Centre of Excellence
for Science and Innovation Studies

CESIS Electronic Working Paper Series

Paper No. 190

Agglomeration Externalities and Entrepreneurship

- micro-level evidence from Sweden.

Apostolos Baltzopoulos*

(*CESIS and the Division of Economics, KTH)

August 2009

The Royal Institute of Technology
Centre of Excellence for Science and Innovation Studies (CESIS)
http://www.cesis.se

Agglomeration externalities and entrepreneurship. Micro-level evidence from Sweden

Apostolos Baltzopoulos

Division of Economics, Royal Institute of Technology (KTH), Stockholm, Sweden

(e-mail: apobal@infra.kth.se)

Abstract

Past research on the effects of agglomeration externalities on regional economic development is inconclusive and has focused mainly on employment growth and innovative output. This paper considers the link between agglomeration externalities and entrepreneurship. It does so by looking at the importance of Marshallian specialization and Jacobian diversity externalities for regional entrepreneurial output implementing an individual level data set that allows considering not only the effect on total number of start-ups but also on the propensity of the entrepreneur to start his new venture in an industry he has previous experience in. The results suggest that while Marshallian externalities have a positive, Jacobian externalities have a negative effect on regional entrepreneurial output. However, Jacobian externalities increase the probability that an entrepreneur will start a firm in an industry he has relevant experience in, especially in the case of knowledge intensive industries.

Keywords: Entrepreneurship, externalities, spatial agglomeration.

JEL classification: O12, O18, R11, R30.

1. Introduction

The present study seeks to shed light on the role agglomeration externalities play on the emergence of entrepreneurship, identifying a mostly overlooked topic in relevant literature. Agglomeration externalities is a term used to describe the specialization and concentration externalities (MAR externalities after Marshall, 1890 – Arrow, 1962 – Romer, 1986), and economic and social diversity externalities (Jacobs externalities after, Jacobs, 1969) that arise from the spatial concentration of economic agents. Moreover, Porter (1990) identifies an important role for the intensity of competition (Porter externalities) and Hoover (1937, 1948) argues for the significance for regional economic development of a fourth externality, urbanization.

Following the seminal contribution by Glaeser et al. (1992) the significance for regional economic activities of agglomeration externalities has become the subject of much empirical research. The literature has mostly focused on the effects of agglomeration externalities on overall employment growth (e.g., Glaeser et al., 1992; Henderson et al., 1995; Rosenthal and Strange, 2003) and regional innovation activities (e.g., Feldman and Audretsch, 1999; Paci and Usai, 1999; van der Panne and van Beers, 2006). At the same time, robust evidence exist of a positive relation between entrepreneurship and local growth (Acs and Armington, 2004; Braunerhjelm. 2007) but the remaining link between agglomeration and entrepreneurship has received little attention.

Knowledge, whose flow is facilitated by agglomeration, is an essential input of entrepreneurship, which is characterized by the exploitation of Kirznerian as well as Schumpeterian opportunities (Kirzner, 1973; Schumpeter, 1934; Shane, 2003) and entrepreneurship is not strictly limited to innovation (Aldrich and Martinez, 2001). Therefore, the results of studies focusing on regional innovative output cannot be easily assumed to hold for entrepreneurship in general.

One study that does examine the link between agglomeration and entrepreneurship is the one by van Oort and Atzema (2004) that considers the effect of agglomeration economies on the location choice of new firms in the Dutch ICT industry. Their results do not provide a strong support for any of the externalities hypothesis since they report positive effects from both specialization and diversity. Given the conflicting results in the literature and the distinctiveness of their approach, van Oort and Atzema (2004) conclude that their results need to be treated with caution and call for further research on the particular issue providing further motivation for the present paper.

In this study I use a comprehensive data set that matches all individuals in Sweden to their place of work over the period 1999-2005. The detail of the data is such that allows several refinements over past efforts. By matching new firms to their founders and their working environment prior to becoming entrepreneurs I am able to assess the effect of the various externalities on the crucial gestation period of new business ventures. Furthermore, I am able to identify whether entrepreneurs will start a new firm in the same industry they were previously employed in and test whether this decision is affected by MAR or Jacobian externalities.

The extent to which an entrepreneur remains in the same industry is of particular interest given evidence that founder experience in the industry reduces the likelihood of new venture failure (Bruderl and Preisendorfer, 1998; Bruderl et al., 1992; Gimeno et al., 1997), and that entrepreneurs with greater industry experience form firms that grow faster (Reynolds, 1993; Dahlstrand, 1997; Bruderl and Preisendorfer, 1998) and are more profitable (Gimeno et al., 1997; Kalleberg, 1986).

The methodology that I use is twofold. First, I regress a count of new entrepreneurs per industry per region on a set of production structure characteristics using a negative binomial model. Secondly, I apply a binomial logit model to capture the effect of the same production structure characteristics on the probability that the entrepreneur will start his new firm in the same industry he was previously employed in.

The results suggest that while it is Marshallian externalities that promote overall regional entrepreneurial output, Jacobian externalities help promote entrepreneurship of "higher quality" by increasing the probability that entrepreneurs will start a firm in an industry they have relevant experience in, especially so in knowledge intensive sectors.

The rest of the paper is organized in the following manner. Section 2 discusses the theories that explain the manner in which agglomeration externalities are assumed to affect regional economic development and gives a brief overview of the results of past empirical studies that test these hypothesis. Section 3 describes the data and section 4 the empirical model used in the present study. Section 5 displays and discusses the results of the analysis while section 6 concludes.

2. Agglomeration externalities. Theory and past empirical results

Theory

The endogenous growth theory (Romer, 1986, 1990; Lucas, 1988; Krugman, 1991) underlines the importance of knowledge spillovers and externalities in inducing self-reinforcing increasing returns to scale within a geographically bounded region supporting the agglomeration of economic activities. The term "externalities" refers to economies of scale external to the firm. A stream of empirical literature that started with the seminal contribution by Glaeser et al. (1992) has sought to examine the respective roles of specialization and diversity regarding local and regional development (Greunz, 2004).

According to Marshall (1890) firms in industries that exhibit a high degree of regional concentration may benefit from a rich pool of specialized workers, the emergence of secondary services specific for the particular industry, easy access to intermediate inputs and technological or knowledge spillovers among firms. According to Griliches (1979), by working on similar things (and hence benefiting from each other's research and know-how) knowledge spillovers increase the stock of knowledge available for each individual firm. The argument first introduced by Marshall (1890), Arrow (1962), and Romer (1986) and later formalized by Glaeser et al. (1992) as the Marshall-Arrow-Romer (MAR) model, is that tacit knowledge, the sort that is spatially bounded in the region in which it was originated since its transmission requires face-to-face social interaction (Feldman and Audretsch, 1999), is industry specific. Therefore, gains from knowledge spillovers may only be realized among same-industry concentrated firms. Henderson et al. (1995) refer to this type of industry-specific externalities that arise from regional agglomeration localization externalities.

Urbanization externalities differ from localization externalities in that they are not industry specific. According to Hoover (1937, 1948) and Israd (1956) all economic actors irrespective of industry stand to gain from the regional agglomeration of general economic activities. Their arguments mirror those original put forward by Marshall (1890) and build their case on the possible gains from a rich labour pool, a big local consumer market, and well developed infrastructure, not limited to but also including public facilities such as universities and research institutes.

In contrast to localization externalities, Jacobs (1989) argues that it is variety/diversity of regional economic activity that fosters innovation and growth. While the positive effects of Jacobs externalities are still realized through the appropriation of knowledge spillovers among different actors, it is inter-industry rather than intra-industry spillovers that are assumed to be of importance. According to Jacobs local economic actors stand to gain more from an availability of a diverse knowledge base rather than industrial specialization. She argues that innovations and new business ideas are more likely to be the product of interaction among actors from different industries resulting in new knowledge combinations.

One additional prediction of the MAR theory is that local monopoly further promotes innovation by allowing externalities to be internalized by the innovator. Porter (1990) agrees on the premise that knowledge spillovers are industry specific but argues that it is local competition rather than monopoly that advances innovation since survival in the face of strong competition requires businesses to pursuit the development and adoption of innovations.

Previous empirical results

In assessing the effects of agglomeration externalities on regional economic development the study by Glaeser et al. (1992) has been extended in several ways.

An array of different units of regional classification and indices has been applied over time. The object of the analysis has not been constant either creating problems for the comparability of results[1].

Focusing on employment growth Glaeser et al. (1992) find positive effects from diversity rather than regional specialization. Henderson et al. (1995) provide evidence for the exact opposite while Combes (2000) argues that specialization has a negative effect but reports inconclusive evidence on the significance of diversity.

A separate branch in this literature tests the effect of agglomeration externalities on regional innovation output, measured as either the number of patents or the number of new products as advertised in technical magazines that can be attributed to each region. Feldman and Audretsch (1999) find that innovative output tends to be lower in cities which are specialized in the industry in question. Past research by the same authors provided similar results (Audretsch and Feldman 1996a, b) and Duranton and Purga (2000) refer to the prevalence of Jacob's thesis in the US as a stylized fact.

However, van der Panne (2004) applies a similar approach in Netherlands and finds evidence to the exact opposite. Moreover, Paci and Usai (1999, 2000), Greunz (2004) and van der Panne and van Beers (2006) report positive effects on innovation output from both Marshall and Jacobs externalities. In particular, van der Panne and van Beers (2004) argue that MAR externalities play a crucial role in the emergence of an innovation while Jacobs externalities are more important for the subsequent market progress of the new product while Greunz (2004) suggests that MAR externalities matter more in sectors with lower technological intensities as opposed to Jacobs externalities that matter for high tech innovations in "high density" areas. It is obvious that the matter is far from resolved.

[1] See de Groot et al. (2007) for a comprehensive meta-analysis

3. Data

The paper utilizes an unbalanced panel data compiled by Statistics Sweden and referred to as FAD which is the acronym of "Firms and establishments dynamics" (Företagens och arebtsställenas dynamik). FAD contains linked information on all firms, establishments and working individuals in Sweden and the present paper uses data describing the period 1999-2005. The primary data sources are the Swedish tax office and firms' annual reports. The data set contains information on individuals' education, occupation and places of origin and residence and organizational and financial data for firms and establishments.

By tracking the employment status of each working individual in Sweden between consecutive years I am able to identify entrepreneurs as the individuals that switch from being employed in a firm to becoming the owners of their own newly found firm. Such an individual-level approach allows measuring entrepreneurship in a more accurate way than using start-up rates as is the case in van Oort and Atzema (2004) since often new firms are the result of mergers, splits or expansions of incumbent firms and are not the least associated with any sort of innovation or new business idea. This is a distinction I take in consideration when identifying entrepreneurs in the case of the Swedish data. See Baltzopoulos (2009) for a detailed analysis of the methodology used.

The unit of analysis is the number of entrepreneurs starting a new firm in industry *i* in region *j*. Excluding the agricultural and public services sectors, the Swedish economy is broken down in 43 industry branches based on the 2-digit level of the Swedish Standard Industrial Classification codes (SNI) used by Statistics Sweden, which correspond to the NACE - Classification of Economic Activities in the European Community - codes. Spatially, Sweden is broken down in 81 functional regions which include the major urban centers and the corresponding surrounding areas, based on commuting patterns of the labor force.

Table 1 presents the summary statistics of the period's average counts for each industry branch in each region. Note that if all industries were present in all functional regions there should be *43*81=3,483* observations. Instead there are only 2,117. Clearly, not all industries are present in all functional regions. Table 1 also makes a distinction between Knowledge Intensive sectors[2] and the rest of the industries. On average, a total of 12,457 entrepreneurs left their previous job position to start their own firm. The knowledge intensive sectors seem to outperform the rest when it comes to spawning entrepreneurs with a mean value of entrepreneurs per industry per region of 8.3 as opposed to 5.2.

In total 66,899 entrepreneurs have been identified over the period 1999-2005. For 40,432 (60%) of those, data availability permits to see the industry in which they started their new firm in. Table 2 reports the percentages of those that started a new firm in the same industry as the one of their previous employment, again distinguishing between the knowledge intensive sector and the rest of the industries. Overall, 40% of the entrepreneurs remain in the same industry they are already familiar with based on their previous experience. This percentage is considerably higher among people working in knowledge intensive sectors (48%) compared to the rest of the industry branches (35%).

Note that while the data also allows the identification of entrepreneurs that were previously unemployed or have just graduated from university I do not take then into consideration when calculating regional entrepreneurial activity. I assume that in order to potentially benefit from knowledge spillovers and the corresponding agglomeration externalities an individual needs to have been part of

[2] SNI codes 29-33 (high-tech manufacturing) and 72-74 (knowledge intensive business services)

the local business environment prior to his decision to start a firm. This approach underestimates the levels of entrepreneurship but it does not do so by a large margin while at the same time it focuses on those startups that are more relevant for a study on the significance of the various externalities.

4. Empirical design

The aim of the present paper is to determine whether the Marshallian model of specialization externalities and local market power or the Jacobian model of diversification externalities and local competition can best explain regional entrepreneurial output[3]. To achieve this, an analysis is carried out on two levels; that of an industry sector in each of Sweden's functional regions and that of the individual entrepreneur. The first step is to regress the count of entrepreneurs in industry i in region j on a set of production structure characteristics and the second step is to assess the effect of the same set on the probability that an individual entrepreneur will start a firm in the same industry he or she was previously employed in. The production structure characteristics of interest are the degree of industrial specialization, the degree of regional diversity, and the degree of local competition.

In order to measure the degree of industrial specialization and subsequently the effect of MAR externalities I use the production structure specialization index (PS) (Feldman and Audretsch, 1999; Paci and Usai, 1999). The PS-index measures the extent to which region j is specialized towards industry i:

[3] The positive significance of urbanization externalities is a very robust result in all relevant studies and is included only as a control variable

$$PS_{ij} = \left[\frac{E_{ij}}{\sum_i E_{ij}} \right] / \left[\frac{\sum_j E_{ij}}{\sum_i \sum_j E_{ij}} \right] \qquad (1)$$

where $i = 1, \ldots, 43$ for each industry branch

$j = 1, \ldots, 81$ for each functional region

E = employment

For practical purposes the PS-index is standardized using the formula (PS-1)/(PS+1) to make it balanced and constrained within the interval (-1,1). This way positive values of the PS-index refer to industries whose share of employment in a particular region is greater to this industry's share in national employment, while negative values refer to the exact opposite.

The degree of a region's industrial diversity is captured through the use of the reciprocal of a Gini coefficient (Paci and Usai, 1999; van der Panne and van Beers, 2006; Greunz, 2004). That is defined as:

$$PD_j = \left[\frac{2}{(n-1)Q_n} \sum_{i=1}^{n-1} Q_i \right] \qquad (2)$$

where Q_i is the cumulative sum of employees up to sector i when sectors are listed in increasing order. The PD-index is defined in the interval (0,1) and larger values correspond to a higher degree of regional industrial diversity, or in other words to stronger Jacobian externalities.

Based on these definitions, specialization and diversity are not mutually exclusive. While the PS-index refers to a particular industry in a particular region, the PD-index refers to the region as a whole. A region may therefore exhibit a relative specialization in e.g. computer-related activities but also a high degree of diversity, through a strong presence of multiple other industrial branches.

The degree of local competition is measured by the competition coefficient, COMP, which like the PS-index refers to industry *i* in region *j*. It is defined as:

$$COMP_{ij} = \left[\left(S_{ij} / E_{ij} \right) / \left(\sum_i \sum_j S_{ij} / \sum_i \sum_j E_{ij} \right) \right] \qquad (3)$$

where S = number of firms

The COMP-index is the ratio of the number of firms in industry *i* in a region *j* over its national equivalent. I apply the same standardization as in the case of the PS-index by calculating (COMP-1)/(COMP+1). High values of the COMP-index are indicative of strong local market competition in a particular industry, while low values are indicative of fewer but larger firms that enjoy more market power.

Table 3 presents the summary statistics of the three indices along with those of the count of entrepreneurs and the log of the size of the region, measured by total local employment. The count of entrepreneurs per industry and region follows a Poisson distribution suggesting the use of a count data model, but there is a clear overdispersion problem with the standard error being more than four times larger than the mean. To deal with this problem the preferred estimator is that of the negative binomial model. The probability distribution function of the model is:

$$P(y_{ij} \mid x) = \frac{\Gamma(y_{ij} + a^{-1})}{y_{ij}!\,\Gamma(a^{-1})} \left(\frac{a^{-1}}{a^{-1} + \mu_{ij}} \right)^{a^{-1}} \left(\frac{\mu_{ij}}{a^{-1} + \mu_{ij}} \right)^{y_{ij}} \qquad (4)$$

where y_{ij} = number of new entrepreneurs per industry *i* per region *j*

Γ = gamma function

a = unobserved heterogeneity parameter among observations

μ_{ij} = $\exp(\mathbf{x'_{ij}\beta})$

$\mathbf{x'_{ij}}$ = $\begin{bmatrix} 1 & PS_{ij} & PD_j & COMP_{ij} & Ln(size)_j \end{bmatrix}$

$\boldsymbol{\beta}$ = vector of coefficients

The vector of control variables $\mathbf{x'_{ij}}$ will also be extended to include a Knowledge Intensive industry dummy plus interaction terms between this dummy and the three production structure indices.

In the second stage of my analysis, in order to determine the effect of the various production structure characteristics on the probability that an entrepreneur will start a business in the same industry he or she was previously employed in, I use a binomial logit model. The control variables included in the negative binomial regression are supplemented by three individual characteristics controls that account for the age, tenure, and sex of the entrepreneur.

An issue that needs to be addressed is the difference in the treatment of the data for the negative binomial compared to the binomial logit regressions. Given that regional industrial specialization and diversity as well as entrepreneurial output do not change much between consecutive years there is no gain from using a panel data approach when the unit of analysis is the entrepreneurial output of an industry branch in a functional region but pooling the data is not an option either since the cross-sections are not independent. For this reason average values of employment and entrepreneurship for the 1999-2005 period for each industry/region are used in applying the negative binomial model. However, since the unit of analysis for the binomial logit model is the individual entrepreneur and these are never the same people between different cross-sections[4] it is safe to run the binomial logit model on the independently pooled cross sections including the appropriate set of time dummies.

The logit individual-specific model which considers the probability that the entrepreneur will stat his firm in the industry he was previously employed in reads:

$$\Pr(z_k = 1 \mid x, h, t) = \Lambda(a_k + \mathbf{x}_{ijk}\boldsymbol{\beta}_1 + \mathbf{h}_k\boldsymbol{\beta}_2 + \mathbf{t}\boldsymbol{\beta}_3 + \varepsilon_k) \qquad (5)$$

where $\Lambda(c) = \exp(c)/\left[1+\exp(c)\right]$

k = 1, … , 40432, for each entrepreneur in the data

$z_k = 1$ if the entrepreneur stars a firm in the same industry he was previously employed in, zero otherwise.

a_k = unobserved heterogeneity parameter among individuals

$\mathbf{x}_{ijk}$ = same as above, controls for the characteristics of the industry/region the entrepreneur k was living/working in before becoming self-employed

$\mathbf{h}'_k = \left[Tenure_k \quad Age_k \quad Male_k\right]$, individual specific controls

$\mathbf{t}$ = vector of time dummies

$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ = vectors of coefficients

ε_k = i.i.d. error term

Table 4 presents the summary statistics of the variables included in the binomial logit regression. It is interesting to note that the profile of the typical Swedish entrepreneur matches almost perfectly that of studies from other countries being a male of around 40 years old.

[4] Here I impose the rather safe assumption that an individual will not have sufficient time to switch from being employed to becoming an entrepreneur to becoming employed again and then start yet again a new firm in the time period considered.

5. Results

The results of the two-level analysis are presented in tables 5 and 6. Table 5 reports the estimation of the negative binomial model based on equation 4 where the dependent variable is the count of new entrepreneurs in industry *i* in region *j*. Table 6 presents the estimation of the binomial logit model of equation 5 where the dependent variable is the probability that an entrepreneur will start his new firm in the same industry he was previously employed in. In both cases three different specifications were fitted and the respective results are listed in columns I, II and III.

In the case of the negative binomial model the basic specification (I) includes only the measures of specialization (PS index), diversity (PD index), competition (COMP index), and urbanization (Ln(size)). In the case of the binomial logit model the basic specification (I) also includes the individual controls *Tenure, Age, Male*. Specification II also includes a Knowledge Intensive dummy (KI) to control for the fact that Knowledge Intensive industries exhibit a higher propensity to concentrate geographically (Audretsch and Feldman, 1996a). Specification III is augmented by three interaction variables between the KI dummy and the measures of specialization, diversity and competition following Greunz (2004) who reports that MAR and Jacobs externalities influence high tech innovations differently than those in sectors with lower technological intensities. The alternative specifications also work as robustness checks.

In table 5 clear support for the Marshallian thesis emerges as expressed by the combination of a positive and significant coefficient on the PS-index and a negative and significant coefficient on the PD-index and COMP-index in all three specifications. Entrepreneurial output is affected positively by local industrial specialization and negatively by regional economic diversity and local market competition. As predicted, Knowledge Intensive industries spawn more entrepreneurs as evidenced by a positive and significant coefficient on the KI dummy for both specification II and III.

Moreover, the interaction variables in the third column show that Knowledge Intensive sectors benefit from industrial concentration even more than the rest of the economy but also suffer higher barriers in entrepreneurial output from local competition (positive and significant effect of PS*KI, negative and significant effect of COMP*KI). Finally, while regional economic diversity seems to have a negative effect on entrepreneurial spawning this effect does not differ between high and low tech industries (the coefficient on the PD*KI interaction dummy is not significant). These results only partly agree with van Oort and Atzema (2004) that report a positive effect from both Marshallian and Jacobs externalities but are in line with van der Panne (2004).

Turning to table 6 and the results of the binomial logit model is worth noting that this approach is novel and to my knowledge no comparable studies exist. The results measure the effect of the agglomeration externalities on the probability that an individual will start a new firm in the same industry he was previously employed in (conditional on having decided to become an entrepreneur). This probability increases with the degree of local industrial specialization (PS index) and decreases with the degree of local market competition (COMP index). These results are robust in all three specifications. However, regional economic diversity, despite having a negative effect on overall entrepreneurial output, has a significantly positive effect on the probability that the individual will remain in the same industry he was previously employed in (coefficient on PD index positive and significant on all three specifications).

Focusing on the interaction variables in specification III two more interesting points arise. The positive effect of diversity is even larger in the case of Knowledge Intensive industries (see coefficient on PD*KI) while that of industrial specialization does not seem to affect entrepreneurs in those industries differently than the rest (coefficient on PS*KI non-significant).

Given the premises that entrepreneurs are more successful in industries they have a relevant experience in and are more likely to contribute to economic development when starting firms in knowledge intensive sectors (Shane, 2008), it appears that Jacobian externalities have not lost the battle. Although diversity externalities seem to have a negative effect on overall entrepreneurial output they have a positive effect on spawning entrepreneurs of "higher quality"; meaning entrepreneurs that have a relevant experience in the industry they start a firm in especially in knowledge intensive sectors. Due to the novelty of the approach these results need to be treated with caution they do however suggest that further research on this subject using data of similar detail should be carried out.

6. Summary and conclusions

The present paper has sought to add a new approach to the research dealing with the effects of agglomeration externalities on regional economic development. Research on this subject has focused mostly on employment growth and innovation output overlooking to some extent the significance for entrepreneurship, as witnessed by the creation of new firms.

The usual approach of regressing employment growth or innovative output on a set of regional production structure characteristics is applied on a count of entrepreneurs instead, similarly to van Oort and Atzema (2004), without however restricting the focus on the ICT industry alone. This traditional approach is supplemented by an assessment of the role these same production structure characteristics play on the probability that the entrepreneur will start a firm in the same industry he was previously employed in. The latter can be motivated by the fact that survival and success rates increase when the entrepreneur has past experience in the industry he starts his firm in.

The results suggest that while Marshallian externalities in the form of industrial concentration and low local market competition provide the grounds for higher rates of regional entrepreneurial output Jacobian externalities will increase the probability that the entrepreneur will start his firm in the industry he has previous experience in. These findings call into attention the qualitative on top of the quantitative effects of agglomeration externalities and suggest an explanation of the incongruity of past results that focus on different aspects of regional economic development.

References

Acs, Z.J., and Armington, C. (2004) Employment growth and entrepreneurial activity in cities. *Regional Studies*, 38: 911-927.

Aldrich, H. E., and Martinez, M. A. (2001) Many are called, but few are chosen: An evolutionary perspective for the study of entrepreneurship, *Entrepreneurship: Theory and Practice*, vol.25, issue 4, p41-57.

Arrow, K. (1962) Economic welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity,* pp. 609-626. Princeton: Princeton University Press.

Audretsch, D. B., and Feldman, M. (1996a). R&D spillovers and the geography of innovation and production. *American Economic Review*, 86, 630-640.

Audretsch, D. B., and Feldman, M. (1996b). Innovative clusters and the industry life cycle. *Review of Industrial Organization*, 11: 253-273.

Baltzopoulos, A. (2009) The firm and the region as breeding grounds for entrepreneurs. CESIS Working Papers Series.

Braunerhjelm, P. (2008) Entrepreneurship, knowledge and growth. *Foundations and Trends in Enterpreneurship*, 4:5.

Bruderl, J., and Preisendorfer, P. (1998) Network support and the success of newly founded businesses. *Small Business Economics*, 10: 213-225.

Bruderl, J., Preisendorfer, P., and Ziegler, R. (1992) Survival chances of newly founded business organizations. *American Sociological Review*, 57:227-302.

Dahlstrand, Å. (1997) Growth and inventiveness in technology-based spin-off firms. *Research Policy*, 26: 331-344.

De Groot, H. L. F., Poot, J., and Smit, M. (2007). Agglomeration, innovation and regional development: Theoretical perspectives and meta-analysis. Tinbergen Institute Discussion Paper No. 07-079/3.

Dumais, G., Ellison, G., and Glaeser, E. L. (2002). Geographic concentration as a dynamic process. *The Review of Economics and Statistics*, 84(2): 193-204.

Feldman, M. P., and Audretsch. D. B. (1999) Innovation in cities: Science-based diversity , specialization and localized competition. *European Economic Review*, 43: 409-429.

Gimeno, J., Folta, T., Cooper, A., and Woo, C. (1997) Survival of the fittest? Entrepreneurial human capital and the persistence of underperforming firms. *Administrative Science Quarterly*, 42: 750-783.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Schleifer, A. (1992) Learning in cities. *Journal of Urban Economics*, 46: 254-277.

Greunz, L. (2004) Industrial structure and innovation – evidence from European regions. *Journal of Evolutionary Economics*, 14: 563-592.

Griliches, Z. (1979) Issues in assessing the contribution of R&D to productivity growth. *Bell Journal of Economics*, 10: 92-116.

Henderson, V., Kuncoro, A., and Turner, M. (1995) Industrial development in cities. *The Journal of Political Economy*, 103(5): 1067-1090.

Hoover, E. M. (1937) *Locatinon Theory and the Shoe and Leather Industry.* Harvard University Press, Cambridge.

Hoover, E. M. (1948) *The Location of Economic Activity*. McGraw Hill, New York.

Israd, W. (1956) *Location and Space Economy*, MIT Press, Cambridge, MA.

Jacobs, J. (1969) *The Economy of Cities*. Random House: New York.

Kalleberg, A. (1986) Entrepreneurship in the 1980s: A study of small business in Indiana. *Advances in the Study of Entrepreneurship, Innovation and Growth*, 1: 157-189.

Kirzner, I. M. (1973) *Competition and Entrepreneurship*, Chicago: University of Chicago Press.

Krugman, P. (1991) Increasing returns and economic geography. *Journal of Political Economy*, 99: 483-499.

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

Marsahll, A. (1890) *Principles of Economics*. MacMillan: London.

van Oort, F. G., and Atzema, O. A. L. C. (2004) On the conceptualization of agglomeration economies: The case of new firm formation in the Dutch ICT sector. *The Annals of Regional Science*, 38: 263-290.

van der Panne, G. (2004) Agglomeration externalities: Marshall versus Jacobs. *Journal of Evolutionary Economics*, 14: 593-604.

van der Panne, G., and van Beers, C. (2006) On the Marshall-Jacobs controversy: It takes two to tango. *Industrial and Corporate Change*, 15(5), 877-890.

Paci, R., and Usai, S. (1999) Externalities, knowledge spillovers and the spatial distribution of innovation. *GeoJournal*, 49: 381-390.

Paci, R., and Usai, S. (2000) The role of specialization and diversity externalities in the agglomeration of innovative activities. *Rivista Italiana degli Economisti*, 2: 237-268.

Porter, M. E. (1990) *The Competitive Advantage of Nations*, MacMillan, Houndmills.

Romer, P. M. (1986) Increasing returns and long-run growth. *Journal of Political Economy*, 94, 1002-1037.

Romer. P. M. (1990) Endogenous technological change. *Journal of Political Economy*, 98 :71-102

Rosenthal, S. S., and Strange, W. C. (2003) Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, 85: 377-393.

Schumpeter, J. A. (1934) *Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. Cambridge: Harvard University Press.

Shane, S. (2008) *A General Theory of Entrepreneurship – The Individual Opportunity Nexus*, Edward Elgar, Cheltenham, UK – Northampton, MA, USA.

Shane, S. (2008) *The Illusions of Enterpreneurship: The Costly Myths That Entrepreneurs Investors, and Policy Makers Live By*, Yale University Press, New Haven & London

Tables

Table 1. Number of entrepreneurs in industry *i* in region *j* (1999-2005 averages).

	Number of observations	Minimum	Maximum	Sum	Mean	Std. Deviation
All 43 industry branches						
	2117	1	905	12303	5.8	27.4
Knowledge intensive sectors (26% of total employment)						
	435	1	905	3608	8.3	48.6
Excluding Knowledge intensive sectors (74% of total employment)						
	1682	1	382	8695	5.2	18.3

Table 2. Percentage of the total number of entrepreneurs in the period 1999-2005 that start a firm in the same industry as their previous employment (where identifiable)

All 43 industry branches	16273 out of 40432 40%
Knowledge intensive sectors	6481 out of 12839 48%
Excluding Knowledge intensive sectors	9792 out of 27593 35%

Table 3. Summary statistics of the variables in the negative binomial regression (Number of observations is 2117).

Variable	Mean	Std. Dev.	Min	Max
$Enterpreneurs_{ij}$	5.9	27.4	1	905
PS_{ij}	-0.07	0.38	-0.96	0.95
PD_{j}	0.35	0.04	0.21	0.45
$COMP_{ij}$	0.10	0.43	-0.93	0.99
$Ln(size)_{j}$	9.4	1.3	5.9	13.0
$Knowledge\ intensive_{ij}$	0.21	0.40	0	1

Table 4. Summary statistics of the variables in the negative binomial regression (Number of observations is 40432).

Variable	Mean	Std. Dev.	Min	Max
Same industry dummy$_k$	0.40	0.49	0	1
Tenure$_k$	3.25	4.25	0	18
Age$_k$	40.1	11.1	16	81
Male$_k$	0.79	0.41	0	1
PS$_{ij}$	0.05	0.23	-0.96	0.95
PD$_j$	0.33	0.04	0.21	0.45
COMP$_{ij}$	-0.03	0.33	-0.87	0.99
Ln(size)$_j$	11.4	1.5	5.8	13.0
Knowledge intensive$_{ij}$	0.32	0.47	0	1

Table 5. Results of negative binomial regression (see model (4)). Dependent variable: Count of entrepreneurs in industry *i* in region *j*

Variable	I	II	III
Constant	-4.68**	-4.78**	-4.78**
	[0.25]	[0.24]	[0.26]
PS index	1.55**	1.65**	1.51**
	[0.06]	[0.06]	[0.07]
PD index	-1.63**	-1.64**	-1.41**
	[0.57]	[0.56]	[0.62]
COMP index	-1.30**	-1.38**	-1.19**
	[0.05]	[0.05]	[0.06]
Ln(size)	0.70**	0.70**	0.69**
	[0.01]	[0.01]	[0.01]
Knowledge intensive	-	0.046**	1.06**
		[0.05]	[0.48]
*PS index * Knowledge*	-	-	0.83**
intensive			[0.16]
*PD index **	-	-	-1.13
Knowledge intensive			[1.38]
*COMP index **	-	-	-1.02**
Knowledge intensive			[0.13]
Number of obs.	2117	2117	2117
Pseudo R^2	0.210	0.218	0.223

**Significant at 5%; *significant at 10%; standard errors in brackets

Table 6. Results of binomial logit regression (see model (5)). Dependent variable: Probability the entrepreneur will start a firm in the same industry he was previously employed in.

Variable	I	II	III
Constant	-2.71**	-2.60**	-2.41**
	[0.21]	[0.21]	[0.23]
Tenure	0.02**	0.02**	0.02**
	[0.00]	[0.00]	[0.00]
Age	-0.01**	-0.01**	-0.01**
	[0.00]	[0.00]	[0.00]
Male	0.44**	0.51**	0.49**
	[0.03]	[0.03]	[0.03]
PS index	1.81**	1.59**	1.80**
	[0.07]	[0.07]	[0.09]
PD index	3.27**	3.36**	2.55**
	[0.37]	[0.38]	[0.42]
COMP index	-4.03**	-3.91**	-4.10**
	[0.06]	[0.06]	[0.08]
Ln(size)	0.06**	0.04**	0.04**
	[0.01]	[0.01]	[0.01]
Knowledge intensive	-	0.58**	-0.35
		[0.03]	[0.27]
*PS index * Knowledge*	-	-	-0.20
intensive			[0.15]
*PD index **	-	-	2.95**
Knowledge intensive			[0.79]
*COMP index **	-	-	0.32**
Knowledge intensive			[0.12]
Number of obs.	40432	40432	40432
Pseudo R^2	0.15	0.16	0.16

**Significant at 5%; *significant at 10%; standard errors in brackets; time-dummies omitted.