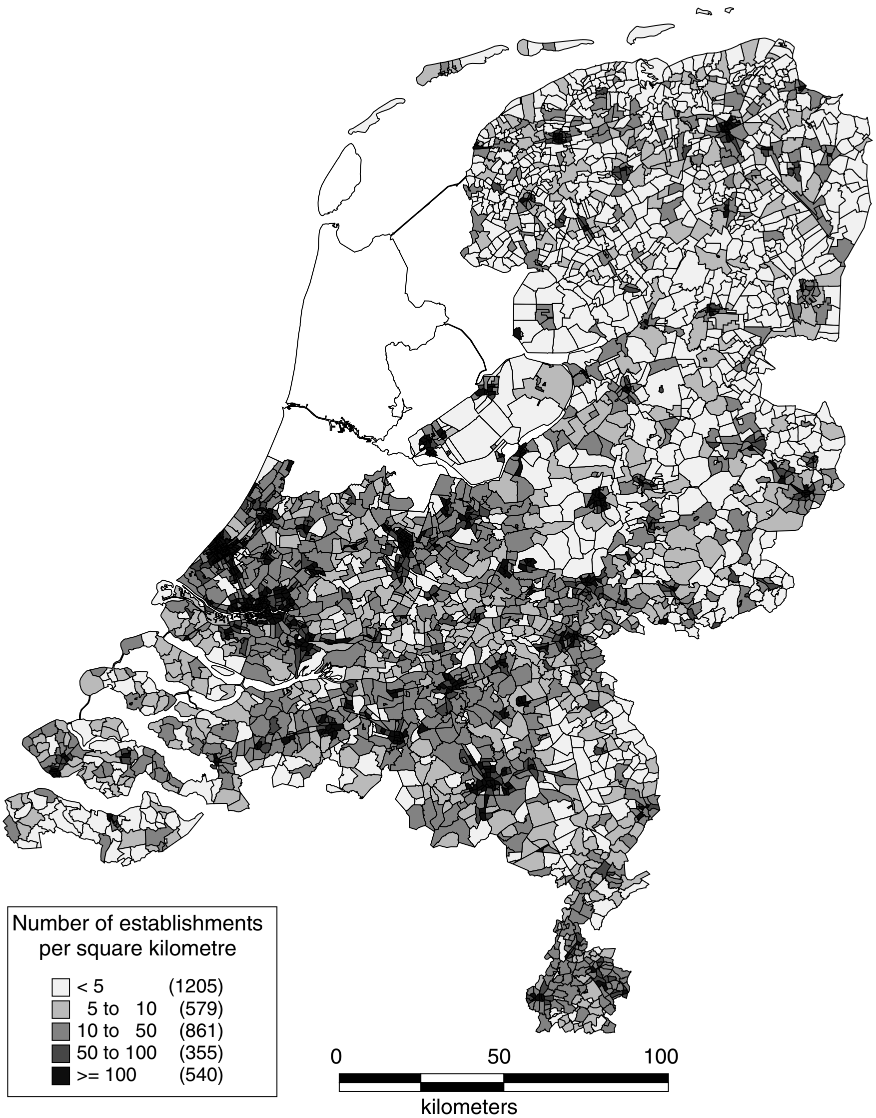

Fig. 3. Number of establishments per square kilometre per postcode region in 1997

Figure 4 shows the K-statistics for three different K_{ij} combinations. In this case the i refers to firms that closed down during 1997, and the j refers to firms that existed at the beginning of 1997. The K-statistics were calculated following functions (7), (8) and (9). The K_{11} function counts the average number of closing firms surrounding a closing firm in circles with increasing distance from the firm, the K_{12} function counts the average number of existing firms surrounding a closing firm in a similar fashion, and the K_{22} function counts the average number of existing firms surrounding an existing firm. We report values up to a radius of 30 kilometres, which is of the order of magnitude of

Table 1. Number of firms by sector and event, 1997

SBI'93 1-digit sector	1997		Number of		In percentage	
	Total number	In percentage	Starters	Closures	Starters	Closures
A: Agriculture-hunting-forestry	560	0.1	98	27	17.5	4.8
B: Fishery	104	0.0	17	6	16.3	5.8
C: Extracting minerals	6,133	1.2	713	295	11.6	4.8
D: Industry	38,232	7.4	3,552	2,429	9.3	6.4
E: Public services	448	0.1	30	48	6.7	10.7
F: Construction industry	40,212	7.8	4,883	1,945	12.1	4.8
G: Repair of consumer goods and trade	166,788	32.2	15,357	11,745	9.2	7.0
H: Catering industry	33,824	6.5	2,603	1,906	7.7	5.6
I: Transport storage and communication	22,841	4.4	2,535	1,710	11.1	7.5
J: Financial institutions	17,885	3.5	1,906	1,322	10.7	7.4
K: Letting and commercial services	91,961	17.8	15,857	7,441	17.2	8.1
L: Public administration and social security	3,715	0.7	168	237	4.5	6.4
M: Education	18,223	3.5	1,180	1,002	6.5	5.5
N: Health care and welfare	32,168	6.2	2,555	1,382	7.9	4.3
O: Culture recreation and other services	44,407	8.6	4,989	2,851	11.2	6.4
Other	154	0.0	15	12	9.7	7.8
Total	517,655	100.0	56,458	34,358	10.9	6.6

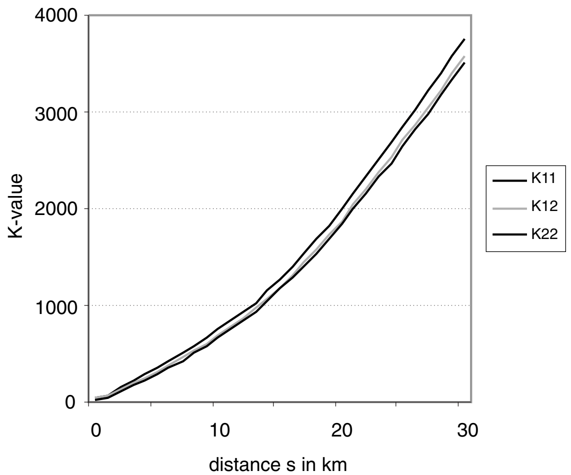

Fig. 4. K-statistics for closing and existing firms (all sectors), 1 = closing firms, 2 = existing firms

the radius of a province. A similar exercise was performed for starting firms: again K-statistics for three different K_{ij} combinations were calculated. Now the i refers to firms that started during 1997, and the j again refers to firms that already existed at the beginning of 1997.

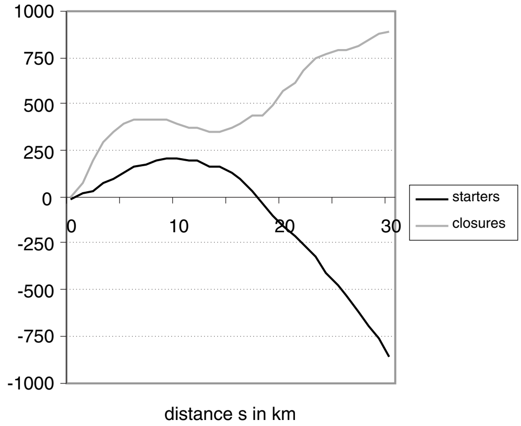

Fig. 5. ACI(s)-values for starters and closures (all sectors)

Since we want to know whether the demographic events enhance clustering of firms or not, we calculated the $D(s)$-function as the difference between the K_{12} and the K_{22} function. Figure 5 reveals the Absolute Concentration Index (ACI) for both starting and closing firms. Positive values indicate that an excess number of existing firms surrounding a starting or closing firm is found, relative to the number of firms expected based on the distribution of existing firms to each other. In the case of starting firms, this means a *concentration* effect. In the case of firms that close down, positive values refer to a *deconcentration* effect. Negative values of D indicate the opposite: a deconcentration effect for starting firms, and a concentration effect for closing firms. A value of for example ACI(10) = 211 for starting firms, means that based on the distribution of existing firms to each other, we find, within a distance of 10 kilometres surrounding a starting firm, 211 additional existing firms (i.e., a concentration effect).

So far, we were able to indicate whether the demographic events birth and death have a deconcentration or concentration effect on the distribution of firms at different distances. But we do not know the relative effects at different distances (is an excess of 211 firms a lot?), or the combined effect of both events. For making statements about the relative impact on clustering we use the RCI(s), as specified in Eq. (6).

Figure 6 shows the results by length of the radius of the circle. It now becomes clear, that as already found, firms that close down have a deconcentration effect on the location of firms (positive values), but what now also becomes apparent is that the effect is most pronounced at short distances. Within distances 2, 3 and 4 kilometres we find more than 10% additional existing firms surrounding a firm that closed down than expected on the basis of the distribution of existing firms to each other. The figure also depicts a deconcentration effect at a distance of zero kilometres for starting firms (5% less existing firms). However, this result may be biased due to the aggregation of distances at the postcode centroids.

A last step needs to be taken in order to make statements about the combined effect of births and deaths in the population of firms. Clearly, if we find positive values for startups and negative values for closures, the combined effect is one of concentration (Table 2), and similarly, if we find

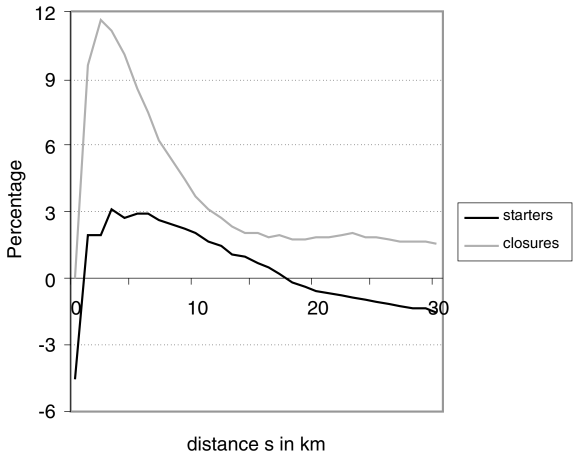

Fig. 6. RCI(s)-values for starters and closures (all sectors)

Table 2. Relation between the signs for WRCI*(s)* of starters and closures and the resulting (de-) concentration effect

		WRCI(s) starters	
		–	+
WRCI(s) closures	+	Deconcentration	Deconcentration or concentration
	–	Deconcentration or concentration	Concentration

negative values for startups and positive values for closures, the result is a net deconcentration effect. For the other possible combinations (positive values for both starters and closures, or negative values for both startups and closures) the net result is not immediately clear. Therefore, it is necessary to know the ratio (r) between the number of firms that started and the number that closed down. The effect for starters and closures go in opposite directions, so the combined effect may go either way. The number of starting firms is 56,458 and the number of closing firms is 34,358. Therefore the RCI(s) for starters is multiplied with 0.62 and the RCI(s) for closures is multiplied with 0.38. If the *Weighted Relative Concentration Index* (WRCI(s) for starters is larger than the WRCI(s) for closures we have a deconcentration effect, and a concentration effect if the opposite is true. In case of negative values the following holds: if a value for starters is more negative than the value for closures we have a deconcentration effect, and a concentration effect the other way around.

In Fig. 7 these WRCI(s)-functions are displayed. At all distances from zero to 30 kilometres, the combined effect of the demographic events birth and death is one of deconcentration. At a distance of zero kilometres it is the effect of starters that is solely responsible for the total deconcentration effect. From distances 1 to 18 kilometres it is the deconcentration effect of closures that outweighs the concentration effect of starters. Thereafter, both startups and closures contribute to the overall deconcentration.

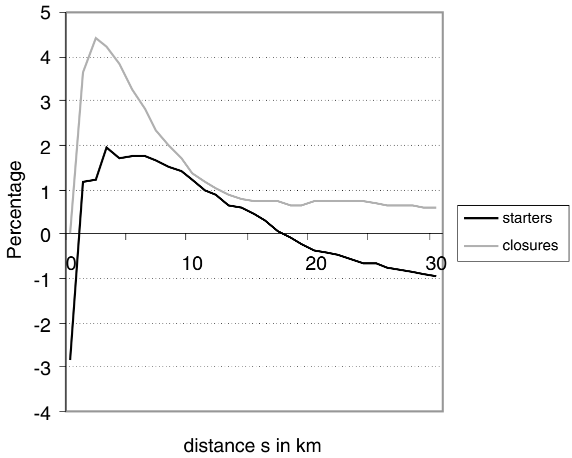

Fig. 7. WRCI(s)-values for starters and closures (all sectors)

5. Results for economic sectors

As for all firms in the Netherlands we performed similar calculations on firms by sector. Instead of giving detailed results, we will focus on specific distances. We will show results for a distance of 2 kilometres (close to the average radius of a postcode area), a distance of 5 kilometres (the average radius of a municipality), a distance of 18 kilometres (the radius of COROP-region) and a distance of 30 kilometres (a province). The COROP classification is a statistical classification designed by the Co-ordinating committee Regional Research-programme. The COROP areas divide the 12 Dutch provinces into 40 nodal regions. This classification matches with the so-called NUTS-III classification, a classification that is often used within the European Union.

Table 3 gives a summary overview of the main results at the chosen geographical scales. This table shows for all 1-digit industries for both demographic components startups and closures whether startups (closures) are spatially relatively more close to existing units (+) or more dispersed (−). The table also shows the overall net effect of both components: If the weighted overall effect of both components is spatial deconcentration this is indicated by a gray shade in the cell of the table. Clearly, as discussed in the previous section, if a cell contains (+ −) this implies overall localization (startups are more concentrated, and closures are more dispersed); if the cell contains (− +) this implies deconcentration (startups are more dispersed and closures are more concentrated); this is indicated by a gray shade in the cell). For the other two possibilities (+ +) and (−−) the overall effect depends on the relative weight of both components. The overall localization tendencies for the aggregate of all industries is given in the bottom lines.

The previous section already reported on the deconcentration trend of the total population of business establishments at all distances between 0 and 30 km, and this is reported in the second row from below. The bottom row gives the percentage of industries (at the 1 digit level) showing a net localization trend based on the combined effects of startups and closures. At the

Table 3. Concentration and deconcentration effects for starters (S) and closures (C) , for one-digit level industries

Economic activity	s = 2 Postcode		s = 5 Munici-pality		s = 18 Nodal region		s = 30 Province	
	S	C	S	C	S	C	S	C
A: Agriculture-hunting-forestry	+	+	+	+	–	–	–	–
B: Fishery	+	–	+	–	–	–	–	+
C: Extracting minerals	+	+	+	+	+	+	+	+
D: Manufacturing	+	+	+	+	–	–	–	–
E: Public services	+	+	+	+	+	+	+	+
F: Construction industry	+	+	+	+	+	+	–	+
G: Repair of consumer goods and trade	–	+	–	+	–	+	–	+
H: Catering industry	–	+	–	+	–	–	–	–
I: Transport storage and communication	–	+	+	+	–	+	–	+
J: Financial institutions	+	+	+	+	+	+	+	+
K: Commercial services	+	+	+	+	+	–	–	+
L: Public administration and social security	+	+	+	–	+	–	+	–
M: Education	+	+	+	+	+	+	+	+
N: Health care and welfare	–	+	–	+	–	+	–	+
O: Culture recreation and other services	–	+	–	+	–	+	–	+
Other	+	–	+	–	+	–	+	–
% Industries with localization by component (+ for startups; – for closures)	69	12	75	19	50	44	38	31
Total plants	+	+	+	+	–	+	–	+
% Industries with overall localization tendency	44		50		44		31	

municipal level (corresponding to a radius of 5 km) half of the industries show a localization tendency as a result of demographic events. At all other spatial scales the percentage is lower. At the provincial scale deconcentration is the rule: only 5 out of 16 industries show a localization tendency. This subset of concentrating industries includes the manufacturing sector. This is interesting since most past studies on agglomeration economies pertain to the manufacturing sector. At both the local level of the postcode and the agglomeration level of the nodal region only 44% of all one-digit economic sectors show a localization trend.

The table also shows the different role of both demographic components. Deconcentration is primarily the result of closing firms, who tend to be relatively more co-located with incumbents than surviving firms. This is most strongly the case at the local level of the postcode and the municipality, but also at the higher spatial scales this tendency prevails. Starting firms tend to choose locations relatively more closely to incumbents than existing firms, but this tendency decreases with higher geographical scales. Startups lead to localization for the majority of industries at the postcode and municipal level, but only for half of the industries at the nodal region level, and even less at the regional level of the province. At the more detailed level of the individual industries, we observe that only two industries show localization at all spatial scales: extracting minerals, public administration and ‘other’. Seven industries show deconcentration at all spatial scales. These include the largest sectors in

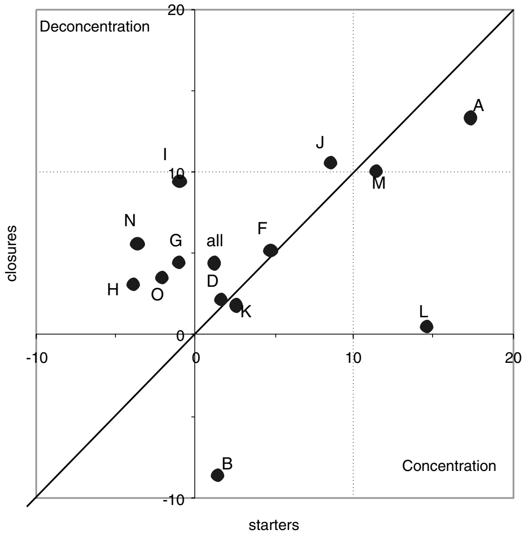

Fig. 8. WRCI(s)-values, at postcode area level (s = 2) for 1-digit economic sectors

terms of the number of firms: trade, catering, and financial services. Figures 8 and 9 show the position of all industries in the two-dimensional space of localization tendencies of both components. Figure 8 shows the relative concentration and deconcentration effects for starters and closures at the postcode level (a radius of 2 kilometres). Economic sectors located in the upper left quadrant of the chart experience deconcentration because of the combined effect of closing and starting firms, whereas economic sectors in the lower right quadrant of the chart experience a concentration trend[4]. In Fig. 9, similar values, but now at the nodal region level (s = 18 kilometres) are plotted. Labels in both figures refer to the sector-codes mentioned in Table 3.

In summary, deconcentration prevails among the industries in this analysis, and this tendency increases with larger geographical scale. Moreover, this spatial deconcentration is primarily driven by plant closures. New firm openings by itself show a tendency of localization, although this tendency is weaker at higher geographical scale. A possible explanation for these diverging tendencies is that new firms tend to be located in designated areas, as a result of strict zoning regulations in the Netherlands. For instance, municipalities tend to concentrate manufacturing firms in industrial areas. Similarly, new office space tends to be concentrated in special areas and building complexes. This may explain why startups show a localization tendency, especially at lower geographical scales. Why closures contribute to deconcentration is less clear. It is tempting to conclude that existing firms in concentrated areas, such as city centres, have a smaller survival probability than more dispersed firms. This would mean that for existing firms the

[4] Three sectors are excluded from the chart, since their values for *WRCI*(*s*) would disturb the picture. These are "C extracting minerals", "E public services" and "other". In Fig. 9, the sector "other" is excluded from the chart.

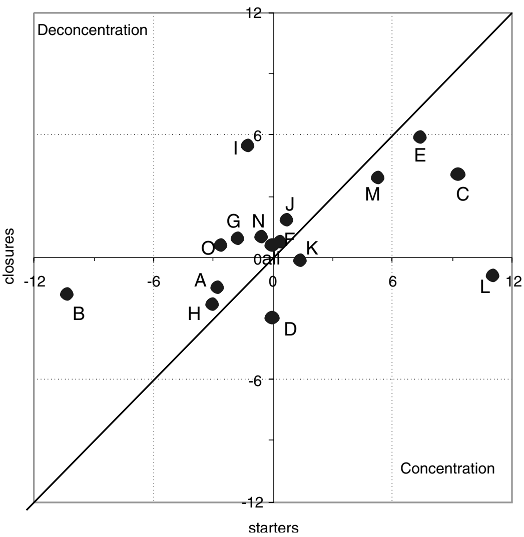

Fig. 9. WRCI(s)-values, at nodal region level (s = 18) for 1-digit economic sectors

negative effect of spatial competition is stronger than the positive effect of localization economies. But if this is true, why would new firms tend to concentrate? Another explanation is that many closures are in fact relocating firms who, in a later stage of the life cycle, move away from clusters to more peripheral locations. More research is necessary to look into these life cycle aspects of geographical concentration tendencies.

An interesting question is whether localization trends, as materialized in relative locations of starters and closures, is related to existing localization patterns of industries. Localization trends are measured relative to the existing pattern of the industry in question. Do industries, which show a high degree of localization in their spatial pattern tend to become relatively less or more spatially clustered? In other words, are existing localization patterns reinforced or weakened as a result of industry dynamics of startups and closures? In order to answer this question, Figs. 10 and 11 show the localization index of both the existing pattern of industries (at the horizontal axis) as well as the localization index of industry dynamics on the vertical axis. The x-axis pertains to the localization score of the incumbent plans for each manufacturing industry, *vis-à-vis* the total population of plants. Positive values indicate that the industry is more clustered than all other industries; negative values indicate that the sector is less clustered than on average. The agglomeration effect for demographic events on the y-axis is calculated as the difference between the weighted net effect of startups and closures. Positive values stand for an overall concentration tendency caused by the two demographic events, relative to the pattern of incumbents in the same industry, and negative values for an overall deconcentration effect relative to incumbents of the same industry. The figures show whether or not existing agglomeration patterns are reinforced by the demographic events birth and death. At the postcode level there is some evidence that industry dynamics weakens existing localization patterns: the number of industries in the first and third quadrant of the figure is larger than that in the second and fourth

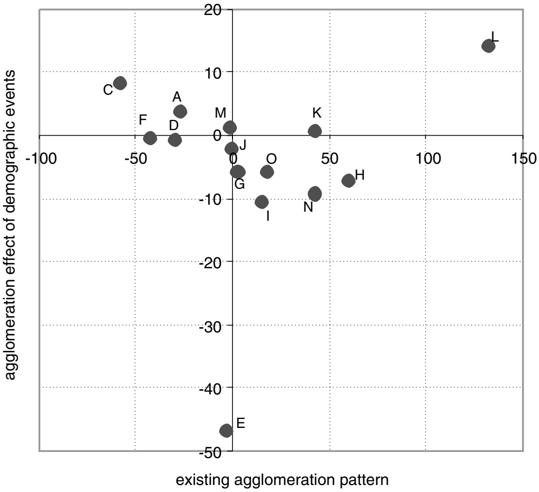

Fig. 10. Localization indices of existing industry patterns (horizontal axis) and net clustering effect of startups and closures (vertical axis) at the postcode (s = 2 km) level

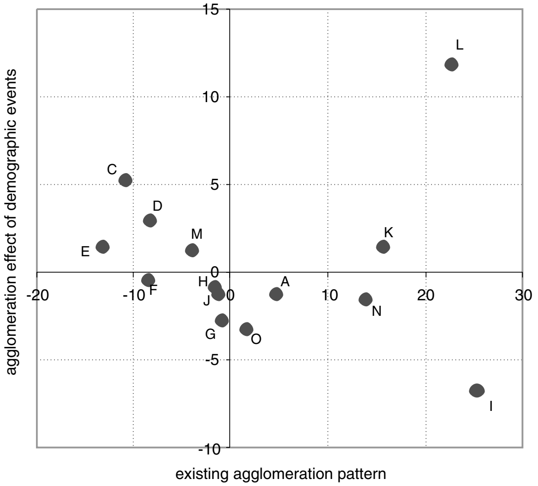

Fig. 11. Localization indices of existing industry patterns (horizontal axis) and net clustering effect of startups and closures (vertical axis) at the nodal region (s = 18 km) level

quadrant. In other words: localized industries tend to become less localized, and dispersed industries tend to become more localized. At the agglomeration level (s = 18) this pattern is less clearly visible.

If we compare these results with that of Dumais et al. (2002), who also looked at the spatial effects of dynamic events at the plant level for manufacturing, we find a number of striking differences. Their analysis is area-based at the state level in the US, so this is a totally different geographical scale. Although differences exist in method and geographical units, they found that "new firm births and expansions of existing plants have a de-agglomerating effect whereas the plant closure process tends to reinforce concentration levels". For our data on the manufacturing sector this is only partly true, as can be seen in Fig. 12. Starters only have a de-agglomeration effect at very short distances (0 and 1 kilometres) as well as (slightly) after 18

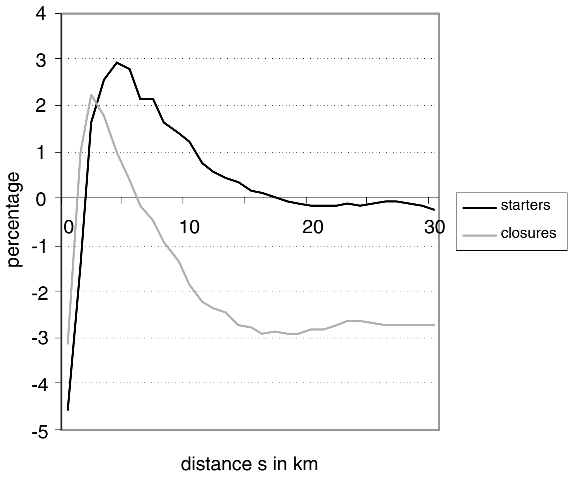

Fig. 12. WRCI(s)-values for starters and closures, sector D manufacturing

kilometres. Within a distance of 2–17 kilometres we actually find a concentration effect for the process of births. This effect is most pronounced at distances around 5 kilometres, which is, on average, the municipality level. The process of firm deaths shows more similarities between their and our analysis: for distances at 0 kilometres and from 6 kilometres onwards we also find a reinforcement of concentration levels. Between 1 and 5 kilometres the reverse is true. The combined effect of both processes is one of concentration from 3 kilometres onwards. In this respect manufacturing establishments behave differently than all firms on average: there we found an overall deconcentration effect for these processes.

All in all the results of Dumais et al. support the conclusion that economic agglomeration is a fairly stable phenomenon in the US. In contrast, in the Netherlands, having a geographical scale which is comparable to the state level of the US, local, urban and regional sprawl is much more dominant, and this is reflected in the results here. Since we are dealing with the localization effects of dynamic events of flows into and out of the population the results cannot directly be compared to results pertaining to agglomeration characteristics of stocks of firm populations. Nevertheless, it is interesting to note that Duranton and Overman (2002), analysing UK data find that at most half of the industries show a tendency to localization, and that localization predominantly takes place at smaller spatial scales, up to 50 kilometres.

6. Conclusions

In this article we studied the localization effects of the dynamic demographic events of startups and closures for economic sectors in the Netherlands. We employed a distance-based statistic, which was originally developed in mathematical biology, but can also be employed for detecting localization patterns of industries. A major advantage of this method is that it does not suffer from aggregation bias that is inherent in area-based methods. This method takes the existing pattern of incumbent plants as the point of reference. In other words, it controls for the existing spatial clustering of the industry. Plant openings and closures can either reinforce or weaken the existing localization pattern. We studied these localization tendencies for industries at the one-digit level.

The main results of this analysis can be summarized as follows:

- The geographical scale is important in measuring localization (spatial concentration/clustering/agglomeration) processes. We found, in line with the conclusions of Duranton and Overman (2002) or Feser and Sweeney (2000), for Dutch economic sectors that processes are different at various spatial scales. For instance, the localization effect of startups is positive up to 18 km, but negative at larger distances.
- All business plants when taken jointly, exhibit a clear deconcentration trend as a result of firm startups and closures. This is mainly the result of closures: surviving firms are relatively less clustered. At a geographical scale up to the nodal region level startups have a positive localization effect, which is outweighted by the effect of closures; at larger geographical scales both startups and closures work in the direction of deconcentration.
- The localization tendencies of startups, especially at smaller spatial scales, may be due to zoning regulations, where industrial areas and office space tends to become more concentrated at the municipal and regional level.
- A possible explanation for the lower survival probabilities of firms in clusters would lead to the conclusion that for existing firms local competition is stronger than the positive effects of localization economies. Another explanation is that relocating firms, who started in clusters, and move to more distant locations in a later phase of the life cycle, have an impact on this process. However, these life cycle aspects of localization need more study.
- The results show that the direction and speed of the process of localization depends on the geographical scale. The concentration effect of startups, and the deconcentration effect of closures are strongest at the municipal level.
- In a majority of the economic sectors at the one-digit level localization as a result of the combined effect of startups and closures is not the rule, and this is even more the case at larger geographical scales. At the municipal level half of the industries show a localization trend, and at the provincial level this is only 31% This is consistent with the tendency towards spatial sprawl at the (sub-) local, agglomeration and regional level in the Netherlands.
- The localization tendency of the manufacturing industry is different from this prevailing trend towards deconcentration. In the Netherlands the manufacturing industry shows a clear concentration trend except at the postcode level. This spatial clustering is mainly the result of closures: here surviving manufacturing business plants tend to be more clustered. Closing firms tend to be located outside the spatial clusters at all spatial scales.
- There is no strong relation between localization tendencies as a result of startups and closures on the one hand, and the degree of localization of the existing pattern of the industry on the other hand. There is some evidence, however, that industries that are more clustered, tend to become less clustered through startups/closures, and industries that are less clustered, tend to become somewhat more clustered. This regression to the mean

process, although it is not a strong tendency, would in the long run lead to a convergence of spatial clustering patterns among economic sectors.

Overall, we find no strong tendencies of clustering trends in the Netherlands in the chosen time period, due to startups and closures, at the 1-digit economic sector level. Nevertheless, for manufacturing, localization tendencies are present, except for the very small level of the postcode. This localization trend is caused by the demographic process of plant closures. Hence, there is some evidence from this analysis that spatial proximity to other manufacturing plants increases the probability of survival.

References

Adams J, Jaffe A (1996) Bounding the effects of R&D: An investigation using matched establishment-firm data. Rand Journal of Economics 27:700–721

Arrow KJ (1962) The economic implications of learning by doing. Review of Economic Studies 29:155–173

Arthur WB (1990) Silicon Valley locational clusters: when do increasing returns imply monopoly? Mathematical Social Sciences 19:235–251

Ciccone A (2002) Agglomeration effects in Europe. European Economic Review 46:213–227

Diggle PJ (1983) Statistical Analysis of Spatial Point Patterns. London Academic Press, London

Diggle PJ, Chetwynd AG (1991) Second-order analysis of spatial clustering for inhomogeneous populations. Biometrics 47:1155–1163

Dumais G, Ellison G, Glaeser EL (2002) Geographic concentration as a dynamic process. The Review of Economics and Statistics 84:193–204

Duranton G, Overman HG (2002) Testing for localisation using micro-geographic data. Centre for economic performance, London School of Economics and Political Science

Ekamper P, Huisman C, van Wissen L (2001) Bedrijvendynamiek in Nederland met behulp van TabFirms version 1.0: Een demografische database van LISA-gegevens in de periode 1996–1999 [Firm dynamics in the Netherlands by means of Tabfirms version 1.0: A demographic database of LISA firm data in the period 1996-1999]. NIDI, The Hague

Ellison G, Glaeser EL (1997) Geographical concentration in US manufacturing industries: A dartboard approach. Journal of Political Economy 107:889–927

Ellison G, Glaeser EL (1999) The geographic concentration of industry: does natural advantage explain agglomeration? The American Economic Review 89:311–316

Feser EJ, Sweeney SH (2000) A test for the coincident economic and spatial clustering of business enterprises. Journal of Geographical Systems 2:349–373

Henderson JV (1986) Efficiency of resource usage and city size. Journal of Urban Economics 18:47–70

Henderson JV (2003) Marshall's Scale Economies. Journal of Urban Economics 53:1–28

Hoogstra G, van Dijk J (2004) Explaining firm employment growth: does location matter? Small Business Economics 16 (forthcoming)

Jaffe A, Tratjenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108:577–598

Klepper S (2002) The capabilities of new firms and the evolution of the US automobile industry. Industrial and Corporate Change 11:645–666

Marshall A (1890) Principles of Economics. Prometheus Books, New York

Nakamura R (1985) Agglomeration economies in manufacturing. Journal of Urban Economics 17:108–124

Van Oort FG (2004) Urban growth and innovation. Spatially Bounded Externalities in the Netherlands. Aldershot, Ashgate

Romer PM (1986) Increasing returns and long-run growth. Journal of Political Economy 94:1002–1037

Van Wissen LJG (2002) Demography of the firm: A useful metaphor? European Journal of Population 18:263–279