# Agglomeration Economies and the Location of New Information and Communication Technology (ICT) Firms in the Netherlands

Frank van Oort[1] and Oedzge Atzema[2]

[1] RPB Netherlands, Institute for National Spatial Policy Analysis, The Hague, and Utrecht University, Faculty of Geographical Sciences, The Netherlands.
E-mail: f.vanoort@geog.uu.nl
[2] Utrecht University, Faculty of Geographical Sciences, The Netherlands.
E-mail: o.atzema@geog.uu.nl

Paper prepared for the ERSA-conference, August 2002, Dortmund.

May, 2002

***Agglomeration Economies and the Location of***

***New Information and Communication Technology***

***(ICT) Firms in the Netherlands***


**Abstract**

This paper looks at the determinants of births of new establishments in the information and communications technology (ICT) sector among 580 municipalities in the Netherlands. In particular, we examine the role of agglomeration economies and other locational attributes in determining where new firms locate. Agglomeration economies facilitate knowledge transfer and are thus expected to be important determinants of entrepreneurial activity. We find that more industrially diverse urban municipalities that are already relatively specialised in ICT and producer services attract more start-up establishments than other municipalities. Both proximity and heterogeneous (non-contiguous) urban structures at the local, regional and national level are significantly attached to localised firm formation rates. This result supports previous evidence that high-technology enterprises tend to collocate in areas where economic activity is spatially dense.


## 1 Introduction

The role of agglomeration economies in economic growth has long been a central theme in urban and regional economics. Additionally, this topic has taken on greater importance in years since seminal contributions by Romer (1986) and Lucas (1988) modelled growth in an endogenous framework. In these types of models, knowledge spillovers between economic agents, an important source of agglomeration economies, play a crucial role in the growth process leading to external economies of scale in production. A large and growing empirical literature has been built around testing this idea using data from cities (Glaeser *et al.* 1992, Lambooy 2002 and Henderson *et al*. 1995). The reasoning here is that if knowledge spillovers are important to growth and firm dynamics, they should be more easily identified in cities where many people are concentrated into a relatively small geographic space so that knowledge can be transmitted between them more easily. Most studies along these lines, however, have focused on overall employment growth (see for an overview Van Oort 2002a), so they do not consider the role of spatial externalities in fostering the formation of new firms, or entrepreneurship. This paper looks at how agglomeration economies, actually indicators of knowledge spillovers, affect births of new establishments, using a unique data set for the Netherlands. The focus of the analysis is on the information and communications technology (ICT) sector. Development of high-technology industries has long been the subject of general interest among social scientists and there have been related investigations of the computer industry (i.e., Beardsell and Henderson 1999). The data utilised in the present study provide counts (relative to population) of newly established and incumbent businesses and their employment levels by industry for 580 municipalities (cities) over a five-year period extending from 1996-2000. The approach taken here is quite similar to that in Rosenthal and Strange (2002), who analysed determinants of establishment births in United States zip codes using Dun & Bradstreet Marketplace data, and Van Soest *et al.* (2002), who analysed new firm agglomeration determinants in the Dutch province of Zuid-Holland, using similar data as applied in this paper. While the U.S. data have the advantage that more is known about each establishment in the data set, the Dutch data provide information about establishment births and growth over a longer time period. A longer time period over which to measure births is expected to provide a clearer picture of which types of areas and which regional characteristics are most attractive to entrepreneurs. The remainder of this paper is divided into three sections. Section 2 describes the data used. Section 3 provides some background for the study concerning agglomeration hypotheses tested for in the remainder of the paper. In section 4 spatial contiguity and spatial heterogeneity as econometric modelling tools are introduced, and applied to the Dutch data. Section 5 then focuses on Explanatory Spatial Data Analysis (ESDA) as a tool to statistically map contiguity-based agglomeration patterns over space. Applications to the dataset are presented, especially on new firm formation rates of ICT-firms. This section also introduces shortly the indicators used in the econometric analyses applied in the subsequent section. The econometric results then are presented in section 6. This section also compares estimates developed to those obtained in related previous studies. Section 7 concludes.

## 2 Data and industry classification of ICT firms

Utrecht University and ETIN Consultants Tilburg have initially created the dataset of ICT firms in the Netherlands used in this paper (Atzema 2001). Although ICT as being a current 'leading enabling technology' is used in almost all sectors of the Dutch economy (Van der Laan *et al.* 2001, Bouwman and Hulsink 2000), we have limited our research to ICT providing firms, which also includes services industries. Even within this limitation, the presentation of a statistical overview of the ICT firms in the Netherlands is severely hampered by the deficiency of official sources of data. Official Dutch data files arranged on industries do not appoint unambiguous SIC-codes to "the" ICT sector. The research population for this paper therefore had to be composed of combinations of detailed (5-digit) existing SIC-codes. Unfortunately, there is no unity in definition about such a composition between several authorities. The European Information Technology Observatory (EITO) for example includes software production, but excludes the consumers' electronics. The OECD on the contrary, chose the opposite in its' statistics. In this research we apply the definition of Statistics Netherlands (CBS), which makes, just like comparable institutions in Canada and Finland, a distinction between ICT service industries and ICT goods industries. We have refined this definition and made a distinction into initially eight main groups of ICT activities, summing up 25 sectors at the 5-digit level of definition (see table 1). Nevertheless, we are aware of the problem that sectors on the base of SIC-codes, even on the 5-digit level, are made up of collections of firms whose outputs are quite disparate in terms of their involvement in the ICT business. In order to avoid the indicative value of the research by industry incomparability, the sectors are aggregated into ICT-manufacturing (of hardware and software), ICT-distribution and ICT-service activities (see also Van Soest *et al.* 2002).

Table 1 Employment in ICT-sectors in the Netherlands, 1998

| Industry (SIC-code)* | N of jobs | % of jobs |
|---|---|---|
| Production: | | |
| Production of hardware (7220,72101) | 9,154 | 4,7 |
| Production of software (2233,3002,3220) | 46,196 | 24,1 |
| Trade: | | |
| Wholesale trade of ICT products (51641, 51642, 51657) | 27,603 | 14,4 |
| Retail trade of ICT products (52454, 52481, 52494) | 4,443 | 2,3 |
| Services: | | |
| Internet/(multi)media, telecom (6420) | 35,722 | 18,7 |
| Data- and computer centres (7230, 7240) | 10,701 | 5,6 |
| ICT Consultancy (72102, 74141, 74143, 74204, 74846) | 54,498 | 28,5 |
| Other kinds of (ICT) producer services (7133, 7250, 7260) | 3,149 | 1,6 |
| Total | 191.466 | 100 |

* SIC-codes are the SBI93-codes as used by Statistics Netherlands (see Van Oort 2002a)

The population of ICT firms has been collected in a two-step procedure. In a time consuming first step the Yellow Pages for all regions in the Netherlands have been screened for selecting firms from the following business categories: software, automation, internet, tele- and data communication. This selection consists of 12,878 ICT firms in the Netherlands (Atzema 2001). This method has two disadvantages: not every company will be entered in the Yellow Pages and information about the extant employment lacks. Therefore we completed the dataset in a second step of the procedure, in which the file obtained through the Yellow Pages has been joined with the nationally covered LISA-file (Erdman 2000, Van Oort 2002a). This LISA-file yearly registers the employment of over 750,000 companies and institutions in the Netherlands. So we used the first step to select relevant SIC-codes in the second, more comprehensive step. Both files have been compared with one another and the definite file has been extended with other companies from the LISA-file. This results in a file of 18,985 ICT firms for 1998. The number of jobs in ICT firms contributes nearly 3% of the total employment in the Netherlands. The ICT providing sector is still a relatively small sector in the Netherlands and the same is true for other Western Europe countries (Van Ark 2000). Furthermore, it becomes clear that the employment in the Dutch ICT sector is dominated by service activities like consultancy, internet providing and whole sale trade. Within the field of production activities, the production of software dominates.

Several additional alterations on the data were carried out for this paper especially. Employment function, location quotients and concentration and specialisation indicators are calculated as average over the years 1966-2000. Growth indicators (defined as in Van Oort 2002a) compare the average stock of firms over 1996 and 1997 with the average stock of firms over 1999-2000. This in order to minimise (spatial or temporal) outlier dependency. The firm level data are aggregated into three broad ICT-industries: production, distribution and services (see above)[1]. Further, the firm level data are aggregated into 580 locations, which represent municipalities (situation 1998). The four largest municipalities (Amsterdam, Rotterdam, The Hague and Utrecht) are split into 3-digit zip-code areas in order to make distinctions in harbour, central location and edge-city locations within municipalities possible. This resulted in 36 observations for Amsterdam, The Hague, Rotterdam and Utrecht. In 1998 The Netherlands is build up by 548 municipalities, the four largest are replaced by the 36 3-digit zip code areas (still referred to as municipalities), making in total 580 observations. Because of the longitudinal, firm level database, new and incumbent firm populations could be distinguished.

## 3 Agglomeration hypotheses

In line with the international literature, as indicators of agglomeration economies, economic diversity, spacialisation and local competition indicators are applied in this paper. These statistical indicators are broader than commonly used 'pure' innovation indicators, like patent-citation (Van Oort 2002b, Van Soest *et al.* 2002). For example, knowledge spills over between firms via informal contacts between

[1] Notice that growth rates are used *as reference* for new firm formation rates, since both indicators are complementarily to each other in terms of regionalised employment growth (Ashcroft and Love 1996). Both components of change- and industrial classification detail are not (yet) fully used in analyses.

employees, or because employees switch jobs and take their knowledge with them. Indeed, the most important type of knowledge that plays a role in growth and innovation processes is not necessarily path-breaking innovations, but may be learning opportunities for everyday people (Glaeser 1999). Empirical tests of this theory often have looked at cities to identify settings in which these external factors most effectively foster economic firm dynamics.

Results, however, have been sharply divided. On the one hand, Glaeser *et al.* (1992) and Feldman and Audretsch (1999) find that employment growth and firm dynamics is enhanced by diversity of activity across a broad range of sectors. Henderson *et al.* (1995), Black and Henderson (1999), and Beardsell and Henderson (1999), on the other hand, find faster growth when more activity is concentrated in a single sector (specialisation). While endogenous (technological) 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. 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, firm formation and innovation in urban areas (Landry 2000).

Knowledge-based theories of endogenous development are tested at the city (municipal) level in this paper. 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 local or regional specialisation will foster growth and new firm formation. Furthermore, (local) market power is also thought to stimulate firm dynamics as it allows the innovating firm to internalise a substantial part of the rents. A possible conjecture in this regard is that a local competition variable (at the municipal level) is an indicator of both product market and labour market competition for non-manufacturing establishments (e.g. ICT services) that sell goods and services only locally, but an indicator of just labour market competition for manufacturing establishments (e.g. ICT manufacturers) that are more likely to sell in national or even worldwide markets (Feldman and Audretsch 1999, Van Soest *et al.* 2002). The second hypothesis, proposed by Porter (1990), also states that knowledge is predominantly sector-specific, but argues that its effect on growth and firm dynamics is enhanced by local competition rather than market power as firms need to be innovative in order to survive. The 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 other sectors. Table 2 summarises the spatial externality circumstances distinguished in these respective hypotheses. 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. This paper will empirically relate these hypotheses (controlling for sectoral and spatial heterogeneity) to spatial patterns of new firm formation in ICT firms in the Netherlands.

Table 2 Stylised hypothesised relations of agglomeration circumstances with innovation and economic growth

| | MAR | Porter | Jacobs |
|---|---|---|---|
| Concentration | + | + | – |
| Diversity | – | – | + |
| Competition | – | + | + |

## 4 Modelling spatial proximity and spatial heterogeneity

Marshall (1890) in the past and Krugman (1995) in the present, argue that proximity may matter for economic growth, new firm formation and innovation in high-technology industries because of tacit knowledge. The marginal cost of transmitting tacit knowledge rises with distance (Audretsch 1997). As tacit knowledge and human interaction become more valuable in the innovation process, geographical proximity becomes crucial to the innovation process. The exchange of tacit knowledge may require a high degree of mutual trust and understanding. During the 1990s some statistical-empirical evidence on the importance of proximity for innovation is presented in the Anglo-Saxon literature (Feldman 1999). Audretsch and Feldman (1996) conclude that while the cost of transmitting information may be increasingly invariant to distance, presumably the cost of transmitting tacit knowledge rises with distance. Baptista (2000, p.531) concludes that “externalities related to knowledge tend to grow stronger as the geographical units of reference become smaller”. The literature discussed, concludes upon two contradicting agglomeration theses. On is that the location of R&D activity, firm dynamics or growth in a particular place is determined more by the location of innovation or growth in other sectors than by the location of its own production; the other however finds evidence for the opposite (specialisation or concentration) thesis. Most of the relevant empirical papers focus on American states as spatial unit of analyses. Some Anglo-Saxon research, however, focuses on lower scales of analysis. Anselin *et al.* (2000) and Wallsten (2001) use metropolitan statistical areas to analyse the spatial extent of R&D and growth externalities and find that local spatial externalities are present and important. A tentative conclusion of the theory and empirical studies presented so far is that knowledge externalities (spillovers) become less important with increasing distance. Proximity matters in the transmission of innovation- and growth-based knowledge of dynamic (incumbent and new) firms, while distance decays tend to be rather steep (Jaffe *et al.* 1993).

Up to this point, the literature summarised in this and the previous section stresses the proximity hypothesis of R&D intensity and high-technology firm dynamics within agglomeration externalities. But the literature to a large extent discusses non-contiguous (regime types of) spatial dependence as well. Other research (see for an overview Van Oort 2002b) for instance finds that quality of life aspects and city size are significant locational considerations both to professional workers and to growing ICT firms (Verlinde and Van Oort 2002). Also emphasised are city-size (non-contiguous) or metropolitan spatial distributions as a relevant spatial regime for explaining the geography of firm dynamics and innovation. Some studies (e.g. Cortright and Mayer 2001) stress high-tech *employment specialisation* in metropolitan areas as indication of innovative competitiveness, while others (e.g. Frenkel 2001) instead finds a negative relationship between specialisation and R&D (high-tech industry) intensity. Economic *diversity* on the contrary is found most crucial for technological development in metropolitan areas by Florida (2001) and Feldman and Audretsch (1999). Paci and Usai (2000) in a spatial-econometric framework find *both* concentration and diversity indicators of local economic structures important determinants of firm level innovation intensity in labour market areas in Italy.

The spatial structures of proximity (contiguous nearness on the municipal level) and heterogeneity (urban hierarchical and regional spatial dependence) in this paper have been captured in spatial lag (or spatial error) estimates and spatial regimes, respectively. The spatial coefficient in spatial lag estimation shows whether the dependent variable in a model (in our case localised innovation intensity) is dependent on neighbouring values of this dependent variable. If so, conclusions can be reached on the significance and magnitude of this spatial dependence (Anselin 2000). Not all hypothetical spatial dependence is pure contiguous in character though, as becomes clear from careful studying of earlier empirical studies (Van Oort 2002, see especially chapters 2 and 3). Especially urban-hierarchical spatial dependency on meso (regional) and macro levels of spatial aggregation are mentioned in the literature. Spatial heterogeneity is therefore modelled by spatial regimes on meso-level labourmarket (commuting) induced connectedness (figure 5), macro-level national zoning (Randstad core region, intermediate zone and national periphery, figure 6) and degree of urbanisation (dichotomy, cut-off population threshold of 45,000 inhabitants). Within the Randstad core economic region, a division in northwing (Amsterdam-Utrecht) and southwing (The Hague-Rotterdam) is often made, especially concerning the location of ICT firms (Atzema 2001). A fourth spatial regime is therefor applied in the econometric analyses, distinguishing the north- and southwing municipalities from other locations in the Netherlands (figure 7). See also Van Oort (2002, appendix IV) for more exact specifications. These forms of spatial heterogeneity make up four spatial levels of urban constellation: the urban level itself, the functional (commuting) region, the meso-level 'agglomerative fields' of the northwing and soutwing of the Randstad core region (Stroeken *et al.* 2002) and finally the macro-economic core-periphery (Randstad, intermediate zone, national periphery) distinction.

## 5 Descriptive analysis and statistical indicators

A technique most conveniently used for spatial statistical descriptions of data is exploratory spatial data analysis (ESDA). This is a set of techniques aimed at describing and visualising spatial distributions, at identifying atypical localisations or spatial outliers, at detecting patterns of spatial association, clusters or hot spots, and at suggesting spatial regimes or other forms of spatial heterogeneity (Anselin 1995). These methods provide both measures of global and local spatial autocorrelation, which will be technically discussed briefly. In this section exploratory spatial data analysis (ESDA) using global and local indicators of spatial association are applied to the Dutch data on the firm- employment-, and growth structure and distribution over the municipalities.

*Global spatial autocorrelation* can be defined as the coincidence of value similarity with locational similarity (Anselin 2000). Positive spatial autocorrelation occurs when high or low values of a random variable tend to cluster (agglomerate) in space,; negative spatial autocorrelation occurs when geographical areas tend to be surrounded by neighbours with very dissimilar values. The measurement of global spatial autocorrelation in this chapter is based on Moran's *I* statistic, which is the most widely known measure of spatial clustering. For each year or period (of change) of observation, this statistic is given by:

where $x_{it}$ is the observation in region $i$ and year (period) $t$, $m_t$ is the mean of the observations across regions in year (period) $t$, $n$ is the number of regions and $w_{ij}$ is the interregional element of the spatial weight matrix $W$. This matrix contains the information about the relative spatial dependence between the $n$ regions $i$ and $j$. The elements on the $w_{ii}$ diagonal are set to zero whereas the elements $w_{ij}$ indicate the way region $i$ is spatially connected to region $j$. $S_0$ finally is a scaling factor equal to the sum of all elements of $W$. For row-standardised spatial weight matrices, which are the preferred way to implement the Moran's $I$ test statistic, the normalising factor $S_0$ equals $n$, since each row then sums to 1 (see Anselin 1995, p.22-1 and further). The statistic of the above equation then simplifies to the ratio of a spatial cross products to variance. This makes Moran's $I$ similar but not equivalent to a correlation coefficient; it is not centred around 0. The theoretical mean of Moran's $I$ is $-1/N-1$. The expected value is thus negative and is only a function of sample size ($N$). This mean will tend to zero as the sample size increases. The theoretical variance of Moran's $I$ depends on the stochastic assumptions made. Either the assumption of a normal distribution of variables in question (normality assumption), the assumption that each value observed could equally likely have occurred at all locations (randomisation assumption) or a randomisation approach using a reference distribution for $I$ that is generated empirically (permutation assumption) can be tested for[2]. Albeit all three variance assumptions were tested for on employment and population structure and development indicators, in this chapter only the results for the randomisation assumption will be presented[3]. Inference is based on a standardised z-value of $I$ that is computed by subtracting the theoretical mean and dividing the result by the theoretical standard deviation. A positive and significant z-value for Moran's $I$ (as can be judged from accompanying low probability values) indicates positive spatial autocorrelation. Similar values of the variable, either high or low, are more spatially clustered than could be caused purely by chance. In contrast, a negative and significant z-value for Moran's $I$ indicates negative spatial autocorrelation, the opposite of spatial clustering[4]. The results for Moran's $I$ are to a large extent determined by the choice of the spatial weight matrix. In general, a pattern of decreasing autocorrelation with increasing orders of contiguity is typical of many spatial autoregressive processes[5]. Table 3 displays global indicators for spatial autocorrelation for location quotients of (aggregated) employment in ICT firms over the period 1996-2000 (employment function) for new firm formation rates (average 1996-2000, new firms relative to average population, explained variable in section 6 of this paper), for employment growth in all ICT-firms for the period 1996-2000 and in only incumbent firms (present in all years of observation in the dataset). From the table it becomes clear the new ICT-firm formation rates show high degrees of spatial association. Growth functions are

2 The software used for testing is SpaceStat (Anselin 1995).

3 The three approaches implicate the use of different models. Inference results from the normal distribution- and (10000-) permutation approach (see Anselin 1995a, p.22-2) are because of economising reasons not presented in this paper. Results of the three models of inference specification are very similar in terms of significance though; all directions and magnitudes of spatial association are confirmed.

4 The concept of negative spatial autocorrelation is harder to grasp; it reflects a lack of clustering, more so than would be the case in a random pattern. Perfect negative spatial autocorrelation is represented by a checkerboard pattern.

5 In the analyses made for this paper first, second as well as third order distance weight matrices were used for spatial autoregressive modelling, while in the paper itself the emphasis will be on first order weight matrices only. When appropriate, differing conclusions from applying different weight matrices is mentioned.

in general less spatially clustered than the employment function and the firm formation rates (this was also observed in earlier studies, see Van Oort 2002, chapter 4). Employment growth in incumbent firms in ICT distribution activities and ICT services is not significant. The information concerning the specification of spatial dependence in table 3 will be used for econometric model specifications in section 6 of this paper.

Table 3 Moran's *I* statistics for log employment function (location quotient), new firm formation rates (log, percentage population) and log employment growth rates, (Netherlands 1996-2000, n=580, randomisation assumption)

| | Moran's *I* w_1* | Standard. dev. w_1 | standard. value w_1 | standard. value w_2 | standard. value w_3 |
|---|---|---|---|---|---|
| **Employment function (all ICT firms)** | | | | | |
| Total | 0.1075280 | 0.002870 | 38.072 | 21.617 | 14.805 |
| ICT production | 0.0350667 | 0.002870 | 12.819 | 7.962 | 5.589 |
| ICT distribution | 0.0298202 | 0.002870 | 10.993 | 7.283 | 5.319 |
| ICT services | 0.0961176 | 0.002870 | 34.092 | 19.745 | 13.601 |
| **New firm formation rate (% pop.)** | | | | | |
| Total | 0.18749410 | 0.002869 | 65.946 | 40.812 | 27.636 |
| ICT production | 0.12084660 | 0.002869 | 42.716 | 27.162 | 18.565 |
| ICT distribution | 0.07359253 | 0.002870 | 26.248 | 18.462 | 13.331 |
| ICT services | 0.15556050 | 0.002870 | 54.804 | 33.747 | 23.169 |
| **Employment growth all firms** | | | | | |
| Total | 0.01664020 | 0.002856 | 6.431 | 3.927 | 2.906 |
| ICT production | 0.01918721 | 0.002858 | 7.317 | 5.393 | 4.452 |
| ICT distribution | 0.00400914 | 0.002862 | 2.004 | 3.081 | 3.571 |
| ICT services | 0.00477644 | 0.002860 | 2.274 | 1.503* | 1.182* |
| **Employme. growth incumbent firms** | | | | | |
| Total | 0.00450938 | 0.002859 | 2.132 | 1.632* | 1.114* |
| ICT production | 0.01672780 | 0.002858 | 6.457 | 5.437 | 4.303 |
| ICT distribution | 0.00199073 | 0.002845 | 1.307* | 1.691* | 1.449* |
| ICT services | 0.00040811 | 0.002853 | 0.462* | 0.984* | 0.946* |

* The expected value for Moran's *I* statistic is constant over each sector, both for employment and number of firms: $E(I)$=-0.002.
All statistics except those marked * are significant at $p$=0.01.

Moran's *I* statistic is a global statistic: it does not enable us to take into account the regional and local structure of spatial autocorrelation. However, which regions or locations contribute most to the global spatial atocorrelation, if there are specific local or regional clusters of high or low values and to what point the global evaluation of spatial autocorrelation masks atypical localisations or 'pockets of nonstationarity' (deviations from the global pattern) are interesting questions. Analysis of *local spatial autocorrelation* can be carried out using two tools. The first is the Moran scatterplot, which can be used to visualise local spatial instability. The spatial lag $Wz_I$ is plotted against the original values $zi$, resulting in four different quadrants of the scatterplot that correspond to four types of local spatial association between a location and its neighbours. The HH quadrant comprise locations with a high value surrounded by locations with high values. LH locations with a low value surrounded by locations with high values, LL locations with a low value surrounded by locations with low values and HL locations with a high value surrounded by locations with low values. HH and LL refer to positive spatial autocorrelation, indicating spatial clustering of similar values, whereas LH and HL represent negative spatial autocorrelation indicating spatial clustering of dissimilar values. The locations in each quadrant can be mapped. The Moran scatterplot may thus be used to visualise atypical localisations and the use of standardised variables allows Moran scatterplots to be comparable across time. The global spatial autocorrelation may also be visualised in the graph since Moran's *I* is formally equivalent to the slope coefficient of a linear regression of $Wz_i$ on $z_i$ using a row-standardised weight matrix. This regression can therefor be assessed with diagnostics for model fit.

The Moran scatterplot does not give any indications of significance of spatial clustering, and therefor it cannot be considered as a Local Indicator of Spatial Association (LISA, see Anselin 1995). This second tool for local statistics can be used to test the hypothesis of random distribution by comparing values of each specific localisation with the values in the neighbouring locations. Anselin defines a local indicator of spatial association as any statistics satisfying two criteria. First, the LISA for each observation gives an indication of significant spatial clustering of similar values around that observation; second, the sum of the LISA for all observations is proportional to a global indicator of spatial association. The local version of Moran's *I* statistic for each region *i* and year (period) *t* can then be written as:

where the summation over *j* is such that neighbouring values (contiguous analysis) or values within a predefined distance (full distance analysis with or without cut-off) of *j* are included. The sum of local Moran's statistics is then:

From the first presented equation above it follows that the global Moran's *I* statistic (for a row-standardised weight matrix, so $S_0=n$) is indeed proportional to the mean of the local Moran's statistics:

Positive values for $I_{it}$ indicate clustering of similar values (high or low), whereas a negative value indicates clustering of dissimilar values[6]. Anselin (1995) gives two interpretations for local Moran's statistics. They can be either used as indicators of local spatial clusters (called hot spots), which can be identified as locations or sets of neighbouring locations for which the LISA are significant or as diagnostics for local instability, i.e. for significant outliers with respect to the measure of global

[6] For a technical discussion on inference and significance levels see Anselin (1995a) and Le Gallo and Ertur (2000).

spatial autocorrelation. This second interpretation of the LISA statistics is similar to the use of a Moran scatterplot. In this interpretation we will apply the two tools of local spatial autocorrelation in following sections. Localised Moran scatterplot maps are presented for location quotients of (aggregated) employment in ICT firms over the period 1996-2000 (employment function, figure 1) for new firm formation rates (average 1996-2000, new firms relative to average population)[7], for employment growth in all ICT-firms for the period 1996-2000 (figure 3) and in only incumbent firms (present in all years of observation in the dataset, figure 4)[8]. Only figures for all ICT activities aggregated are presented, the figures of individual sectors (production, distribution and services) can be obtained from the authors. Notice that the relative large share of service activities among ICT firms (table 1) determine the spatial pattern of localised autocorrelation to a large extent(see also table 3). The employment function of ICT firms (averaged over 1996-2000, see figure 1) shows a large degree of concentration of firms in the Randstad core region of the Netherlands. But also parts of Gelderland and Noord-Brabant (the intermediate zone) show relative high concentration patterns. The location of newly established ICT firms over the research period (figure 2) differs from the employment function considerably. New ICT firms are not unambiguously represented in the southwing of the Randstad region (The Hague-Rotterdam). Instead, the northwing of the Randstad region and municipalities more remote from the Randstad core region show high concentration values. As remarked in Bleichrodt et al. (1992), who found similar but less profound patterns of new ICT firm location, this indicates a filtering down process of ICT technology using firms, radiating from the Randstad region outwards. New firm formation automatically induces employment growth (as long as these new firms survive, see Ashcroft and Love 1996). Employment growth in incumbent firms though (figure 4) is clustered in the traditional high-technology centres in the Netherlands (compare Wever and Stam 1999). The similarity with figure 1 (spatial *concentrations* of employment) is striking. Figures 4 (growth of incumbent firms) and 2 (new firm formation) together make up the total growth figures presented in figure 3. Remarkably, new firm formation appears to attribute much more to this growth pattern than incumbents growth. In the period 1996-2000, employment in incumbent firms grew from 122,433 to 149,727 persons (22% increase), while total employment in both incumbents and new

[7] This is a quite common definition of new firm formation, see Ashcroft and Love (1996) and Bleichrodt *et al.* (1992). Alternative definitions though relate new firm accounts to the existing stock of firms (e.g. Storey 1982, Van Wissen 2000). The sensitivity of the (spatial) outcomes of our research questions for alterations in these definitions turned out to be limited. Therefore, results for other definitions of populations of new formed firms are not presented, but are obtainable by the authors on request.

[8] See for detailed explanations of (alternative) growth measures and definitions Van Oort (2002a).

firms rose from 161,534 (including later bankrupt firms in the base year) to 226,969 persons (41% increase). The difference in growth rates is accounted for by employment growth created by new ICT firms established during the research period, a pattern commonly found for especially service based activities (compare Van Oort 2002a, chapter 3). Sections 6 will focus on these patterns in more detail, relating indicators of agglomeration economies to the firm formation patterns.

The relative small size of the Netherlands provides a natural control for much location-specific heterogeneity. In fact, several variables enumerated in related studies (Henderson, Kuncoro, and Turner 1995, Glaeser 1999) are potentially important location-specific factors that may affect either employment growth or establishment birth rates either are roughly constant between locations in South-Holland, or else can be at least partially controlled. Cultural differences between locations in the Netherlands are small. Variations in taxes, environmental amenities (such as climate), and environmental regulations between locations are quite small. Differences in prices of non-land inputs exhibit little variation across the country. Prices charged for energy inputs vary by sector, but within a sector, they are the same throughout the Netherlands. Wages also vary by sector, but not much within sectors. Thus, wage rates within a sector would be uniform and there is little need to control for labour force characteristics such as level of education, percent of workers with particular skills, or percent of workers who are union members (see Van Oort 2002a for actual testing of these elements). Cities (in our case: municipalities) are fertile grounds for testing knowledge-based theories of endogenous growth. Dense urban agglomerations provide opportunities for learning because they are frequently centers of knowledge creation. Electronic communications infrastructure generally is well developed and face-to-face meetings between key people desiring to share knowledge are certainly easier to arrange than they would be in rural areas. In fact, if electronic and face-to-face communications are complements rather than substitutes, as Glaeser has hypothesised, many firms may see a decided cost savings from locating in urban areas rather than rural areas. Prior studies looking at employment growth rates in cities have tested three (in some respects competing) hypotheses concerning the way in which knowledge spillovers affect growth, see section 3. This paper tests how well these theories predict one aspect of urban employment growth; i.e., that part of growth arising from entrepreneurship or the birth of new establishments. Focusing on establishment births sets this paper apart from the related literature on employment growth and facilitates analysis in at least two ways. First, initial economic conditions prevailing in an area at the beginning of the sample period can arguably be treated as exogenous determinants of births. In other words, new establishments can be viewed as taking initial conditions as given and then deciding where to locate. Second, new establishments do not have a prior history of location, input, and output choices that needs to be accounted for. Thus, this analysis avoids the frequently intractable problem of collecting historical establishment-level data on capital stocks. A Dutch municipal data set on sectoral employment structures is used to construct indicators of various types of agglomeration economies (as hypothesised in section 3) that are similar or as close as possible reminiscent to those used in prior studies (see especially Glaeser *et al*. 1992 and Henderson *et al*. 1995). The agglomeration indicators are not constructed by means of the ICT-database itself, both for technical reasons (multicolinearity) and for theoretical reasons (agglomeration economies are commonly defined in a national, aggregated setting). The base-year approach applied on variables facilitates testing whether effects of different types of agglomeration economies on growth and innovation persist over time and space (Van Oort 2002a). *CONCENTRATION* is defined as a location quotient showing the percentage of employment accounted for by an industry in a municipality relative to the percentage of employment accounted for by that industry in the Netherlands. This indicator especially comprises localisation or specialisation economies. *COMPETITION* is measured as establishments per worker in a municipality and industry divided by establishments per worker in that industry in the Netherlands It indicates whether establishments in industries tend to be larger or smaller in a municipality compared to the country as a whole. This spatial indicator of relative firm size fits in a tradition of identifying common labour market competition and market structure indicators. Glaeser *et al.* (1992) interpret this variable as a measure of local competition on the assumption that competition is more intense among a larger number of smaller establishments than among a smaller number of larger establishments. This interpretation, however, has been called into question by Combes (2000), who contends that it may measure internal diseconomies of scale, and by Rosenthal and Strange (2000), who view it as a broader measure of local industrial organisation. Several variables were tried as a measure of industrial diversity to indicate how evenly employment in a municipality is spread across economic sectors. *DIVERSITY,* the Gini-coefficient for the distribution of employment by sector in a municipality (or zip code), measures the absence of diversity: The locational Gini-coefficient has a value of zero if employment shares among industries are distributed identically to that of total employment in the reference region (across 49 sectors in the Netherlands, of which the ICT-sector is only a minor part, see section 1). A value of 0.5 results if employment is concentrated in only one sector. Lower values of *GINI* thus implicate higher degrees of diversity. The diversity indicator is treated as indicator of urbanisation economies. Results presented in the next section can be used to make at least a suggestive test of the three hypotheses outlined earlier[9].

[9] Because of space limitations, correlation diagnostics of all explanatory variables used in this paper are not presented. No correlation higher tan 0.5 in absolute terms was allowed in the analyses.

## 6 Empirical results

In tables 4 and 5, the econometric models run are summarised. Below the tables technical explanation on the models is given. The models are numbered over the two tables, models (1) till (4) in table 4, models (5) and (6) in table 5. First, column (1) in table 4 gives the OLS estimation of the percentage of new firm formation. Concentration indicators are given by the location quotient if ICT-firms relative to the total amount of employees on average present in the municipalities (*CONCENTRATION ICT-FIRMS*). Concentration indices for industrial, distribution, business service- and consumer service activities are introduced in the model (see Van Oort 2002a for an exact definition of these activities). The degree of diversity is measured by a Gini-coefficient, which in fact measures the *lack* of diversity. Localised competition is, in line with the Glaeser *et al.* (1992) approach, measured by relative firm size, both for ICT-firm and for all firms in the localised economy aggregated. Employment levels of ICT-firms and all firms in the local economy together are introduced for correcting for absolute differences in starting values for firm dynamics and growth in the base year. The OLS-model shows the significance of both concentration indicators (of the 'own' ICT-sector, as well as in general for business services in a positive sense, and for consumer services in a negative sense) and the diversity indicator.

Table 4 OLS and combined spatial lag and spatial regime models for (log) new firm formation in ICT activities in the Netherlands (n=580, 1996-2000, t-values in parentheses)

| *Explanatory variables* | (1) OLS | (2) Spatial lag | (3) Spatial lag Urban regimes | | (4) Spatial lag Macro-zoning regimes | | |
|---|---|---|---|---|---|---|---|
| | | | Urban | Non-urban | Randstad | Int. zone | Periphery |
| CONSTANT | 0.536<br>(0.933) | -0.375<br>(-0.726) | -0.120<br>(-0.208) | -0.092<br>(-1.455) | -0.377<br>(-0.484) | 0.231<br>(0.092) | -0.195<br>(-0.700) |
| CONCENTRATION ICT-FIRMS | 0.789<br>**(7.867)** | 0.687<br>**(7.614)** | 0.654<br>**(7.361)** | 0.056<br>(0.601) | 0.616<br>**(6.384)*** | -0.062<br>(-0.086)* | 0.056<br>(0.707)* |
| CONCENTR. INDUSTRY | -0.029<br>(-1.239) | -0.022<br>(-1.008) | -0.008<br>(-0.347) | -0.092<br>(-1.455) | -0.041<br>(-1.199) | 0.016<br>(0.403) | -0.062<br>(-1.158) |
| CONCENTRATION DISTRIBUTION | -0.119<br>(-1.144) | -0.199<br>**(-2.665)** | -0.142<br>(-1.771) | --0.660<br>**(-3.544)** | -0.189<br>(-1.585) | -0.030<br>(-0.204) | -0.133<br>(-0.952) |
| CONCENTRATION BUSIN. SERVICES | 0.292<br>**(4.394)** | 0.188<br>**(3.153)** | 0.239<br>**(3.854)*** | -0.314<br>(-1.640)* | 0.257<br>**(2.796)*** | 0.392<br>**(3.382)*** | 0.062<br>(0.485)* |
| CONCENTRATION CONS. SERVICES | -0.234<br>**(-2.244)** | -0.238<br>**(-2.539)** | -0.059<br>(-0.590)* | -1.106<br>**(-4.887)*** | -0.072<br>(-0.505)* | 0.182<br>(0.089)* | -0.418<br>**(-2.575)*** |
| LACK OF DIVERSITY | -1.114<br>**(-2.934)** | -0.559<br>(-1.639) | -0.820<br>**(-2.167)** | -0.575<br>(-0.676) | -1.133<br>**(-2.011)** | -1.503<br>**(-1.940)** | -0.684<br>(-1.116) |
| SIZE ICT-FIRMS (COMPETITION) | 1.029<br>**(19.820)** | 0.815<br>**(17.429)** | 0.793<br>**(15.985)** | 0.091<br>**(7.133)** | 0.731<br>**(10.192)** | 0.744<br>**(7.268)** | 1.106<br>**(11.967)** |
| SIZE ALL FIRMS (COMPETITION) | -0.465<br>**(-6.353)** | -0.352<br>**(-5.338)** | 0.011<br>(0.052)* | -0.357<br>**(-5.108)*** | -0.404<br>**(-3.907)** | -0.280<br>**(-1.943)** | -0.406<br>**(-3.358)** |
| EMPLOYMENT 1996 ICT-FIRMS | 0.148<br>(1.534) | 0.073<br>(0.847) | 0.084<br>(1.002) | -0.052<br>(-0.590) | 0.084<br>(0.917) | 0.068<br>(0.096) | -0.055<br>(-0.694) |
| EMPLOYMENT 1996 ALL FIRMS | -0.073<br>(-0.678) | 0.007<br>(0.008) | -0.038<br>(-0.383) | 0.056<br>(0.595) | -0.016<br>(-0.130) | -0.068<br>(-0.094) | 0.556<br>(0.595) |
| SPATIAL COEFF. ρ | - | 0.969<br>**(26.843)** | 0.973<br>**(22.948)** | | 0.974<br>**(23.241)** | | |
| *Sum. Statistics* | | | | | | | |
| N | 580 | 580 | 580 | | 580 | | |
| $R^2$/*ML* | 0.657/*-474.63* | *-421.386* | *-400.74* | | *-404.96* | | |
| LM (BP) | 4.095 (0.393) | 6.840 (0.077) | 1.527 (0.216) | | 2.981 (0.170) | | |
| LM (ρ) | 324.78 (0.000) | - | - | | - | | |
| LM (λ) | 56.61 (0.000) | - | - | | - | | |
| LR (ρ) | - | 106.49 (0.000) | 114.54 (0.000) | | 106.86 (0.000) | | |
| Chow-Wald | - | - | 42.822 (0.000) | | 33.810 (0.051) | | |

Values of log-likelihood are not comparable over populations of all and old establishments. LM (ρ) and LM (λ) are statistics for the presence of a spatial lag in the dependent variable and in the residual respectively, following Anselin *et al*. (1996) with a critical value of 3.84 at 5% level of significance (marked +). LR(ρ) tests for the significance of the spatial dependence coefficient. LM (BP) tests for homoscedasticity of regression errors using the Breusch-Pagan Lagrange multiplier test for normal distributed errors. The spatial weight matrix used is w_1 (row standardised), probability levels (p-values) are presented in the tables. Significant p-levels are printed bold. The spatial Chow-Wald test is distributed as an F variate and tests for structural instability of the regression coefficients over regimes (Anselin 1995a, p.32-2). Significant results (95% confidence interval) of the spatial Chow-Wald in general and on individual coefficients (rejection of $H_0$ of joint equality of coefficients over regimes) are marked (*). All variables are log transformed and corrected for extreme values (found in ESDA analyses discussed in section 5).

Table 5 Combined spatial lag and spatial regime models for (log) new firm formation in ICT activities in the Netherlands (n=580, 1996-2000, t-values in parentheses)

| *Explanatory variables* | (5) Spatial lag Connectedness regimes | | (6) Spatial lag Randstad regimes | | |
|---|---|---|---|---|---|
| | Connected | Non-connected | Northwing | Southwing | Other |
| CONSTANT | -0.357 | 0.198 | -0.393 | 0.376 | -0.277 |
| | (-0.419) | (0.254) | (-0.666) | (0.096) | (-0.066) |
| CONCENTRATION ICT-FIRMS | 0.683 | 0.626 | 0.639 | -0.472 | 0.082 |
| | **(3.629)** | **(5.968)** | **(7.188)** | (-0.093) | (0.068) |
| CONCENTR. INDUSTRY | -0.004 | -0.056 | -0.026 | -0.070 | 0.040 |
| | (-0.014) | (-1.765) | (-1.104) | (-1.228) | (0.600) |
| CONCENTRATION DISTRIBUTION | -0.280 | --0.119 | -0.156 | -0.600 | 0.189 |
| | **(-2.848)** | (-1.068) | (-1.772)* | **(-3.068)*** | (0.825)* |
| CONCENTRATION BUSIN. SERVICES | 0.237 | -0.038 | 0.225 | -0.439 | 0.580 |
| | **(2.821)*** | (-0.433)* | **(3.290)*** | **(-2.153)*** | **(2.904)*** |
| CONCENTRATION CONS. SERVICES | -0.383 | -0.127 | -0.078 | -01.008 | -0.040 |
| | **(-3.086)** | (-0.864) | (-0.698)* | **(-4.882)*** | (-0.144)* |
| LACK OF DIVERSITY | -1.005 | -0.132 | -0.928 | -1.046 | -1.646 |
| | **(-2.253)*** | (-0.241)* | **(-2.140)** | (-1.246) | (-1.401) |
| SIZE ICT-FIRMS (COMPETITION) | 0.909 | 0.757 | 0.755 | 0.886 | 0.952 |
| | **(13.637)** | **(11.893)** | **(14.247)** | **(5.815)** | **(5.011)** |
| SIZE ALL FIRMS (COMPETITION) | -0.215 | -0.509 | -0.382 | 0.002 | -0.309 |
| | **(-2.487)** | **(-5.035)** | **(-4.940)** | (0.011) | (-1.306) |
| EMPLOYMENT 1996 ICT-FIRMS | 0.209 | 0.035 | 0.077 | 0.112 | -0.746 |
| | (1.081) | (-0.368) | (0.915) | (0.100) | (-0.062) |
| EMPLOYMENT 1996 ALL FIRMS | -0.097 | -0.026 | -0.005 | -0.111 | 0.076 |
| | (-0.488) | (-0.214) | (-0.006) | (-0.098) | (0.063) |
| SPATIAL COEFF. ρ | 0.964 | | | 0.967 | |
| | **(20.125)** | | | **(23.005)** | |
| *Sum. Statistics* | | | | | |
| N | 580 | | | 580 | |
| $R^2$/*ML* | *-405.126* | | | *-397.379* | |
| LM (BP) | 2.092 (0.148) | | | 0.053 (0.974) | |
| LR (ρ) | 97.302 (0.000) | | | 94.105 (0.000) | |
| Chow-Wald | 50.057 (0.000) | | | 50.057 (0.000) | |

Values of log-likelihood are not comparable over populations of all and old establishments. LM (ρ) and LM (λ) are statistics for the presence of a spatial lag in the dependent variable and in the residual respectively, following Anselin *et al*. (1996) with a critical value of 3.84 at 5% level of significance (marked +). LR(ρ) tests for the significance of the spatial dependence coefficient. LM (BP) tests for homoscedasticity of regression errors using the Breusch-Pagan Lagrange multiplier test for normal distributed errors. The spatial weight matrix used is w_1 (row standardised), probability levels (p-values) are presented in the tables. Significant p-levels are printed bold. The spatial Chow-Wald test is distributed as an F variate and tests for structural instability of the regression coefficients over regimes (Anselin 1995a, p.32-2). Significant results (95% confidence interval) of the spatial Chow-Wald in general and on individual coefficients (rejection of $H_0$ of joint equality of coefficients over regimes) are marked (*). All variables are log transformed and corrected for extreme values (found in ESDA analyses discussed in section 5).

The third agglomeration indicator, that measures localised (labour market or service market) competition circumstances, shows a positive relationship with new firm formation rates when measured for the 'own' ICT-sector. But this indicator shows a strong negative relationship when measured in general terms, taking all firms into account in a municipality, independent of sectoral composition. Interestingly, these results do not provide unambiguous support for any of the three endogenous development theories discussed in section 3. Results for (own, ICT-) sectoral specialisation support the MAR and Porter hypotheses, but results for industrial diversity do not. Results for industrial diversity support the Jacobs hypothesis. Results for (own, ICT-) levels of localised competition support Porter and Jacob's hypotheses of growth, and not the MAR hypothesis. The general indicators of concentration stress the importance of business service specialisation as important correlate to new firm formation, and the negative influence of consumer service specialisation in general. The general competition indicator is clearly negatively related to firm formation rates, concluding upon MAR's hypothesis of economic dynamics. A very confusing picture, indeed. Yet, results presented still are of interest from the broader perspective of those concerned with the location tendencies of start-up establishments in the ICT sector. These firms tend to cluster in municipalities that already are employment centres, rich in industrial diversity (compare also Van Soest *et al.* 2002).

Column (1)'s test-statistics of LM($\rho$) and LM($\lambda$) clearly reveal the presence of spatial autocorelation dependency of the model (as did table 3). In column (2) therefore the model is estimated using a spatial lag specification. Spatial lag models make use of maximum likelihood estimation techniques, in which the explained variance is no longer an adequate measure for model fitting. The spatial coefficient turns indeed out to be highly significant. Introducing spatial dependency in the model alters the coefficients slightly when compared to the OLS base model. Especially relative specialisation of distribution activities hampers firm dynamics, while industrial diversity no longer is unambiguously connected to new firm formation rates. The likelihood based measure (*ML*, in the summary statistics of the tables), can be used to compare the model fit with that of the basic OLS-model. It turns out that for the new firm formation model, the fit considerably improves when the spatial lag is added to the model, as indicated by an increase in the log likelihood. Heteroscedasticity turns out to be no problem in any of the models estimated (see the LM(BP) statistics in the tables). The interpretation of the model outcomes change when the spatial lag specification is applied: the significance of specialisation and competition indicators, together with the insignificance of the diversity indicator favours the MAR hypothesis.

Columns (3) till (6) give spatial lag estimation, but with the allowance of structural change of coefficient estimates between spatial regimes. Column (3) shows that the concentration indicators work out more favourably in connection to new firm formation in urban municipalities, as opposed to non-urban ones. The re-signification of industrial diversity makes the result theoretically again more ambiguous. The Spatial Chow-Wald test confirms the significance of the spatial regime. The model-fit again improves considerably when compared to the OLS and spatial lag model without the urbanisation regimes. The relations found thus work out most profound in urban environments. This conclusion confirms the urban setting of the endogenous development theories as outlined in section 3. But other definitions of urbanisation appear to be significant as well for ICT business development. Column (4) in table 4 shows that the Randstad region most profoundly 'exhibits' the significant set of agglomeration economies, as opposed to the national periphery and (to a lesser extent) the intermediate zone. The model-fit is slightly less than in the urban regimes model, but still considerably better than the OLS and spatial lag (sec) model. Column (7) in table 5 shows that *within* the Randstad region, especially the northwing (Amsterdam-Utrecht) is characterised by significant agglomeration indicators. This is confirmed by other research as well (e.g. Van der Laan *et al.* 2001, Atzema 2001). The southwing of the Randstad (Rotterdam-The Hague) shows of quite opposite, less favourable agglomeration circumstances than the northwing, especially concerning industrial diversity and the specialisation in business services. Column (6) finally shows the significance of the connected spatial regime, as opposed to the unconnected regime. The analyses show that urbanisation matters for new ICT-firm formation on all different scales of urban analyses in the Netherlands, both defined by contiguous proximity (as envisaged by the spatial lag significance) and by the spatial heterogeneous regimes. This extends the current debate on urbanisation- and localisation externalities (which focuses mainly on proximity based spillovers and knowledge transfer) considerably.

## 7 Conclusions

This paper has empirically investigated determinants of growth in the information and communications technology (ICT) sector in the Netherlands. The empirical investigation makes use of a unique and highly detailed (longitudinal) data set on births of new establishments in these sectors in each of 580 municipalities. The relative small size of the Netherlands South Holland offers control for certain types of unobserved heterogeneity, such as aspects of labor market conditions, that have plagued earlier studies. Results from this analysis suggest that new establishments in the ICT sector tend to be concentrated in urban areas that are already relatively specialised in this sector and that are relatively rich in the presence of other industries. This outcome does not fully support or contradict three theories, of knowledge spillovers, attributed to Marshall-Arrow-Romer, Porter, and Jacobs that frequently have been tested using data from urbanised areas. Yet, it does provide some insights into the types of areas where ICT establishments choose to locate. The analyses show that urbanisation matters for new ICT-firm formation on different scales of urban analyses in the Netherlands, both defined by contiguous proximity (as envisaged by the spatial lag significance) and by the spatial heterogeneous regimes. This extends the current debate on urbanisation- and localisation externalities (which focuses mainly on proximity based spillovers and knowledge transfer) considerably. This result, however, should be treated cautiously because to date, most studies of location determinants have focused on employment growth; relatively few have looked at the component of employment growth arising from establishment births. Additional studies of establishment births will be necessary before it can be known whether results presented here will carry over to other settings. It will be most helpful if some of the needed studies are conducted in an international context.

**References:**


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

Anselin, L. (2000), Spatial econometrics. In: Baltagi, ed., *Companion to econometrics*. Oxford: Basil Blackwell.

Anselin, L., A. Varga and Z.J. Acs (2000), Geographic and sectoral characteristics of academic knowledge externalities. *Papers in Regional Science* 79, pp. 435-43.

Van Ark, B. (2000), The renewal of the old economy. The Netherlands in a cross country perspective. In: L. Soete (ed.), *ICT and the new economy*. Utrecht: Lemma, pp.21-60 (in Dutch).

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

Ashcroft, B. and J.H. Love (1996), Firm births and employment change in the British counties: 1981-89. *Papers in Regional Science* 75, pp. 483-500.

Atzema, O. (2001), Location and local networks of ICT firms in the Netherlands. *TESG Journal of Economic and Social Geography* 92(3), pp.369-78.

Audretsch, D.B. (1997), Technological regimes, industrial demography and the evolution of industrial clusters. *Industrial and Corporate Change* 6, pp. 49-82.

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

Baptista, R. (2000), Do innovations diffuse faster within geographical clusters? *International Journal of Industrial Organisation* 18, pp. 515-35.

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

Black, D. and V. Henderson (1999), Spatial evolution of population and industry in the United States. *AEA Papers and Proceedings* 89, pp. 321-27.

Bleichrodt, H., P.J. Louter and W.F. Sleegers (1992), *Young firms in the Netherlands*. EGI Research Memorandum 02, Erasmus University Rotterdam (in Dutch).

Bouwman, H. and W. Hulsink (2000), eds., *Silicon Valley in the polder. ICT-clusters in the Low Countries*. Utrecht: Lemma (in Dutch).

Combes, P.P. (2000), Economic structure and local growth: France 1984-1993. *Journal of Urban Economics* 47, pp. 329-55.

Cortright, J. and H. Mayer (2001), High tech specialization: a comparison of high technology centers. *The Brookings Institution Survey Papers* January 2001, pp. 1-18.

Erdman, P.H. (2000), *The spatial structure of ICT-industries in the Netherlands*. Utrecht University: Faculty of Geographical Sciences (in Dutch).

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

Florida, R. (2001), Technology and tolerance: the importance of diversity to high-technology growth. *The Brookings Institution Survey Series*, pp. 1-12.

Frenkel, A. (2001), Why high-technology frims choose to locate in or near metropelitan areas. *Urban Studies* 38, pp. 1093-101.

Glaeser, E.L. (1999), Learning in cities. *Journal of Urban Economics* 46, pp. 254-77.

Glaeser, E.L., H.D. Kallal, J.A. Scheinkman and A. Schleifer (1992), Growth in cities. *Journal of Political Economy* 100, pp. 1126-52.

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

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

Jaffe, A.B., M. Trajtenberg and R. Henderson (1993), Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* 36, pp. 577-98.

Krugman, P. (1995), Innovation and agglomeration: two parables suggested by city-size distributions. *Japan and the World Economy* 7, pp. 371-90.

Van der Laan, L., J. de Boom and F.G. van Oort (2001), *Spatial labour markets and the knowledge economy*. Den Haag: RPD Research Report 7 (in Dutch).

Lambooy, J.G. (2002), Knowledge and urban economic development: an evolutionary perspective. *Urban Studies* 39(5), pp.1019-35.

Landry, C. (2000), *The creative city*. London: Comedia Earthscan.

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

Marshall, A. (1890), *Principles of economics*. New York: Prometheus Books.

Van Oort, F.G. (2002a), *Agglomeration, economic growth and innovation. Spatial analysis of growth- and R&D externalities in the Netherlands*. Rotterdam: Tinbergen Institute.

Van Oort, F.G. (2002b), Innovation and agglomeration economies in the Netherlands. *TESG Journal of Economic and Social Geography* 93(3), pp.344-60.

Paci, R. and S. Usai (2000), Externalities, knowledge spillovers and the spatial distribution of innovation. *GeoJournal* 49, pp. 381-90.

Rosenthal, S.S. and W.C. Strange (2002), Geography, industrial organization and agglomeration. *Review of Economics and Statistics* (forthcoming).

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.

Van Soest, D.P., S. Gerking and F.G. van Oort (2002), Knowledge transfer and the location of new information and communication technology establishments in the Dutch province of Zuid-Holland. *Journal of Small Business Management* (forthcoming).

Stroeken, F., C. Nauta and F.G. van Oort (2002), *Urban networks in 2030. Scenario analyses on national scale*. Utrecht: Royal Haskoning (in Dutch).

Storey, D.J. (1982), *Entrepreneurship and the new firm*. London: Croom Helm.

Verlinde, H. and F.G. van Oort (2002), The knowledge worker as locational guide? *Rooilijn* 4, pp.178-183 (in Dutch).

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.

Wever, E. and E. Stam (1999), Clusters of high technology SME's: the Dutch case. *Regional Studies* 33(4), pp.391-400.

Van Wissen, L. (2000), A micro-simulation model of firms: applications of concepts of the demography of the firm. *Papers in Regional Science* 79, pp. 111-34.