Ann Reg Sci (2004) 38:263–290
DOI: 10.1007/s00168-004-0195-8

The Annals of
Regional Science

On the conceptualization of agglomeration economies: The case of new firm formation in the Dutch ICT sector

Frank G. van Oort[1,2], **Oedzge A. L. C. Atzema**[2]

[1] Netherlands Institute for Spatial Research (RPB), P.O. Box 30314, 2500 GH Den Haag, The Netherlands (e-mail: oort@rpb.nl)
[2] Urban and Regional Research Centre Utrecht (URU), Utrecht University, P.O. Box 80115, 3508 TC Utrecht, The Netherlands (e-mail: o.atzema@geog.uu.nl)

Abstract. In this research we look at the factors that determine new firm formation 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. Both proximity (contiguous) and heterogeneous (non-contiguous) structures at the local, regional and national level are significant when considering localised firm formation. This result supports previous evidence that high-technology enterprises tend to co-locate in areas where economic activity is spatially dense. The major point of our argument is that controversial research results in the literature concerning explanatory spatial circumstances that most favorably induce dynamic and innovative externalities (to a large extent) can be attributed to the lack of consistent spatial research designs that allow the modelling of multiple spatial scale and composition effects. More specifically, we argue that the incubation hypothesis needs adjusting to the appropriate spatial levels and units of analysis: that of the agglomerated region. Finally, we argue that the lack of consistent inclusion of life-cycle aspects of firms in the present mainstream literature on dynamic externalities also contributes to controversies in research outcomes. These findings are important for spatial economic policy indicating that investment in new technologies and economic structures should enhance the prospects for spillover effects at the local level.

JEL classification: O12, O18, R11, R30

This article benefited considerably from comments on earlier versions made by participants at the 42nd European Regional Science Conference 2002 in Dortmund, participants of the workshop "Modern Entrepreneurship, Regional Development and Policy" at the Tinbergen Institute in Amsterdam (May 2003), attendants at a presentation at the Deutsches Institut für Wirtschaftforschung (DIW) in Berlin (August 2003), Leo van Wissen (Rijkuniversiteit Groningen) and two anonymous referees. The usual disclaimers apply.

1. Introduction

The role accorded to agglomeration economies in determining growth has long been a central theme in urban and regional economics. In theoretical terms, the topic has acquired greater importance in years following seminal contributions by Romer (1986) and Lucas (1988) where economic growth was modelled 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 and innovation process leading to external economies of scale in production. At the core of the new economic growth theory lies the concept of technological change as a non-rival and partially excludable good (as opposed to the neoclassical view of knowledge as an entirely public good). On this basis, new technological knowledge (as applied in the ICT sector) is usually tacit, meaning that its accessibility, as well as its growth spillovers, are bounded by geographic proximity of high-tech firms or knowledge institutions, and by the nature and extent of the interactions among these actors in an innovation system (Acs 2002). A large and growing empirical literature has grown around testing this idea using data from cities (Glaeser et al. 1992; Henderson et al. 1995; Dumais et al. 2002; Van Oort 2004). The assumption 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 and confined space and where knowledge is transmitted more easily.

Most studies along these lines, however, tend to focus on overall employment growth and as a consequence they do not consider the role of spatial externalities in fostering new firm formation or entrepreneurial activity (Acs et al. 2003). This paper examines how agglomeration economies, actually indicators of knowledge spillovers, affect the creation of new establishments, and draws upon a unique data set for the Netherlands. The analysis focuses on the information and communications technology (ICT) sector, an important growth industry where business start-up and innovation are pronounced (Beardsell and Henderson 1999). Three questions are central in our paper. First we want to determine which measurable agglomeration factors (that are comparable with earlier studies) are connected to new firm formation in the ICT industry in the Netherlands. Second, our analysis focuses on the conceptual spatial configurations that best describe the new firm formation patterns. We ask ourselves what additional role network-based spatial regimes (such as urban hierarchy, labour market areas, national core-periphery distinctions) play compared to the localised proximity-thesis stressed in the literature. In turn our approach is applied to the incubation hypothesis that states that large(r) cities are breeding grounds for new firms and entrepreneurship because of localised knowledge spillovers. And third, we try to evaluate whether entrepreneurship as measured by firm dynamics contributes to the spatial externalities debate. The spatial, longitudinal and sectoral detail of the data allows for more sophisticated testing of these questions than within previous studies. The data provide counts (relative to the 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 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. While the U.S. data have the advantage that more is known about each establishment, the Dutch data provide information about all establishment births and growth. For example, in the Henderson et al. (1995) study, the strategy of analysing all cities in a given industry presented many problems. As a result of disclosure rules, employment data for as many as 30% of cities were censored. The Dutch data set concerns a longitudinal survey on employment in all ICT establishments in the Netherlands; it is therefore expected to provide a clearer picture of the types of areas and the local and regional characteristics that are most attractive to entrepreneurs.

The remainder of the paper consists of five sections. Section 2 provides the background for the agglomeration and incubation hypotheses tested in the analysis. We focus on testable indicators of (sector-specific) localisation economies and urbanisation economies, also called spatial externalities, in line with the international literature. The indicators are important means for answering the first question of the paper ('what agglomeration indicators should be tested for, and how do they perform?'). Section 2 also describes how the agglomeration variables used in the regression models are constructed. A short overview of control variables concerning sectoral and regional economic structures is provided. Section 3 describes in detail the dataset. In Sect. 4 spatial contiguity and spatial heterogeneity as econometric modelling tools are defined and applied descriptively to the Dutch ICT data. This application allows us to answer the second research question ('what spatial configuration best describes the firm formation pattern and how does this configuration confirm to the incubator hypothesis of new firm formation?'). Spatial patterns of new firm formation and employment growth in ICT firms are discussed, using Exploratory Spatial Data Analysis (ESDA) as a tool to statistically map proximity based agglomeration patterns over space. In Sect. 5 a descriptive analysis is presented concerning employment specialisation and growth in ICT firms as well as new ICT firm formation rates, with special reference to the concepts and degrees of urbanisation presented as spatial regimes in Sect. 4. Section 6 then presents econometric analyses concerning the relation between new firm formation of ICT firms and agglomeration indicators defined in Sect. 2, using the spatial (contiguity- as well as non-contiguity) configurations defined and described in Sects. 4 and 5. Section 7 concludes the paper, focusing explicitly on the third research question ('do concepts of firm life-cycles contribute to the spatial externalities debate?') and on implications for spatial policy.

2. Agglomeration economies: Hypotheses and indicators

2.1. Agglomeration hypotheses

In this section we suggest the agglomeration indicators that should be used in order to test for the relation between agglomeration and new firm formation in the ICT industry. In line with the international literature, economic diversity, specialisation and local competition indicators are chosen. These statistical indicators are broader than commonly used 'pure' innovation indicators, like patent-citation (Van Oort 2002; Capello 2002). For example, knowledge spills over between firms via informal contacts between employees, or because employees switch jobs and take their knowledge with them. Indeed, the most important type of knowledge that plays a role in 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 have often looked at cities to identify settings in which these external factors most effectively foster (endogenous) economic firm dynamics. Previous 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 of the 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, diversity and their spatial composition sends an ambiguous message regarding policy choices to promote or manage growth, firm formation and innovation in urban areas (Parr 2002; Rosenthal and Strange 2001).

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) (collectively: "MAR"), contends that knowledge is predominantly sector-specific and hence that local or regional specialisation will foster growth and new firm formation. The theory of Marshallian externalities states that intra-regional spillover effects occur alongside agglomeration effects due to labour market pooling and input sharing (see for recent elaboration Feser 2002 and Rosenthal and Strange 2001). 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). This spatial embedding should ideally be incorporated in empirical analysis. 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 1 summarises the

Table 1. Stylised and hypothesised relations of agglomeration circumstances with innovation and economic growth

	MAR	Porter	Jacobs	Fourth hyp.
Concentration	+	+	−	−
Diversity	−	−	+	+
Competition	−	+	+	−

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. Note that while case study research is able to obtain many specific organisational details but is poor on the question of applicability of high-tech spatial regularities to situations elsewhere, the measurable set of indicators in our paper allows for interregional comparison of locations at the loss of organisational detail.

2.2. *The incubator hypothesis*[1]

The incubation hypothesis represents a comparative and static framework in which new and small firms are compared with older and larger firms. The hypothesis was an attempt by Hoover and Vernon (1959) to use neoclassical location theory to explain the location of new firms in metropolitan areas and their subsequent spatial development. The central process in the incubation hypothesis "is one in which persons aspiring to go into production on a small scale have found themselves less obviously barred by a high cost structure at the centre of the urban area than at the periphery" (Hoover and Vernon 1959, p. 47). Leone and Struyk (1976) restated the incubator hypothesis as: "small manufacturing establishments beginning operations will find it to their comparative advantage to locate at highly centralized locations within the metropolis". This advantage was mainly derived from the availability of rentable production space (land costs may be high, but property costs are relatively low; see Fagg 1980), inputs, labour, and other services at central urban locations, but also lower supply risks and rapid communication possibilities with customers and suppliers. In short, the explanatory framework was primarily that of agglomeration externalities. The incubation hypothesis was transformed into a general theory of intra-urban location behaviour. This theory was not just applicable to small manufacturing firms but to all new firms. The theory consisted of two central hypotheses: the 'simple' and the 'dynamic' (or 'complex') hypothesis. The simple hypothesis proposed that highly centralized locations attract a disproportionate number of new firms and/or the employment associated with new firms. The dynamic hypothesis proposed that new firms which are formed in high density areas move outward from such sites in their early years of existence in order to expand their productive activities (Leone and Struyk 1976). The dynamic hypothesis

[1] This section is based largely on Stam (2003, p. 32).

introduces some new elements as it implies that the central urban areas not only have higher establishment birth rates (simple hypothesis) but also have higher rates of business failure. Furthermore, successful, mature firms move to lower density areas for at least three reasons. First, physical expansion is easier and cheaper in lower density areas. Second, firms are more able to achieve internal economies of scale in a single-storey factory built on cheaper land in the outer suburbs. Third, their dependency on others is thought to reduce (Alperovich and Katz 1988; Carlton 1982; Fagg 1980). Formal testing of the simple incubation hypothesis needs information on firm formation data according to intra-urban location. Testing the dynamic hypothesis needs de-aggregation of firm data according to life-cycle stage, relocation over intra-urban locations (event-history), growth and firm failure rates. As a consequence of a scarcity of data on relocations of firms in a sufficient longitudinal period (1996–2000 is too short a period for reliable figures), our research will test only the simple hypothesis, embedded in the general framework of agglomeration hypotheses discussed in the previous section.

2.3. Agglomeration indicators: Background and definition

The relatively small size of the Netherlands provides a natural control for much location-specific heterogeneity. In fact, several variables enumerated in related studies (Henderson et al. 1995; Cortright and Mayer 2001; Glaeser 1999) – that are potentially important location-specific factors affecting either employment growth or establishment birth rates – are either roughly constant between locations in the Netherlands, 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, proportion of workers with particular skills, or percentage of workers with union membership (see Van Oort 2004 for actual testing of these elements). In principle cities are fertile grounds for testing knowledge-based theories of endogenous growth. In the Netherlands numerous dense urban agglomerations provide opportunities for learning because they are frequently centres 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, many firms may observe a marked cost saving from locating in urban areas rather than rural areas. Focusing on establishment births as an important element of urban employment dynamics (Ashcroft and Love 1996) sets this paper apart from the related literature on employment growth. It facilitates analysis in the sense that 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, being less affected by large 'sunk costs', that is to say, irrevocably committed costs or investments which are not recoverable in the case of exit or relocation (Cameron 1973; Clark and Wrigley 1995).

A Dutch municipal data set on sectoral employment structures is used to construct indicators of various types of agglomeration economies (as hypothesised earlier in this section) that are as reminiscent as possible 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 (multi-colinearity) and for theoretical reasons (agglomeration economies are commonly defined in a national, aggregated setting). As we want to test whether initial spatial circumstances are connected to new firm formation (a 'sources of growth' analysis, see Glaeser et al. 1992) explanatory variables are constructed using data from the base year (1996) to reduce problems of simultaneity. *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 in particular comprises (sector-specific) 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. This measure 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 (2003), who view it as a broader measure of local industrial organisation. For consistency reasons (optimal comparison with the influential Glaeser et al. 1992 and Henderson et al. 1995 papers) we apply the relative firm size definition of localised competition. 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, 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). A value of 0.5 results if employment is concentrated in only one sector. Lower values of *GINI* thus indicate higher degrees of diversity. The diversity indicator is treated as indicator of urbanisation (not sector specific) economies. Results presented in the next section can be used to make at least a suggestive test of the four sets of agglomeration hypotheses (Table 1)[2]. *EMPLOYMENT*

---

[2] Due to restrictions of space, correlation diagnostics of all explanatory variables used in this paper are not presented. No correlation higher tan 0.5 in absolute terms was permitted in the analyses.

1996 measures absolute employment values per municipality, and controls for localised start-of-period development bases. Spatial variations in wage structures (and development in wage structures) were not found significant in any analysis (in contrast to Glaeser et al. 1992), and for reasons of parsimonious presentation left out of the final analyses.

3. Data: ICT establishments in the Netherlands 1996–2000

The dataset of ICT firms in the Netherlands used in this paper has been created at Utrecht University (Atzema 2001). Although ICT as a current 'leading enabling technology' is used in almost all sectors of the Dutch economy, 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, uses the opposite in their statistics. In this research we apply the definition of Statistics Netherlands (CBS), which makes a distinction between ICT service industries and ICT hardware production industries. We decided to make a distinction into three main groups of ICT activities, aggregating 22 industries at the 5-digit level of definition (see Table 2). We are aware of the problem that sectors on the basis of SIC-codes, even at 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. Table 2 distinguishes ICT manufacturing (of hardware and software), ICT distribution and ICT service activities (see also Van Soest et al. 2002). Earlier research showed that high-tech *industry* in the

Table 2. Employment in the ICT sectors in the Netherlands (average 1996–2000)

Industry (SIC code)*	No. of jobs	% of jobs
Production		
Production of hardware (2233, 3002, 3220)	9,154	4,7
Production of software (7220, 72101)	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.

Netherlands is dominated by a few large corporations (Van Oort 2002). Trade and services, which make up some 70% of the research population, is not affected by this firm size determinism. These preliminary remarks are important, since this affects research outcomes, both theoretically and technically (through the spatial competition indicator). As will be explained in Sect. 4, formal testing will be undertaken at the aggregated sectoral scale as a result of insufficient spatial dependency in new firm formation and growth rates over individual industries or the three broad sectors 'production', 'trade' and 'services'. New firm formation in *the* ICT sector is therefore measured by the number of new establishments aggregated over all 22 industries.

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 were screened for the selection of 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: the Yellow Pages does not contain information on every company and has no information on existing employment levels. We therefore completed the dataset in a second step, in which the file obtained through the Yellow Pages was linked to the nationally covered LISA file (Van Oort 2004). This LISA file registers on an annual basis the employment of over 750,000 companies and institutions in the Netherlands. Both files have been compared with one another and the Yellow Pages file has been extended with other companies from the LISA-file. This results in a file of 18,985 ICT firms on average for the period 1996–2000. The number of jobs in ICT firms contributes nearly four per cent of the total employment in the Netherlands. The ICT providing sector is still a relatively small sector in the Netherlands. Furthermore, it becomes clear that employment in the Dutch ICT sector is dominated by service activities like consultancy, internet providing and wholesale trade. Within the field of production activities, the production of software dominates.

Several additional alterations to the data were carried out for this paper. Employment function, location quotients and concentration and specialisation indicators are calculated as average over the years 1966–2000. Growth indicators (defined in Van Oort 2004) compare the average stock of firms over 1996 and 1997 with the average stock of firms over 1999–2000 in order to minimise (spatial or temporal) outlier dependency. Growth rates are used *as reference* for new firm formation rates, since both indicators are complementary to each other in terms of regionalised employment growth (Ashcroft and Love 1996). Furthermore, the firm level data are aggregated into 580 locations that 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, resulting in 36 observations for Amsterdam, The Hague, Rotterdam and Utrecht. In 1998 The Netherlands consists of 548 municipalities, the four largest being replaced by the 36 3-digit zip code areas (still referred to as municipalities), making in total 580 observations. The longitudinal, firm level database allows a distinction to be made between new and incumbent firm populations.

4. Spatial dependence: Proximity, urbanisation regimes and visualisation

4.1. Spatial proximity and spatial heterogeneity (regimes)

Spatial proximity (clustering) is considered important by many observers for 'explaining' localised growth and new firm formation in high-tech sectors, the ICT-industry in particular (Audretsch and Feldman 1999; Baptista 2000). The marginal cost of transmitting tacit knowledge rises with distance. As tacit knowledge and human interaction become more valuable in the innovation process, geographical proximity becomes crucial to the innovation and growth process. The exchange of tacit knowledge may require a high degree of mutual trust and understanding. It is especially in the contiguous proximity way of thinking about space that the two contradicting outcomes concerning agglomeration economies (urbanisation *versus* localisation) discussed in Sect. 2 come to the fore. Most of the relevant empirical literature focuses on American states as the spatial unit of analysis. Some Anglo-Saxon research, however, focuses on lower scales of analysis. Anselin et al. (2000) and Wallsten (2001), for instance, 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. 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).

The (geographic) literature also provides clues for non-contiguous (regime) types of spatial dependence. Quality of life aspects, regional labour markets, specialised networks and city size appear as significant locational considerations, both to professional workers and to growing ICT firms (Van Oort et al. 2003; Atzema 2001). The spatial structures of proximity (contiguous nearness at the municipal level) and heterogeneity (urban hierarchical and regional, not necessarily contiguous, spatial dependence) have been captured in this paper 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 new firm formation) 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 1988). Spatial heterogeneity on the other hand is modelled by spatial regimes, involving change-of-slope regression estimation over various types of locations that theoretically 'perform' differently. Four sets of spatial regimes are distinguished, each indicating aspects of urban structures at different spatial scales. On the meso-level we distinguish a labour market induced connectedness regime from a non-connectedness regime (Fig. 1). This spatial regime concerns commuting based labour market relations. In the figure, core and suburban municipalities together comprise the connected regime, as opposed to the other types of locations that are characterised as non-connected. The four types of locations have been distinguished, initially based on municipal data for 1990–1999. The classification is based on the dependency of a municipality's population upon employment and services proximity and accessibility. Urban core areas have an important employment function. More than 15,000 persons commute into these municipalities (while living somewhere else) on a daily basis. Municipalities where more than 20% of residents commute to central core locations are labelled suburban. The literature finds in general that urban areas in the

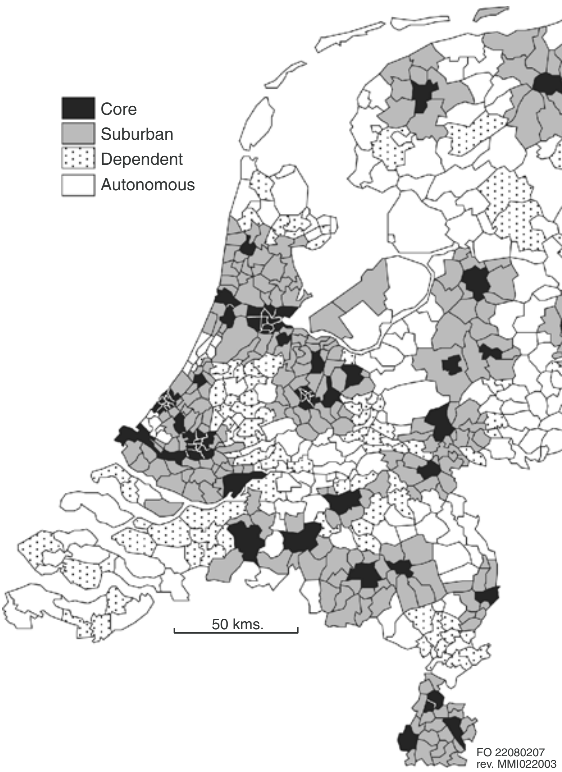

Fig. 1. The (labour market) connectedness spatial regimes

connected regime show higher economic growth and innovation rates than areas in the non-connected regime (e.g., Anselin et al. 2000). As becomes clear from Fig. 1, locations in the connected regime are not necessarily adjacent to each other. On the macro-level, three national zoning regimes have been distinguished: the Randstad core region, the so-called intermediate zone and the national periphery (Fig. 2) Distinguishing between macro-economic zones in the Netherlands is based on a gravity model of total employment concerning data from 1997. The Randstad region in the Netherlands historically comprises the economic core provinces of Noord-Holland, Zuid-Holland and Utrecht, the intermediate zone mainly comprises the growth regions of Gelderland and Noord-Brabant, while the national periphery is built up by the northern and southern regions of the country. This zoning distinction is hypothesised as important in many studies on endogenous growth (e.g., Van Oort 2004) and the incubator hypothesis (e.g., Brouwer et al. 1999) in the Netherlands, in the sense that the Randstad region traditionally is thought to have better economic potential for development. Within the Randstad core economic region, a division in north wing (Amsterdam-Utrecht) and south wing (The Hague-Rotterdam) is often made, especially concerning the location of ICT firms (Atzema 2001). A third spatial regime is therefore applied in the econometric analyses, distinguishing the north and south wing municipalities from other locations in the Netherlands (Fig. 3). The fourth set of

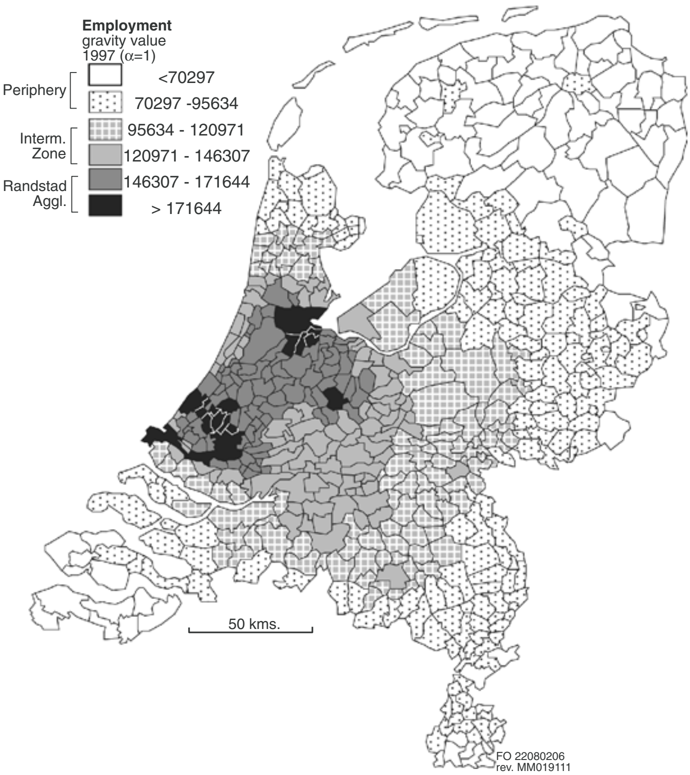

Fig. 2. National zoning spatial regime

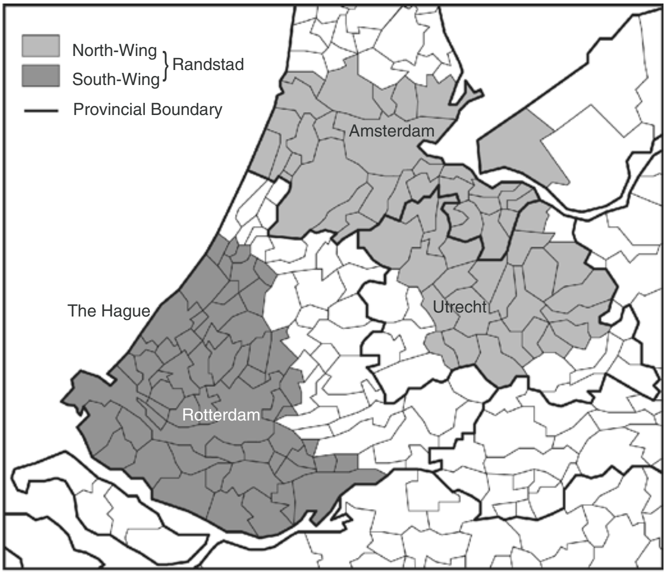

Fig. 3. North and south wing of the Randstad spatial regimes

spatial regimes is constructed using the degree of urbanisation of municipalities. A cut-off population threshold of 45,000 inhabitants, standard for distinguishing medium-sized cities in the Netherlands, is used for the distinction in urban and non-urban regimes respectively. These forms of spatial heterogeneity constitute four spatial levels of urban constellation: the urban level itself, the functional (commuting) region, the meso-level 'agglomerative fields' of the north wing and south wing of the Randstad core region, and finally the macro-economic core-periphery (Randstad, intermediate zone, national periphery) distinction.

4.2. Descriptive analysis: Global indicators of proximity patterns

A technique most conveniently used for spatial statistical descriptions of data is exploratory spatial data analysis (ESDA). ESDA 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 1988). These methods provide measures of local spatial autocorrelation, the technical aspects of which we now briefly discuss. In this section 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.

Spatial autocorrelation can be defined as the coincidence of value similarity with locational similarity. 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 neighbors with very dissimilar values. The measurement of global spatial autocorrelation in this section 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:

$$I_t = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij}(x_{it} - \mu_t)(x_{jt} - \mu_t)}{\sum_i (x_{it} - \mu_t)^2} \qquad (1)$$

where x_{it} is the observation in region i and year (period) t, μ_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. Finally S_0 is a scaling factor equal to the sum of all elements of W. For row-standardized spatial weight matrices, which are the preferred way to implement the Moran's I test statistic, the normalizing factor S_0 equals n, since each row then sums to 1 (see Anselin 1995). The statistic of Eq. (1) then simplifies to the ratio of a spatial cross products to variance, making Moran's I similar but not equivalent to a correlation coefficient; it is not centered 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

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

	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*
Employment 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 $p = 0.01$.

depends on the stochastic assumptions made. Either the assumption of a normal distribution of variables in question (normality assumption), or the assumption that each value observed could equally likely have occurred at all locations (randomization assumption), or a randomization approach using a reference distribution for *I* that is generated empirically (permutation assumption) can be tested. While all three variance assumptions were tested in our research, only the results for the randomization assumption are presented since all three models gave very similar results. Inference is based on a standardized z-value of *I* that is calculated 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 might result from pure chance. In contrast, a negative and significant z-value for Moran's *I* indicates negative spatial autocorrelation, the opposite of spatial clustering[3].

Table 3 displays Moran's *I* statistics for spatial autocorrelation for: (1) location quotients of (aggregated) employment in ICT firms over the period

[3] 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 checker board pattern.

1996–2000 (employment function); (2) new firm formation rates (average 1996–2000; new firms relative to average potential working population of 15 till 65 years of age; variable explained in section 5)[4]; (3) employment growth in all ICT firms for the period 1996-2000; and (4) employment growth in incumbent firms only (present in all years of observation in the dataset). 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. In Table 3, first (w_1), second (w_2) as well as third order (w_3) full distance weight matrices are used for spatial autoregressive modelling. In all research populations presented in Table 3 first order distance weights show the highest significance in spatial autocorrelation. In following sections emphasis will therefore be placed only on first order weight matrices. From the table it becomes clear that the new ICT firm formation rates show high degrees of spatial association. Growth functions are in general less spatially clustered than the employment function and the firm formation rates. Employment growth in incumbent firms in ICT distribution activities and ICT services is not significant.

4.3. Descriptive analysis: Visualisation of local proximity patterns

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, interesting questions that remain include: which regions or locations contribute most to the global spatial autocorrelation; 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). Analysis of *local* spatial autocorrelation can be carried out using the tool of the Moran scatterplot, which can be used to visualise local spatial instability. The spatial lag (using the first order distance matrix w_1 for spatial proximity) is plotted against the original values, 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 comprises locations with a high value surrounded by locations with high values. LH indicates locations with a low value surrounded by locations with high values, while LL denotes 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.

Localised Moran scatterplot maps are presented for location quotients of (aggregated) employment in ICT firms over the period 1996–2000

[4] This combination reflects a common definition of new firm formation (see Ashcroft and Love 1996; Armington and Acs 2002). Alternative definitions however relate new firm accounts to the existing stock of firms (e.g., Storey 1982; Dumais et al. 2002). The sensitivity of the (spatial) outcomes of our research questions for alterations in these definitions turned out to be limited.

(employment function, Fig. 4), for new firm formation rates (average 1996–2000, Fig. 5) for employment growth in incumbent firms (present in all years of observation in the dataset, Fig. 6) and for employment growth in all ICT-firms (new and incumbent) for the period 1996–2000 (Fig. 7). Notice that the relatively large share of service activities among ICT firms (Table 2) determines 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 Fig. 4) 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 relatively high concentration patterns. The location of newly established ICT firms over the research period (Fig. 5) differs from the employment function considerably: new ICT firms are not present in the south wing of the Randstad region (The Hague-Rotterdam). Instead, the north wing of the Randstad region and municipalities more remote from the Randstad core region show high concentration values. New firm formation automatically induces employment growth (as long as these new firms survive, see Ashcroft and Love 1996). Employment growth in incumbent firms (Fig. 6), though, is clustered in the traditional high-technology centres in the Netherlands (compare Wever and Stam 1999). The

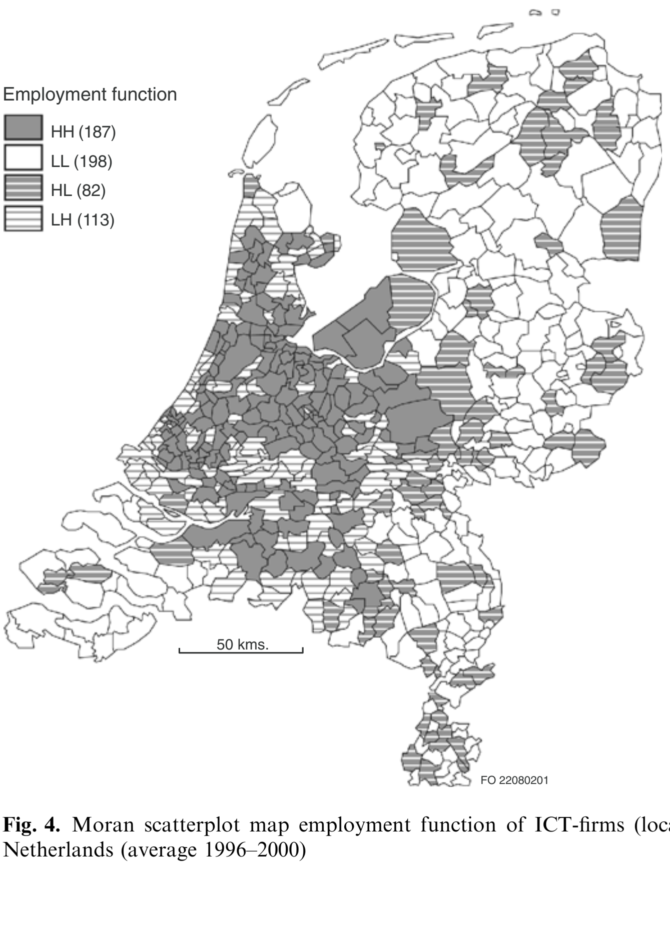

Fig. 4. Moran scatterplot map employment function of ICT-firms (location quotients) in the Netherlands (average 1996–2000)

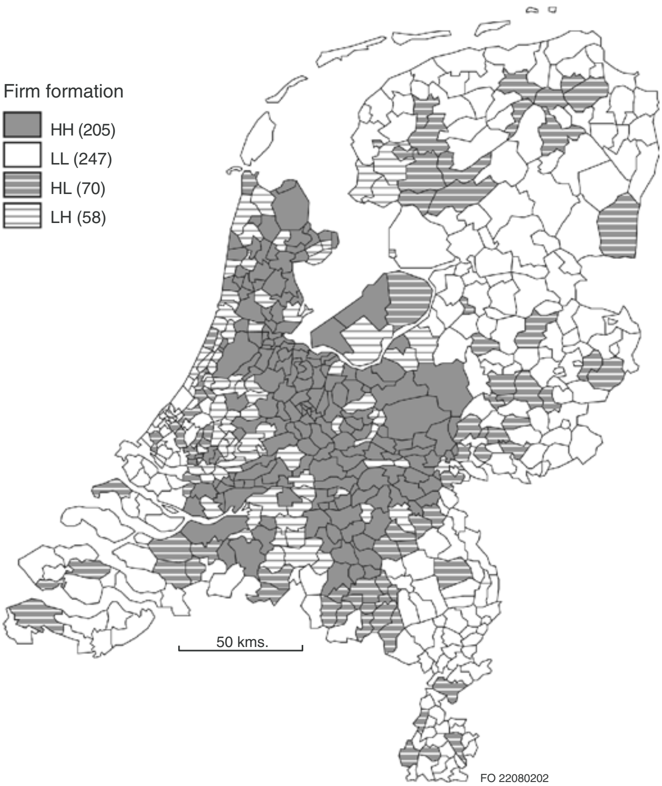

Fig. 5. Moran scatterplot map log new firm formation rate in ICT-industries in the Netherlands (average 1996–2000)

similarity with Fig. 4 (spatial *concentrations* of employment) is striking. Figures 6 (growth of incumbent firms) and 5 (new firm formation) together make up the total growth figures presented in Fig. 7. Remarkably, new firm formation appears to attribute much more to this growth pattern than growth among incumbents. 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 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 service-based activities.

5. Descriptive analyses using size, zoning and connectedness regimes

In this section we take a closer look at the characteristics of the different concepts and degrees of urbanisation presented as spatial regimes in the previous section. We are especially interested in patterns of new firm formation in urban areas that could hint at confirmation of the incubation hypothesis. At first instance, no significant correlation is found between urban population size

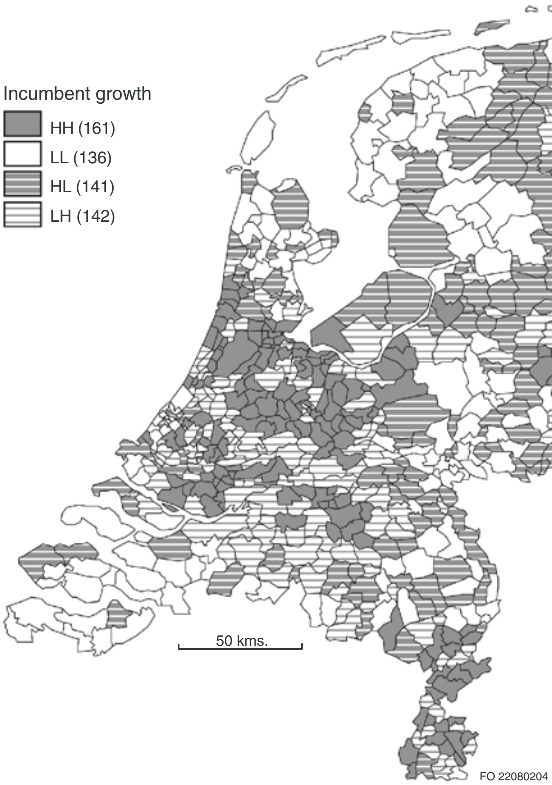

Fig. 6. Moran scatterplot map log employment growth *incumbent* ICT-firms in the Netherlands (1996–2000)

and new firm formation in municipalities (n = 580). Areas with large population numbers do not necessarily have the highest scores with regards to new firm formation. In order to test whether *employment* density is positively related to firm formation and employment growth in ICT firms, several cross-tabulations are presented in Tables 4 and 5. In Table 4 the zoning-regimes (Randstad core region, intermediate zone and national periphery) are crossed with the connectedness classification (core location, suburban location and non-connected location). In Table 5 the zoning regime is crossed with an urban size regime (large cities: more than 200,000 inhabitants, medium-sized cities: between 45,000 and 200,000 inhabitants, small cities: less than 45,000 inhabitants)[5]. The tables provide information on total employment in the ICT sector, local over-representation in terms of sectoral structure (location quotient), growth rates for all and incumbent firms and new firm formation rates. From Tables 4 and 5 some remarkable insights come to the fore. The ICT sector is

[5] Recall that the urban size *spatial regime* is defined as urban (more than 45,000 inhabitants) versus non-urban (less than 45,000 inhabitants). In Sect. 5 we make use of a more detailed classification than in the econometric models in Sect. 6, mainly for reasons of model comparison with previous research.

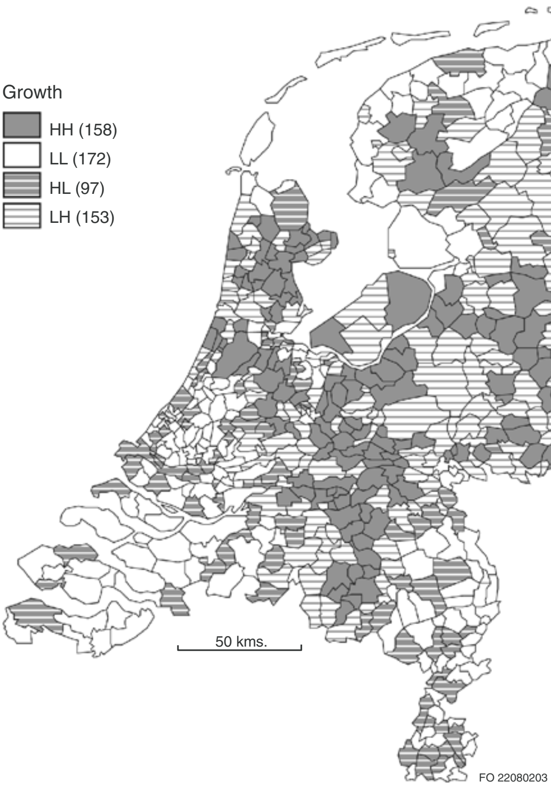

Fig. 7. Moran scatterplot map log employment growth all ICT-firms in the Netherlands (1996–2000)

over-represented (measured as a location quotient related to total employment) in core and large locations, especially in the Randstad region. But also medium sized cities and small cities in the Randstad show a considerable over-representation of ICT firms. Within the Randstad, a more polycentric locational pattern exists. Growth rates of ICT employment are highest in suburban, non-connected, medium-sized and small locations in the intermediate zone and national periphery in the period 1996-2000. Less central locations on all defined urban levels are thus catching up in terms of growth rates with more central locations. Growth rates of incumbent ICT firms (presented over the whole period 1996–2000) show the same localised pattern, but to a much lesser degree. Growth in the ICT sector as a whole is approximately 40%, while among incumbent firms the figure is only 20%. New firm formation rates are highest in all Randstad locations (except non-connected ones), but also in medium sized cities and suburban locations in the Intermediate zone. These descriptive analyses lead us to argue that the ('simple') incubation hypothesis, stating that only larger cities are breeding places for new firms and entrepreneurship, needs to be adjusted to the appropriate spatial levels and units of analysis: that of the agglomerated region. In this way, central core locations, suburban locations, large and medium sized cities in the Randstad and intermediate zone of the Netherlands together make up a complex and detailed

Table 4. Employment, growth and new firm formation (1996–2000) in zoning and connectedness classifications

Connected ↓	Zoning →	Randstad core region	Intermediate zone	National periphery	Total
Core location	Employment ICT	69,086	27,388	13,590	110,064
	Lq ICT employm.*	1.622	1.176	0.952	1.373
	% growth ICT firms	+34.7%	+40.1%	+51.4%	+38.0%
	Index growth ICT	0.856	0.989	1.269	0.937
	% growth inc.**	+17.7%	+19.9%	+25.1%	+19.2%
	Index growth inc.	0.796	0.891	1.126	0.860
	New firm form.***	5.844	4.218	2.659	4.821
	Index firm form.	1.591	1.148	0.724	1.312
Suburban	Employment ICT	44,120	11,342	1,865	57,327
	Lq ICT employm.	1.676	0.605	0.202	1.055
	% growth ICT firms	+36.8%	+55.4%	+51.1%	+40.8%
	Index growth ICT	0.909	1.369	1.262	1.077
	% growth incumb.	+30.4%	+22.1%	+25.6%	+28.4%
	Index growth inc.	1.362	0.991	1.148	1.274
	New firm form.	5.226	4.080	1.131	3.930
	Index firm form.	1.422	1.111	0.308	1.070
Non-connected	Employment ICT	9,189	9,455	7,665	26,309
	Lq ICT employm.	1.138	0.500	0.238	0.444
	% growth ICT firms	+28.8%	+68.4%	+59.4%	+50.9%
	Index growth ICT	0.712	1.688	1.466	1.256
	% growth incumb.	+23.8%	+32.8%	+17.9%	+25.2%
	Index growth inc.	1.067	1.470	0.805	1.129
	New firm form.	2.465	2.960	1.882	2.298
	Index firm form.	0.671	0.806	0.512	0.625
Total	Employment ICT	122,395	48,185	23,120	193,700
	Lq ICT employm.	1.590	0.791	0.415	1.000
	% growth ICT firms	+35.0%	+48.7%	+54.0%	+40.5%
	Index growth ICT	0.864	1.202	1.333	1.000
	% growth incumb.	+22.1%	+22.6%	+22.7%	+22.3%
	Index growth inc.	0.989	1.012	1.020	1.000
	New firm form.	5.186	3.758	1.860	3.674
	Index firm form.	1.412	1.023	0.506	1.000

* Lq = Location quotient.
** Inc. = Incumbent ICT-firms (1996–2001).
*** New firm formation per 10,0000 population.

polycentric agglomerative field having more than average propensities for ICT-firms to start new businesses. At the same time we should be aware of the fact that within this agglomerative field, 'unconnected' and small municipalities clearly have more limited propensities for new firm formation in the ICT sector. The next section focuses on these patterns in more detail, relating indicators of agglomeration economies to the firm formation patterns.

6. Empirical results

In Tables 6 and 7, the econometric models that we ran are summarised. Below the tables technical explanation on the models is provided. The models are numbered over the two tables – models (1) to (4) in Table 6 and models

Table 5. Employment, growth and new firm formation (1996–2000) in zoning and urban size classifications

Urban size ↓	Zoning →	Randstad core region	Intermediate zone	National periphery	Total
Large (>200,000)	Employment ICT	52,162	x****	x****	52,162
	Lq ICT employm.*	1.656	x	x	1.656
	% growth ICT firms	+38.2%	x	x	+38.2%
	Index growth ICT	0.944	x	x	0.944
	% growth inc.**	+16.9%	x	x	+16.9%
	Index growth inc.	0.759	x	x	0.759
	New firm form.***	5.117	x	x	5.117
	Index firm form.	1.393	x	x	1.393
Medium (45,000–200,000)	Employment ICT	42,463	35,212	14,855	92,530
	Lq ICT employm.	1.747	1.055	0.820	1.221
	% growth ICT firms	+30.9	+43.7%	+51.1%	+38.8%
	Index growth ICT	0.762	1.080	1.261	0.958
	% growth incumb.	+19.5%	+20.8%	+25.3%	+21.0%
	Index growth inc.	0.874	0.931	1.133	0.940
	New firm form.	5.717	4.388	2.521	4.384
	Index firm form.	1.556	1.194	0.686	1.193
Small (<45,000)	Employment ICT	27,770	12,973	8,265	49,008
	Lq ICT employm.	1.311	0.471	0.220	0.567
	% growth ICT firms	+35.3%	+63.1%	+59.3%	+46.2%
	Index growth ICT	0.872	1.559	1.465	1.140
	% growth incumb.	+39.1%	+28.8%	+18.2%	+32.2%
	Index growth inc.	1.755	1.291	0.815	1.442
	New firm form.	4.770	3.221	1.634	2.875
	Index firm form.	1.298	0.877	0.445	0.782
Total	Employment ICT	122,395	48,185	23,120	193,70
	Lq ICT employm.	1.590	0.791	0.415	1.000
	% growth ICT firms	+35.0%	+48.7%	+54.0%	+40.5%
	Index growth ICT	0.864	1.202	1.333	1.000
	% growth incumb.	+22.1%	+22.6%	+22.7%	+22.3%
	Index growth inc.	0.989	1.012	1.020	1.000
	New firm form.	5.186	3.758	1.860	3.674
	Index firm form.	1.412	1.023	0.506	1.000

* Lq = Location quotient.
** Inc. = Incumbent ICT-firms (1996–2001).
*** New firm formation per 10,0000 population.
**** Combination size & zoning non-existent.

(5) and (6) in Table 7. The degree of localised concentration, diversity and competition are introduced according to the definitions given in Sect. 2. Besides concentration indices of ICT firms, concentration indices for industrial, distribution, business service and consumer service activities are introduced in the model (see Van Oort 2004 for an exact definition of these activities). Likewise, localised competition, in line with the Glaeser et al. (1992) approach, is measured by relative firm size both for ICT firms and for all firms in the localised economy in an aggregated sense.

The Ordinary Least Squares model for the percentage of new firm formation (column (1) in Table 6) shows the significance of both concentration indicators (of the ‘own’ ICT sector, as well as in general for business services

Table 6. OLS and combined spatial lag and spatial regime models for 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	Rand-stad	Int. Zone	Periphery
Constant	0.536	−0.375	−0.120	−0.092	−0.377	0.231	−0.195
	(0.933)	(−0.726)	(−0.208)	(−1.455)	(−0.484)	(0.092)	(−0.700)
Concentration	0.789	0.687	0.654	0.056	0.616	−0.062	0.056
ICT Firms	**(7.867)**	**(7.614)**	**(7.361)**	(0.601)	**(6.384)***	(−0.086)*	(0.707) *
Concentration	−0.029	−0.022	−0.008	−0.092	−0.041	0.016	−0.062
Industry	(−1.239)	(−1.008)	(−0.347)	(−1.455)	(−1.199)	(0.403)	(−1.158)
Concentration	−0.119	−0.199	−0.142	−0.660	−0.189	−0.030	−0.133
Distribution	(−1.144)	**(−2.665)**	(−1.771)	**(−3.544)**	(−1.585)	(−0.204)	(−0.952)
Concentration	0.292	0.188	0.239	−0.314	0.257	0.392	0.062
Business services	**(4.394)**	**(3.153)**	**(3.854)***	(−1.640)*	**(2.796)***	**(3.382)***	(0.485)*
Concentration	−0.234	−0.238	−0.059	−1.106	−0.072	0.182	−0.418
Consumer serv.	**(−2.244)**	(−2.539)	(−0.590)*	**(−4.887)***	(−0.505)*	(0.089)*	**(−2.575)***
Lack of	−1.114	−0.559	−0.820	−0.575	−1.133	−1.503	−0.684
Diversity	**(−2.934)**	(−1.639)	(−2.167)	(−0.676)	**(−2.011)**	**(−1.940)**	(−1.116)
Size ICT Firms	1.029	0.815	0.793	0.091	0.731	0.744	1.106
(competition)	**(19.820)**	**(17.429)**	**(15.985)**	**(7.133)**	**(10.192)**	**(7.268)**	**(11.967)**
Size all firms	−0.465	−0.352	0.011	−0.357	−0.404	−0.280	−0.406
(Competition)	**(−6.353)**	**(−5.338)**	(0.052)*	**(−5.108)***	**(−3.907)**	**(−1.943)**	**(−3.358)**
Employment 1996	0.148	0.073	0.084	−0.052	0.084	0.068	−0.055
ICT Firms	(1.534)	(0.847)	(1.002)	(−0.590)	(0.917)	(0.096)	(−0.694)
Employment 1996	−0.073	0.007	−0.038	0.056	−0.016	−0.068	0.556
All firms	(−0.678)	(0.008)	(−0.383)	(0.595)	(−0.130)	(−0.094)	(0.595)
Spatial	–	0.969	0.973			0.974	
Coefficient. ρ		**(26.843)**	**(22.948)**			**(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	6.840	1.527			2.981	
	(0.393)	(0.077)	(0.216)			(0.170)	
LM (ρ)	324.78	–	–			–	
	(0.000)						
LM (λ)	56.61	–	–			–	
	(0.000)						
LR (ρ)	–	106.49	114.54			106.86	
		(0.000)	(0.000)			(0.000)	
Chow-Wald	–	–	42.822			33.810	
			(0.000)			(0.051)	

Values of log-likelihood are not comparable over populations of all and old establishments. Following Anselin et al. (1995), LM (ρ) and LM (λ) are statistics for the presence of a spatial lag in the dependent variable and in the residual respectively,with a critical value of 3.84 at the 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 in bold. The spatial Chow-Wald test is distributed as an F variate and tests for structural instability of the regression coefficients over regimes (Anselin 1995, p. 32). 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 Sect. 4).

Table 7. Combined spatial lag and spatial regime models for 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	North wing	South wing	Other
Constant	−0.357	0.198	−0.393	0.376	−0.277
	(−0.419)	(0.254)	(−0.666)	(0.096)	(−0.066)
Concentration	0.683	0.626	0.639	−0.472	0.082
ICT firms	**(3.629)**	**(5.968)**	**(7.188)**	(−0.093)	(0.068)
Concentration	−0.004	−0.056	−0.026	−0.070	0.040
Industry	(−0.014)	(−1.765)	(−1.104)	(−1.228)	(0.600)
Concentration	−0.280	−0.119	−0.156	−0.600	0.189
Distribution	**(−2.848)**	(−1.068)	(−1.772)*	**(−3.068)** *	(0.825) *
Concentration	0.237	−0.038	0.225	−0.439	0.580
Business services	**(2.821)** *	(−0.433) *	**(3.290)** *	(−2.153) *	**(2.904)** *
Concentration	−0.383	−0.127	−0.078	−01.008	−0.040
Consumer services	**(−3.086)**	(−0.864)	(−0.698) *	**(−4.882)** *	(−0.144) *
Lack	−1.005	−0.132	−0.928	−1.046	−1.646
of Diversity	**(−2.253)** *	(−0.241) *	**(−2.140)**	(−1.246)	(−1.401)
Size ICT firms	0.909	0.757	0.755	0.886	0.952
(Competition)	**(13.637)**	**(11.893)**	**(14.247)**	**(5.815)**	**(5.011)**
Size all firms	−0.215	−0.509	−0.382	0.002	−0.309
(Competition)	**(−2.487)**	**(−5.035)**	**(−4.940)**	(0.011)	(−1.306)
Employment 1996	0.209	0.035	0.077	0.112	−0.746
ICT firms	(1.081)	(−0.368)	(0.915)	(0.100)	(−0.062)
Employment 1996 all	−0.097	−0.026	−0.005	−0.111	0.076
Firms	(−0.488)	(−0.214)	(−0.006)	(−0.098)	(0.063)
Spatial	0.964			0.967	
Coefficient. ρ	**(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)	

See technical explanation below Table 6.

in a positive sense, and for consumer services in a negative sense) and the diversity indicator. The third agglomeration indicator, measuring 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 within a municipality into account and independent of sectoral composition. Interestingly, these results do not provide unambiguous support for any of the three endogenous development theories discussed in Sect. 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, but 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 on the MAR hypothesis of economic dynamics – a very confusing picture, indeed. Yet, results presented are still very much 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, and rich in industrial diversity (compare also Van Soest et al. 2002). The test statistics of LM(ρ) and LM(λ) in Column (1) reveal the presence of spatial autocorelation dependency of the model (also the case in 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 indeed turns out to be highly significant. Introducing spatial dependency in the model alters the coefficients slightly when compared to the OLS base model. Relative specialisation of distribution activities in particular 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 does not emerge as a 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 (3a–b) and (4a–c) 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 with new firm formation in urban municipalities, as opposed to non-urban ones. The significance of industrial diversity makes the result, again, theoretically 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 profoundly in urban environments. This conclusion confirms the urban setting of the endogenous development theories as outlined in Sect. 3. But other definitions of urbanisation appear to be significant as well for ICT business development. Column (4) in Table 6 shows that the Randstad region most notably '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 (5a–b) shows the significance of the connected spatial regime, as opposed to the unconnected regime. Column (6a–c) in Table 7 finally shows that *within* the Randstad region, especially the north wing (Amsterdam-Utrecht) is characterised by significant agglomeration indicators. The south wing of the Randstad (Rotterdam-The Hague) shows quite opposite, less favourable agglomeration circumstances than the north wing, particularly concerning industrial diversity and the specialisation in business services. 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 considerably the current debate on urbanisation and localisation externalities, which focuses mainly on proximity based spillovers and knowledge transfer.

7. Conclusions

This paper has empirically investigated spatial determinants of new firm formation in the ICT sector in the Netherlands. We contribute to the growth and dynamic discussion on ICT clusters by addressing and answering three questions. The first question asked whether measurable agglomeration factors that are comparable with earlier studies are connected to new firm formation in the ICT industry. Secondly, we asked what conceptual spatial configurations best describe the new firm formation patterns. More specifically, we tested the ('simple') incubation hypothesis that states that larger cities are breeding grounds for new firms and entrepreneurship because of localised knowledge spillovers. And third, we tried to evaluate whether entrepreneurship as measured by firm dynamics contributes to growth theory. In this section we answer these questions according to the research outcomes presented in the previous sections and draw some conclusions for spatial policy. The empirical investigation made use of a unique and highly detailed (longitudinal) data set of births of new establishments in these sectors in each of 580 municipalities. The relatively small size of the Netherlands offers control for certain types of unobserved heterogeneity, such as aspects of labour market conditions that have plagued earlier studies.

Regarding the first question of this paper ('what agglomeration indicators should be tested for, and how do they perform?'), results from the analysis suggest that new establishments in the ICT sector tend to be concentrated in urban areas that are already relatively specialized in this sector and that are relatively rich in the presence of other industries. These results support previous evidence that high technology enterprises tend to co-locate in areas where economic activity is spatially dense. But these outcomes do not fully support or contradict four theories of knowledge spillovers, attributed to Marshall-Arrow-Romer, Porter, Jacobs and combinations of these three that frequently have been tested using data from urbanized areas. Yet, it does provide some insights into the types of areas where ICT establishments choose to locate. It also has implications for practical conclusions that might be drawn by policy makers regarding urban planning and development. This leaves room for policy makers to act and in turn lay the foundations for and develop a favourable local environment for growth and innovation (Gibbs and Tanner 1995). The models explored in this article confirm that 'urbanisation economies' may offer advantages of flexibility secured in a local diversity of activities. But also provision of ICT infrastructure in urban areas for the specialised ICT sector might be a fruitful policy. Our research however shows an important constraint to this general view. Spatial policy, then, should not be restricted to local *or* regional environments alone. Spatial externalities relevant in local contexts can be at work at higher levels. In order to 'gain' from them, location nearby larger cities or within the regional or spatial hierarchical setting (of labour market 'connectedness', urban size or national zoning heterogeneity) then suffices (compare Phelps et al. 2001). This also means that certain local endowments in larger regional settings can be critically poor for the development of (new) ICT clusters, such as the south wing of the Randstad and the national periphery. In those regions local policy makers should be open to the argument of spillover effects from nearby (not necessarily adjacent) agglomerations instead of promoting 'own' ICT clusters.

The elements for answering the second research question ('what spatial configuration describes the firm formation pattern best and does this confirm the incubator hypothesis of new firm formation?') come from descriptive and econometric analyses. ESDA analyses in Chapt. 4 and the significance of spatial lag estimators in the models presented in Sect. 6 indicate that proximity and contiguity based spatial autocorrelation is significantly attached to new firm formation rates in the Dutch ICT sector over municipalities. At the same time, the descriptive analyses of Sect. 5 and the change-of-slope econometric analyses of Sect. 6 show that urbanisation defined in spatial heterogeneous (network-based, non-contiguous) regimes matters for new ICT firm formation on different scales of analyses in the Netherlands. In general, this leads to the major point of our argument that controversial research results in the literature concerning explanatory spatial circumstances that most favorably induce dynamic and innovative externalities (to a large extent) can be attributed to the lack of consistent spatial research designs that allow the modelling of multiple spatial scale and composition effects. More specifically the analyses lead us to argue that the ('simple') incubation hypothesis, stating that only larger cities are breeding grounds for new firms and entrepreneurship, needs to be adjusted to the appropriate spatial levels and units of analysis: that of the agglomerated region. In this way, central core locations, suburban locations, large and medium sized cities in the Randstad and intermediate zone of the Netherlands together constitute a complex and detailed polycentric agglomerative field having more than average propensities for ICT firms to start new businesses. Our findings and conclusions considerably extend the current debate on urbanisation and localisation externalities, which hitherto focuses mainly on proximity based knowledge transfer.

Our results, however, should be treated with caution because, to date, most studies of location determinants have focused on employment growth and innovation intensity; relatively few have looked at the component of employment growth arising from establishment births (Dumais et al. 2002, Acs et al. 2003, Van Oort 2004). The relative importance of various types of externalities in fostering new firm formation, economic growth and innovation, locally as well as among more geographically dispersed areas, has implications for the formulation and interpretation of endogenous growth models. This brings us to the third research question ('do concepts of firm life-cycles potentially contribute to the spatial externalities debate?'). Our answer to this question is confirmative. No satisfying formulation has been developed (yet) to incorporate life-cycle aspects (of entrepreneurship) fully in endogenous growth models (Acs et al. 2003). But there clearly is a rationale to do so: based on our research, we argue that the lack of consistent inclusion of life-cycle aspects of firms (like the distinction between new firms and incumbent firms) in the present mainstream literature on dynamic externalities contributes to unnecessary controversies in research outcomes, as does an undifferentiated conceptualisation of urban space.

References

Acs Z (2002) Innovation and the growth of cities. Edgar Elgar, Cheltenham

Acs Z, Audretsch D, Braunerhjelm P, Carlsson B (2003) The missing link: the knowledge filter and endogenous growth. Paper presented at the workshop "Modern Entrepreneurship, Regional Development and Policy". Tinbergen Institute, Amsterdam

Alperovich G, Katz E (1988) The location decision and employment suburbanzation. Urban Studies 25: 243–247

Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dordrecht

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

Anselin L, Varga A, Acs ZJ (2000) Geographic and sectoral characteristics of academic knowledge externalities. Papers in Regional Science 79: 435–443

Armington C, Acs ZJ (2002) The determinants of regional variation in new firm formation. Regional Studies 36: 33–45

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

Ashcroft B, Love JH (1996) Firm births and employment change in the British counties: 1981–89. Papers in Regional Science 75: 483–500

Atzema O (2001) Location and local networks of ICT firms in the Netherlands. TESG Journal of Economic and Social Geography 92(3): 369–378

Audretsch DB, Feldman MP (1996) R & D spillovers and the geography of innovation and production. The American Economic Review 86: 630–640

Baptista R (2000) Do innovations diffuse faster within geographical clusters? International Journal of Industrial Organisation 18: 515–535

Beardsell M, Henderson V (1999) Spatial evolution of the computer industry in the USA. European Economic Review 43: 431–456

Black D, Henderson V (1999) Spatial evolution of population and industry in the United States. AEA Papers and Proceedings 89: 321–327

Brouwer E, Budil-Nadvornikova H, Kleinknecht A (1999) Are urban agglomerations a better breeding place for product innovation? An analysis of new product announcements. Regional Studies 33: 541–549

Cameron GC (1973) Intraurban location and the new plant. Papers of the Regional Science Association 31: 125–143

Capello R (2002) Entrepreneurship and spatial externalities: theory and measurement. The Annals of Regional Science 36: 387–402

Carlton DW (1982) The location and employment choices of new firms: an econometric model with discrete and continuous endogenous variables. The Review of Economics and Statistics 27: 440–449

Clark GL, Wrigley N (1995) Sunk costs: a framework for economic geography. Transactions of the Institute of British Geographers 20: 204–223

Combes PP (2000) Economic structure and local growth: France 1984–1993. Journal of Urban Economics 47: 329–355

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

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

Fagg JJ (1980) A re-examination of the incubator hypothesis: a case of Greater Leicester. Urban Studies 17: 35–44

Feldman MP, Audretsch DB (1999) Innovation in cities: science based diversity, specialization and localized competition. European Economic Review 43: 409–429

Feser EJ (2002) Tracing the sources of local external economies. Urban Studies 39: 2485–506

Gibbs D, Tanner K (1995) Communications technology, local economies and regulation theory. TESG Journal of Economic and Social Geography 88: 29–49

Glaeser EL (1999) Learning in cities. Journal of Urban Economics 46: 254–277
Glaeser EL, Kallal HD, Scheinkman JA, Schleifer A (1992) Growth in cities. Journal of Political Economy 100: 1126–152
Henderson V, Kuncoro A, Turner M (1995) Industrial development in cities. Journal of Political Economy 103: 1067–1085
Hoover EM, Vernon R (1959) Anatomy of a metropolis. Harvard University Press, Cambridge Mass
Jacobs J (1969) The economy of cities. Vintage, New York
Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics 36: 577–598
Leone RA, Struijck R (1976) The incubator hypothesis: evidence from five SMSA's. Urban Studies 13: 325–331
Lucas RE (1993) Making a miracle. Econometrica 61: 251–272
Marshall A (1890) Principles of economics. Prometheus Books, New York
Van Oort FG (2002) Innovation and agglomeration economies in the Netherlands. TESG Journal of Economic and Social Geography 93: 344–360
Van Oort FG (2004) Urban growth and innovation. Analysis of spatially bounded externalities in the Netherlands. Ashgate, Aldershot
Van Oort FG, Weterings A, Verlinde H (2003) Residential amenities of knowledge workers and the location of ICT-firms in the Netherlands. TESG Journal of Economic and Social Geography 94: 516–523
Parr JB (2002) Agglomeration economies: ambiguities and confusions. Environment and Planning A 34: 717–731
Phelps NA, Fallon RJ , Williams CL (2001) Small firms, borrowed size and the urban-rural shift. Regional Studies 35: 613–624
Porter M (1990) The competitive advantage of nations. Free Press, New York
Romer PM (1986), Increasing returns and long-run growth. Journal of Political Economy 94: 1002-1037
Rosenthal SS, Strange WC (2001) The determinants of agglomeration. Journal of Urban Economics 59: 191–229
Rosenthal SS, Strange WC (2003) Geography, industrial organization and agglomeration. Review of Economics and Statistics (forthcoming)
Van Soest DP, Gerking S, van Oort FG (2003) Knowledge transfer and the location of new ICT establishments in the Dutch province of Zuid-Holland. International Journal of Entrepreneurship and Small Business (forthcoming).
Stam E (2003) Why butterflies don't leave Locational evolution of evolving enterprises. Dissertation, Utrecht University
Storey DJ (1982) Entrepreneurship and the new firm. Croom Helm, London
Wallsten SJ (2001) An empirical test of geographic knowledge spillovers using geographic information systems and firm-level data. Regional Science and Urban Economics 31: 571-599
Wever E, Stam E (1999) Clusters of high technology SME's: the Dutch case. Regional Studies 33(4): 391–400