<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *Regional studies*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Wixe, S. (2015)

The impact of spatial externalities: Skills, education and plant productivity.

Regional studies, 49(12): 2053-2069

<http://dx.doi.org/10.1080/00343404.2014.891729>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-26260>

The Impact of Spatial Externalities: Skills, Education and Plant Productivity

Sofia Wixe

Centre for Entrepreneurship and Spatial Economics (CEnSE), Jönköping International
Business School, P.O. Box 1126, SE-551 11 Jönköping, Sweden.
E-mail: sofia.wixe@jibs.hj.se

ABSTRACT

This paper analyses the role of a broad range of spatial externalities in explaining average labour productivity of Swedish manufacturing plants. The main findings show positive effects from general urbanization economies and labour market matching, as well as a negative effect from within-industry diversity. These results confirm previous research despite methodological differences, which implies wider generalizability. Additionally, the empirical findings support MAR and Porter externalities, i.e. positive effects from specialization and competition. No evidence is found of Jacobs externalities, neither when measured as between-industry diversity nor as within-industry diversity. Finally, plant-specific characteristics play a key role in explaining plant-level productivity.

Keywords: plant productivity, spatial externalities, manufacturing, Sweden.

JEL Classification Codes: D24, L25, L60, R32

INTRODUCTION

Agglomeration economies capture the benefits of the co-localization of firms. These benefits are commonly analysed in terms of externalities, such as knowledge spillovers affecting productivity in nearby firms.¹ This implies that the characteristics of the local economic environment provide determinants for the productivity of individual firms. In the literature on this subject four main sources of spillover effects can be identified. JACOBS (1969) argues that among geographically close firms diversity drives innovation and growth, which is part of urbanization economies. A second view, referred to as Marshall-Arrow-Romer (MAR) externalities by GLAESER *et al.* (1992), is that specialization in only one industry promotes growth, which is part of localization economies. Third, PORTER (1998) maintains that the most important aspect for firms to innovate and become more productive is competition. The last source of externalities concerns the labour market, whose importance was acknowledged already by MARSHALL (1890). Labour market matching gives rise to knowledge spillovers since labour mobility plays an important role in knowledge diffusion (MASKELL and MALMBERG, 1999; POWER and LUNDMARK, 2004).

Starting with the seminal papers by GLAESER *et al.* (1992) and HENDERSON *et al.* (1995) there are numerous studies on externality effects. GLAESER *et al.* (1992) find positive effects from diversity, and negative effects from specialization, while HENDERSON *et al.* (1995) find positive effects from both diversity and specialization, on employment growth in US industries. The inconclusiveness has continued throughout the research on spatial externalities. As discussed by BEAUDRY and SCHIFFAUEROVA (2009), among others, this is mainly due to differences in methodological approaches. Hence, despite a vast amount of research there are still difficulties to draw major conclusions. A particular issue concerns the unit of analysis. Like GLAESER *et al.* (1992) and HENDERSON *et al.* (1995), most studies use aggregated data on, for instance, industry or regional level. These results cannot be applied to make inferences about the nature of firms, a problem referred to as ecological fallacy. In more recent years data

availability has improved and there are studies on spatial externalities at firm- and plant level. However, the inconclusiveness continues and there is still need for further research.

Sweden provides an excellent case to study, partly because it is a good example of an average industrialized country. Sweden has been a member of the European Union since 1995 and has had a right-wing government since 2006. This has reduced the historically strong welfare state and has led to privatization of public goods and services. In addition, in 2010, Sweden had the fastest growing economy, the highest level of innovation and was the most competitive economy in the European Union (WORLD ECONOMIC FORUM, 2010). Furthermore, the unique data availability allows for micro-level analyses.

The purpose of the present paper is hence to analyse the role of spatial externalities in explaining the productivity levels of Swedish manufacturing plants, for the years 2002 to 2010. Fixed- and random-effects estimates are applied, in order to capture both short-term and long-term effects. In addition, a broader than usual range of externalities is tested for. The most robust results concern general urbanization economies and labour market matching, both of which are found to enhance productivity in the short term as well as in the long term. The results also show that within-industry diversity has a negative long-term effect on productivity. These results are consistent with previous studies, despite methodological differences, which allows for a wider generalization. This is the main contribution of the present study. Additionally, region-wide industrial diversity significantly decreases productivity, both in the short term and the long term, while positive effects from specialization and competition are found for the long term only.

In addition, plant-specific characteristics, including the characteristics of the workforce, are important for explaining plant productivity. The employees have the potential to affect the way in which plants absorb and use potential spillover effects and they are therefore a crucial component for channelling knowledge to the plant as a whole. Tacit knowledge in

particular is transferred through human interaction (GERTLER, 2003). This reasoning is in line with COHEN and LEVINTHAL (1990), who assert that a firm's ability to utilize external knowledge is dependent on its prior relevant knowledge, its absorptive capacity. This in turn depends on the absorptive capacity of the owners, managers and employees of the firm, which is determined by their different perceptive powers, divergent insights and dissimilar attitudes (MASKELL, 2001). It is thus important to take account of the characteristics of the employees in a study such as the present one. The idea to include both internal (firm-specific or microeconomic) and external (regional or macroeconomic) determinants to explain firm performance is applied in BALDWIN *et al.* (2008) and ERIKSSON and LINDGREN (2009), concerning labour productivity, and RODRÍGUEZ-POSE *et al.* (2013), concerning export propensity and intensity, among others.

EXTERNALITIES

Both JACOBS's (1969) and MAR's (GLAESER *et al.*, 1992) views of externalities concern the effects of knowledge spillovers. JACOBS (1969), in her historical account of cities, highlights diversity while the MAR theory supports industrial specialization. JACOBS (1969) claims that cities are the main driving force for the economy because it is in cities that innovation and technological progress take place. The reason for this is that cities are diverse; they are comprised of a wide variety of industries and people and, according to JACOBS (1969), the most productive spillovers are those that transcend industry boundaries. There is a greater flow of different ideas in diversified environments and firms can learn from innovations in other industries.

However, diversity is but one particular aspect of urbanization economies. Urbanization economies in more general terms concern knowledge spillovers arising from the concentration of economic activity per se, measured e.g. as population density. CICCONE and

HALL (1996) and CICCONE (2000) find a significant positive relationship between average labour productivity and employment density for the USA and five European countries, respectively. Regarding Swedish municipalities KARLSSON and PETTERSSON (2005) find that access to own population is significant in explaining the regional gross domestic product (GDP) per square kilometre.

Turning to localization economies and specialization, the first contributor to the theory of MAR externalities is MARSHALL (1890). According to MARSHALL (1890), industrial concentration promotes knowledge spillovers within the industry, which increases growth in the industry and the city as a whole. In ARROW's (1962) model, knowledge is created as a by-product of ordinary production and learning is equal to gaining work experience. Hence, only firm-specific knowledge is accounted for. The last contribution to the MAR theory is by ROMER (1986, 1990) who states that new technology is re-invested in firms and knowledge is therefore internalized. Hence, neither ARROW (1962) nor ROMER (1986, 1990) take knowledge flows between industries into account.²

JACOBS (1969) also asserts that competition is important for an economy to prosper since competition forces firms to innovate in order to survive. This view of competition is in line with PORTER (1998), but opposite to the MAR theory. ARROW (1962) and ROMER (1990) consider technology and knowledge as non-rival goods; the former views them as completely non-excludable and the latter as partially excludable. Competition is negative for the economy since the incentives for firms to innovate are reduced when free-riding is possible and there are risks of not gaining full return on innovations. PORTER (1998) agrees with the MAR theory that knowledge spillovers occur within industries but disagrees that competition decreases innovation. PORTER (1998) argues that competition is positive since even though it reduces the returns on innovations it puts pressure on the firms to become more productive.

CAPELLO (2002) finds that for high-tech firms in Milan the effect of urbanization economies on total factor productivity is negative, while the effect of localization economies is positive. For Spanish manufacturing branches DE LUCIO *et al.* (2002) show that specialization initially affects productivity growth negatively. However, as specialization grows stronger it has a positive effect. No effects are found for either diversity or competition. BALDWIN *et al.* (2008) find that industrial concentration enhances labour productivity in Canadian manufacturing plants, supporting localization economies. A recent study by MARTIN *et al.* (2011) concerns the effects of urbanization and localization economies, diversity and competition on the total factor productivity of French manufacturing plants. Evidence is found of localization economies only.

FRENKEN *et al.* (2007) took the question of regional diversification, or variety, one step further. It was argued that for knowledge spillovers to enhance growth there needs to be some cognitive proximity between firms. A distinction was thus made between related and unrelated variety where related variety is defined as within-industry diversity and unrelated variety as between-industry diversity. FRENKEN *et al.* (2007) argue that Jacobs externalities should be measured as related variety, which had not been the case in most previous studies. Regarding Dutch regions, they find that related variety enhances employment growth while it decreases productivity growth. Other studies have followed in the footsteps of FRENKEN *et al.* (2007), such as BISHOP and GRIPAIS (2010) and BOSCHMA *et al.* (2012).

Another important externality that is embedded in both urbanization and localization economies concerns labour market matching. Already MARSHALL (1890) acknowledged the importance of the labour market, and this source of externalities is referred to as “Marshall labour pooling” (WORLD BANK, 2009). The link to urbanization economies springs from workers bringing innovations from one industry to another, which implies that the benefits arise from industrial diversity. Marshall labour pooling is related also to localization

economies since workers with industry-specific skills cluster in locations specialized in the corresponding industry. According to COMBES and DURANTON (2006), firms that localize in clusters have access to larger labour markets and to employees who already have the relevant knowledge, which reduces training costs. At the same time they face the costs of losing their own knowledge to other firms as well as the costs of having to pay higher wages in order to retain their workers.

Labour mobility is not directly considered in the present paper, but it is connected with labour market matching. Industrial performance and innovation are dependent on the movement of people between labour markets, sectors and firms (POWER and LUNDMARK, 2004). Partly due to tacit knowledge being embedded in individuals rather than being “in the air”, and partly due to labour mobility playing a key role in knowledge diffusion, labour mobility provides an important source for firms to acquire new knowledge (cf. MASKELL and MALMBERG (1999), GERTLER (2003), POWER and LUNDMARK (2004) and BOSCHMA *et al.* (2009), among others). This is related to labour market matching since better matching increases the potential for beneficial labour mobility.

ANDERSSON *et al.* (2007) show that the matching of firm and worker quality contributes to the urban productivity premium in the USA. BALDWIN *et al.* (2008) measure labour pooling in terms of similarity between the occupation mix of the metropolitan areas and the occupation mix of the industries in those areas. The results show that labour productivity is higher for plants located in areas with a stronger match between the local labour pool and the labour pool of the industry that the plant belongs to. Using Swedish data, ERIKSSON and LINDGREN (2009) find that externalities from the labour market are more important for firm productivity than externalities from concentration and diversity. It is argued that what is important is not labour mobility in itself but in combination with labour market matching.

In addition, the employment rate is a potential motivator for employee productivity. Low employment rates are associated with recessions, which implies decreasing real wages. According to AKERLOF and YELLEN (1990), employees respond by lowering their effort, especially if the wage falls below the level considered to be fair. DARITY and GOLDSMITH (1996) argue that being unemployed has a negative effect on psychological well-being, which affects productivity if the unemployed person becomes employed. On the other hand, a higher employment rate might affect productivity negatively due to the entrance of lower-skilled workers. BELORGEY *et al.* (2006) find this effect when comparing 25 countries.

The sources of agglomeration economies can be connected to three types of mechanisms behind agglomeration economies, identified by DURANTON and PUGA (2004) as sharing, matching and learning. The agglomeration of firms increases the possibilities of *sharing* indivisible goods and facilities. This concerns investment and risk as well as labour pooling. Agglomerations of firms attract labour, and vice versa, which provides a market for skilled employees. Increased agglomeration also increases the quality of the *matching* on the labour market, which enhances productivity, innovation and growth. Finally, since urban environments gather a large number of firms and people, the potential for knowledge spillovers, human capital accumulation and thus *learning* increases.

DATA

This study was made possible through the use of micro-level data, obtained by Statistics Sweden, containing detailed information about all firms, plants and employees in Sweden. These are connected by identity numbers, through which each employee is linked to both a plant and a firm. 98 per cent of the firms in Sweden comprise only one plant. However, the last two per cent account for 50 per cent of the firms' total value added, which makes it difficult to disregard them. Since most of these firms have plants in different municipalities it is also

problematic to conduct the analysis at the firm level. What is the location of a firm with, for example, one plant in Stockholm and one in Jönköping? This is crucial for the present study. Accordingly, productivity is measured at the plant level, which raises another issue since data such as value added are at the firm level. This is dealt with by using the plants' shares of employment in their respective firms as weights to distribute the firm value added, as done by MARTIN *et al.* (2011).³ The figures obtained do not perfectly reflect the true ones but are the closest approximations possible. A dummy variable is introduced to identify the plants belonging to multi-plant firms. In addition, all the estimates are performed for single-plant firms only, with consistent results.

Plants with a non-positive value added are excluded, as are those with fewer than one employee and a non-positive physical capital level.⁴ Last, industrial classifications are restricted to two digits and only the manufacturing sector is included, industries 15 to 37 according to the 2002 standard industrial classification (SIC) by Statistics Sweden. Following this procedure, 39,730 individual plants and a total of 205,087 observations remain. These constitute the unbalanced panel for the nine years ranging from 2002 to 2010. On average, the plants show up in the dataset 5.2 times, which implies that there are both start-ups and close-downs of plants. The distribution of plants among industries together with a description of the industries included is found in Table A1. The period starts with an upward trend in the Swedish economy as it recovers from the burst of the dot-com bubble. The economy shows positive growth, which is especially strong in 2004 to 2007, until the worldwide financial crisis hits also Sweden in 2008 and 2009. Finally, 2010 is a year of strong recovery, with a growth rate exceeding the rates of the mid-2000s. To control for these annual differences yearly dummies are introduced.

The regions referred to throughout this paper are, unless stated otherwise, the 290 Swedish municipalities. These are the smallest administrative units in Sweden used for self-

governing purposes. Municipalities are chosen over the larger labour market areas as the unit for the calculation of the regional variables. This is due to recent research showing that agglomeration economies, in particular knowledge spillovers, are bounded in space (ANDERSSON *et al.*, 2012; BALDWIN *et al.*, 2008; ROSENTHAL and STRANGE, 2008). This implies that externalities are local or neighbourhood effects rather than regional effects, which motivates a geographical unit smaller than labour market areas. At the time of this research, municipalities are the smallest geographical units identified in the data set. In addition, there is a large literature on the role of interaction and face-to-face contacts for learning and innovation, especially considering tacit knowledge (see e.g. MASKELL and MALMBERG (1999) and GERTLER (2003)). BOSCHMA (2005), in a conceptual discussion on different dimensions of proximity, states that even though geographical proximity is neither necessary nor sufficient for knowledge spillovers to occur, it facilitates interactive learning.

VARIABLES

The dependent variable in this study is average labour productivity, which is measured at the plant level as value added per employee.

Plant characteristics

To analyse plant performance it is necessary to account for plant-specific characteristics. Capital and labour are the two original factors in production functions. In the present paper, capital is measured as the book value of material assets and labour as the number of full-time equivalent employees. In addition, number of years since establishment, average age of employees, percentage of females, and industry dummies at the two-digit level are introduced as control variables.

As discussed in the introduction, the employees play a prominent role in determining productivity levels. To assess the potential for absorptive capacity (COHEN and LEVINTHAL, 1990) at the plant level the abilities, in terms of skills and education, of the employees are taken into account. Education is measured as the share of employees at each plant with three or more years of university education. Regarding the skills, the division by JOHANSSON and KLAESSON (2011) is followed. Occupations are sorted into four different categories based on the work tasks performed in each occupation, where the work tasks are assumed to reflect the abilities of the employees, beyond formal education. Based on their occupation, employees are thus categorized as having cognitive skills, management and administration skills, social skills, or motoric and other skills. Typical occupations for each category of skills are given in Table 1.

Table 1. Examples of occupations within skill categories.

Cognitive skills	Management and administration skills	Social skills	Motoric and other skills
Engineers Natural and computer scientists	Directors Production managers Office clerks	Sales personnel Client information clerks	Construction workers Machine operators Manufacturing workers

Source: JOHANSSON and KLAESSON (2011).

The plants' percentages of employees in the first three categories are used as explanatory variables. Percentage of motoric skills is excluded due to mutual exclusiveness among the four categories. The inspiration for this division of skills originates from BACOLOD *et al.* (2009), who saw a need to distinguish worker skills in more aspects than differences in education levels.

Skills, measured as occupation, and education to some extent overlap. Table B1 shows that the percentage of highly educated employees is positively correlated with the percentage of employees using cognitive, management and social skills, while it is negatively

correlated with the percentage of employees using motoric skills. The highest correlation is with cognitive skills, which is logical since these occupations commonly require a degree in higher education. However, even though there is a correlation between education and skills, it is not high enough to cause problems with multicollinearity, as shown by the VIF values in Table C1. In addition, education and occupation measure different things. Education shows the formal background of the employees, while skills, in terms of occupation, measure what the employees are actually doing. Even though educational background influences the occupation of an individual, there are many more determinants, such as work experience and personal characteristics.

Regional characteristics

General urbanization economies are commonly measured by density, e.g. population per square kilometre. This is due to density being a measure of economic activity per se, irrespective of its composition. However, for a country like Sweden with a relatively large area and a relatively small population concentrated in urban areas, conventional density measures do not describe the real economic structure. The density of economic activity is instead measured as the size of the accessible market, in terms of wage sums (WS). JOHANSSON *et al.* (2002) divide the accessible market into a local, an intra-regional and an extra-regional part. The local market consists of the municipality in question and the intra-regional market is the functional economic region, which typically comprises four to five municipalities. The extra-regional market consists of the municipalities outside the functional region. The different accessibility measures are calculated as follows (ANDERSSON and KLAESSON, 2009):

$$A_r^l = WS_r \exp\{-\lambda_r t_{rr}\}, \quad (1)$$

$$A_r^{ir} = \sum_{R-r} WS_k \exp\{-\lambda_{ir} t_{rk}\}, \quad (2)$$

$$A_r^{er} = \sum_{W-R} WS_k \exp\{-\lambda_{er} t_{rk}\}, \quad (3)$$

in which A_r^l denotes the local, A_r^{ir} the intra-regional and A_r^{er} the extra-regional market accessibility for municipality r . R constitutes all the municipalities within a functional economic region and W is the set of all Swedish municipalities; t_{rk} is the travel time distance between municipality r and municipality k , where $r \neq k$. The market potential is thus adjusted for travel times between locations. Finally, the λ 's are measures of time-distance sensitivity. Using Swedish commuting data for 1998, JOHANSSON *et al.* (2003) estimated λ_r to 0.02, λ_{ir} to 0.1 and λ_{er} to 0.05.

In addition to controlling for the market potential of municipalities, the accessibility measures model how locations within a municipality are spatially related to other locations in that municipality, as well as to locations in other regions. ANDERSSON and GRÅSJÖ (2009) find that the inclusion of accessibility as a representation of spatial interaction captures the spatial dependence between locations.

To capture Jacobs externalities, or industrial diversity, an entropy measure is commonly applied, see e.g. JACQUEMIN and BERRY (1979), and ATTARAN (1986)):

$$D_r = - \sum_{i=1}^n \left(\frac{e_{i,r}}{e_r} \right) \ln \left(\frac{e_{i,r}}{e_r} \right), \quad (4)$$

in which D_r measures diversity in municipality r , $e_{i,r}$ the number of employees in two-digit industry i and municipality r and e_r the total number of employees in the manufacturing sector in municipality r . The entropy measure ranges from zero (no diversity) to $\ln n$ (maximum

diversity) where n is the total number of two-digit industries in the manufacturing sector in municipality r . However, in terms of FRENKEN *et al.* (2007), Equation 4 is a measure of unrelated variety. To measure within-industry diversity, or related variety, Equation 4 is applied for each two-digit industry using employment in the respective five-digit sub-industries. This follows the procedure by BISHOP and GRIPAIS (2010), who argue that when conducting the analysis at industry level, what is relevant is the related variety within the industry in question, not the region-wide related variety. In the present paper the observations are at an even finer level. However, since industry belonging is known for each plant, it makes sense to account for the related variety in the relevant industry.

To measure localization economies, or more specifically industry specialization, a location quotient, $LQ_{i,r}$, is applied (as in FELDMAN and AUDRETSCH (1999)):

$$LQ_{i,r} = \frac{e_{i,r}/e_r}{e_i/e}, \quad (5)$$

in which e_i measures the number of employees in two-digit industry i , e the total number of employees in the manufacturing sector in Sweden and $e_{i,r}$ and e_r as above. The location quotient is a relative measure in that it measures the regional share of workers relative to the national share of workers in a specific industry. If the location quotient is larger than one, the industry has a larger share of the employees in a region than the country as a whole, implying that the municipality is more specialized than average in that industry. Regarding knowledge spillovers, it could be argued that what matters is the absolute, and not the relative, concentration of people. However, even though specialization and diversity are not mutually exclusive, absolute measures of them are correlated, which imposes problems with multicollinearity. A relative measure is thus chosen for specialization. On the other hand, Equation 4 can be regarded as a proxy for inverse industrial specialization. An increase in industrial diversity implies a decrease

in industrial specialization, and vice versa. In addition, the related variety concept can be applied. Unrelated variety is then a proxy for inverse region-wide, or between-industry, specialization while related variety is a proxy for inverse within-industry specialization. The entropy measure is not as commonly applied to measure specialization as e.g. the Herfindahl index but the two are strongly correlated (PALAN, 2010).

Competition arises when there are many producers of similar products, and competition can thus be measured by the distribution of employees over plants in an industry. Hence, an entropy measure at the industry level is constructed:

$$C_{i,r} = - \sum_{j=1}^l \left(\frac{e_{j,i,r}}{e_{i,r}} \right) \ln \left(\frac{e_{j,i,r}}{e_{i,r}} \right). \quad (6)$$

$C_{i,r}$ measures the competition in two-digit industry i in municipality r , $e_{j,i,r}$ denotes the number of employees in plant j , l the number of plants in two-digit industry i and municipality r , and $e_{i,r}$ as above. This entropy is a measure of competition on the output market since it measures to what extent other plants (or rather employees in other plants) producing similar products are present in the same municipality. To account for the different sizes of the plants, the number of employees, rather than the number of plants, is used in the calculation of competition. A similar transformation of the Herfindahl index to measure local competition at the industry level is done by MARTIN *et al.* (2011).

A well-functioning labour market is crucial for both plants and potential employees. It is self-evident that productivity will be higher if the right person is in the right job. As well as finding suitable employees in the first place, plants need to be able to replace them if the circumstances change. Hence, the abilities of the potential employees, i.e. the regional workforce, need to match the requirements of the regional plants. In addition, labour market matching gives rise to knowledge spillovers due to labour mobility playing an important

role in knowledge diffusion. How well labour market matching works at the plant level is assessed by measuring the concordance between the employees at a plant and the employees in the respective region regarding levels of education combined with types of skills. The skill categories are given by Table 1 and for education six levels are used. This means that for each plant and each region the share of employees for all twenty-four possible combinations of education levels and skills are calculated. To produce a single measure of labour market matching these are weighted together according to the following formula:

$$LM_j = \frac{1}{(\sum_{a=1}^{24} (se_a^{est} - se_a^r)^2)}, \quad (7)$$

in which LM_j gives the labour market matching value for plant j , se_a^r the combinations of education and skills at the municipal level and se_a^{est} the corresponding combinations of education and skills at the plant level. The interpretation of this measure is that the larger it is for a plant, the better that plant's employment needs match the regional labour pool. Hence, the higher is the probability that the right person is in the right place and the greater are the possibilities for labour mobility. This measure of labour market matching follows the same line of thinking as that of BALDWIN *et al.* (2008). A difference is that Equation 7 uses the plant-level labour composition rather than the industry level. The motivation behind this is that it is the plant and not the industry that employs workers. In addition, this measure incorporates both education and occupation, while BALDWIN *et al.* (2008) consider the latter only. Finally, the employment rate is introduced as an explanatory variable.

Variable overview

A summary of the explanatory variables is given by Table 2.

Table 2. Summary of explanatory variables.

Variable	Definition
Plant characteristics	
Capital	Material assets.
Labour	Number of employees.
Multi-firm	Dummy=1 if belonging to a multi-plant firm.
Maturity	Years since establishment.
Age	Average age of employees.
Female	Percentage of females.
Education	Percentage of employees with three or more years of university education.
Skills:	
Cognitive	Percentage of employees classified as cognitive skill workers.
Management	Percentage of employees classified as management and administration skill workers.
Social	Percentage of employees classified as social skill workers.
Motor	Base, percentage of employees classified as motor skill workers.
Industry:	
SCI 15-37	Dummy=1 if the plant belongs to one of industries 15-37, (one dummy for each industry, 15=base).
Regional characteristics	
Urbanization economies:	
Local accessibility	Access to municipal economic activity.
Intra-regional acc.	Access to regional economic activity.
Extra-regional acc.	Access to extra-regional economic activity.
Diversity	Distribution of the employees across industries at municipal level.
Industry diversity	Distribution of the employees across sub-industries at industry level.
Localization economies:	
Specialization	Municipal share of workers in a specific industry relative to the national share of workers in that industry.
Porter externalities:	
Competition	Distribution of the employees across plants at industry level.
Labour market:	
Labour matching	Concordance between plant and municipal workforce.
Employment rate	Municipal employment rate in per cent.

A correlation matrix for all variables is presented in Table B1, while Table D1 provides descriptive statistics.

MODEL AND METHOD

Since the dependent variable is productivity, a natural point of departure is a production function, given by Equation 8:

$$Y_j = F(K, AL) = K_j^\beta (AL_j)^\gamma, \quad (8)$$

in which Y_j denotes value added, K_j capital, A the efficiency of the employees and L_j the number of employees, all for plant j . However, since productivity is measured as value added per employee, Equation 8 is divided by L_j :

$$\frac{Y_j}{L_j} = y_j = \frac{K_j^\beta (A_j L_j)^\gamma}{L_j} = K_j^\beta A_j^\gamma L_j^{\gamma-1} = K_j^\beta A_j^\gamma L_j^\delta. \quad (9)$$

The variables described in the above section are contained in A_j since they are all factors that have the potential to affect average labour productivity. The above model is therefore extended by substituting A_j for these variables. To facilitate the empirical estimates Equation 9 is linearized by transformation into logarithmic form. Hence, the model used as a basis for the estimations of parameters is given by Equation 10. The t subscript is added since the data are longitudinal.

$$\ln y_{jt} = \beta \ln K_{jt} + \delta \ln L_{jt} + \gamma_1 \ln \Gamma_{jt} + \gamma_2 \ln Z_{irt} + \gamma_3 \ln H_{rt} + \gamma_4 D_t + \alpha_j + \varepsilon_{jt} \quad (10)$$

Γ_{jt} contains the plant characteristics, Z_{irt} the industry- and region-specific characteristics and H_{rt} the region-specific characteristics. D_t contains the industry and time dummies, of which the latter are introduced to control for year-specific effects. α_j captures unobserved time-invariant plant-specific effects and ε_{jt} is the usual error term, which by assumption is uncorrelated with the explanatory variables. An advantage of the double-log form of Equation 10 is that the estimated coefficients can be interpreted as elasticities.

As Equation 10 indicates, all explanatory variables are measured at time t .⁵ This raises the issue of reverse causality, i.e. that plant productivity is causing some of the right-hand-side variables. Since it is intuitively unlikely that the productivity of an individual plant has a significant effect on the region as a whole, unless the plant constitutes a major part of the regional economy, this mostly concerns the plant-specific variables. It may be argued that the productivity of the plant has an effect on the composition of the plant labour force as well as the capital stock. However, even though these variables do not change as slowly over time as the region-specific variables, there are substantial time lags due to hiring and firing processes, as well as investment processes. This decreases the problem of reverse causality, an issue that, due to the nature of the world, cannot be fully eliminated.

There are various methods for the estimation of Equation 10. Considering the panel structure of the data two obvious choices are fixed and random effects. The fixed-effects model applies within transformation, which eliminates the α_j 's. Hence, an advantage of the fixed-effects model is that it allows for endogeneity in terms of correlation between the unobserved plant-specific effects and the explanatory variables. In the random-effects model the α_j 's are treated as random variables, and consequently correlation between them and the additional explanatory variables implies biased and inconsistent estimates. For the present case, the Hausman test rejects the hypothesis of no correlation. However, the choice between fixed and random effects is not clear-cut since random-effects estimation has advantages as well. One

advantage is that random effects estimate the impact of explanatory variables with none or little within variation. Due to the within transformation, fixed effects do not work well for these types of variables. In the present case there are no completely time-invariant variables, but the regional characteristics have a tendency to change slowly over time, which implies low within variation when the period is rather short. Table E1 shows that the within variation is considerably lower than the overall and between variation for these variables. In addition, since α_j is in the composite error term ($\alpha_j + \varepsilon_{jt}$) in every period, the residuals are serially correlated. This is an issue especially for datasets with a large number of cases over relatively few periods. Random effects solve this by using feasible generalized least-squares estimators (FGLS) (RODRÍGUEZ-POSE and TSELIOS, 2012). Moreover, random-effects coefficients can be interpreted as long-run effects since the cross-sectional variation is retained, while fixed-effects coefficients can be interpreted as short-run or time-series effects (DURLAUF and QUAH, 1999; RODRÍGUEZ-POSE and TSELIOS, 2012). In other words, fixed effects models emphasize the within effect only, while random effects models also consider the between effect. This implies that random effects draw inferences based on more information, which increases the efficiency of the estimates.

Hence, there are advantages and disadvantages of both types of models. The discussion above shows that fixed and random effects are complements rather than substitutes, and therefore both models are estimated in the present case.⁶ Random effects allows for a generalization of the results beyond the specific sample used to draw the inferences. This implies that the results in the present paper can be discussed in a broader context, especially if Swedish plants are regarded as representative for European, or even Western world, plants.

As mentioned in the section ‘Variables’, the accessibility measures have been shown to capture the spatial dependence between locations. However, as a further control for spatial autocorrelation of the error terms across plants, the standard errors are clustered on 93

labour market regions. KEZDI (2004) shows that cluster-robust estimators are both unbiased and consistent in a fixed-effects setting. The clustering of the standard errors does not alter the results.

REGRESSION RESULTS

Table 3 presents the results from the estimations. Models 1-3 provide the results for the fixed-effects estimations (FE). Model 1 includes the variables describing plant characteristics only, while Model 2 introduces most of the regional characteristics. In Model 3 the accessibility measures are added as explanatory variables, to control for general urbanization economies as well as to model spatial interaction. Models 4-6 present the corresponding for the random-effects estimations (RE). Since Models 1-2 and Models 4-5 exclude the accessibility variables, they suffer from omitted variable bias. Models 3 and 6 can hence be regarded as the full models for the present case. However, considering the rather low R-squares, especially for the fixed-effects estimations, there are omitted variables in all six models. Considering endogenous growth models (e.g. ROMER (1986, 1990), LUCAS JR. (1988), and AGHION and HOWITT (1992)), missing variables concern technology and innovation. These are widely recognized as important drivers for productivity and growth. In the present case, technology and innovation are to some extent captured by industry belonging since certain industries are generally more high-tech and innovative than others. At the plant level, innovativeness is partly captured by the abilities variables since employees working with research and development are in general highly educated with cognitive skills. However, these are imperfect proxies and the models would benefit from direct plant-level measures of technology and innovation.

Table 3. Estimated impact of plant and regional characteristics. Dependent variable: average labour productivity.

	(1) FE	(2) FE	(3) FE	(4) RE	(5) RE	(6) RE
Plant characteristics:						
Capital	0.0705*** (0.0031)	0.0691*** (0.0030)	0.0692*** (0.0030)	0.0946*** (0.0030)	0.0921*** (0.0030)	0.0933*** (0.0027)
Labour	-0.2958*** (0.0126)	-0.3186*** (0.0130)	-0.3192*** (0.0131)	-0.2019*** (0.0048)	-0.2323*** (0.0054)	-0.2352*** (0.0055)
Multi-firm	-0.0212** (0.0098)	-0.0198* (0.0101)	-0.0200** (0.0101)	0.1271*** (0.0102)	0.1350*** (0.0106)	0.1340*** (0.0105)
Maturity	0.1523*** (0.0135)	0.1487*** (0.0128)	0.1484*** (0.0128)	0.1372*** (0.0085)	0.1320*** (0.0081)	0.1324*** (0.0083)
Age	-0.1447*** (0.0151)	-0.1224*** (0.0149)	-0.0122*** (0.0148)	-0.1762*** (0.0108)	-0.1521*** (0.0106)	-0.1496*** (0.0107)
Female	0.0037*** (0.0004)	0.0029*** (0.0004)	0.0029*** (0.0004)	0.0020*** (0.0004)	0.0013*** (0.0004)	0.0013*** (0.0004)
Education	0.0035*** (0.0004)	0.0030*** (0.0003)	0.0030*** (0.0003)	0.0053*** (0.0004)	0.0050*** (0.0004)	0.0047*** (0.0004)
Cognitive	0.0029*** (0.0003)	0.0026*** (0.0003)	0.0026*** (0.0003)	0.0043*** (0.0003)	0.0041*** (0.0002)	0.0041*** (0.0002)
Management	0.0034*** (0.0004)	0.0025*** (0.0003)	0.0024*** (0.0003)	0.0052*** (0.0005)	0.0044*** (0.0004)	0.0042*** (0.0004)
Social	0.0023*** (0.0003)	0.0017*** (0.0003)	0.0017*** (0.0003)	0.0034*** (0.0003)	0.0030*** (0.0002)	0.0029*** (0.0002)
Regional characteristics:						
Local accessibility			0.0216*** (0.0053)			0.0329*** (0.0040)
Intra-regional acc.			-0.0004 (0.0024)			0.0017* (0.0010)
Extra-regional acc.			-0.0001 (0.0126)			0.0120*** (0.0043)
Diversity		-0.0313* (0.0172)	-0.0492** (0.0188)		0.0193 (0.0126)	-0.0369** (0.0150)
Industry diversity		0.0042 (0.0121)	0.0057 (0.0119)		-0.0508*** (0.0161)	-0.0441*** (0.0149)
Specialization		0.0014 (0.0076)	0.0019 (0.0075)		0.0437*** (0.0099)	0.0439*** (0.0090)
Competition		0.0017 (0.0011)	0.0010 (0.0010)		0.0048*** (0.0011)	0.0017* (0.0009)
Labour matching		0.0593*** (0.0043)	0.0594*** (0.0044)		0.0614*** (0.0045)	0.0640*** (0.0048)
Employment rate		0.3057** (0.1400)	0.4484*** (0.1280)		-0.0143 (0.0943)	0.2488*** (0.0800)
Industry dummies	No	No	No	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	5.9224***	4.4979***	3.4268***	5.3937***	5.2614***	3.1296***
R-square ^a	0.0929	0.0977	0.0979	0.1706	0.1750	0.1821
Observations	205,087	205,087	205,087	205,087	205,087	205,087
Plants	39,730	39,730	39,730	39,730	39,730	39,730

Notes: Robust standard errors, clustered on 93 labour market regions, are given in parentheses.

***p < 0.01, **p < 0.05, and *p < 0.1.

^a within for FE, overall for RE.

Plant characteristics

The significance and sign of the coefficients for the plant characteristics are robust throughout all models, with the exception of the multi-firm dummy. Capital has, as expected, a positive impact on average labour productivity. Regarding labour, what is estimated is δ , which implies that γ is approximately equal to 0.7 using fixed effects and 0.8 using random effects ($\delta = \gamma - 1$). Hence, diminishing marginal productivity of labour is found. Additionally the age of the plant influences productivity positively, which is expected since the maturity of the plant measures experience. Also the positive coefficient for females is logical due to the self-selection of females with greater than average abilities in the manufacturing sector, which is male-dominated employment-wise. On the other hand, an increased average age of the employees has a negative impact on productivity. Regarding the multi-firm dummy, the results of the fixed-effects estimations in Models 1-3 show that belonging to a multi-plant firm has a negative impact on productivity. However, only the within variation is accounted for, which implies that the effect is for plants that change status during the period. Table E1 shows that the within variation is low for this variable, implying that this type of change does not frequently occur. When the cross-sectional variation is retained, shown by the random-effects estimations in Models 4-6, the effect of belonging to a multi-plant firm is positive, larger in absolute terms, and has a higher level of significance. Hence, even though the short-run effect is slightly negative, the long-run effect of being part of a larger firm is strongly positive. As already touched upon, this difference is probably due to that fixed effects accounts for intra-plant variation over time while random effects also includes inter-plant differences.

Regarding the variables describing the abilities of the employees, the percentage with at least three years of higher education influences productivity positively. Education measures human capital, a key input factor in production functions. In addition, an increased share of employees with cognitive, management or social skills has a positive effect, the largest

impact coming from employees with cognitive and management skills. This is logical since positions requiring cognitive and management skills are often filled with educated and/or experienced employees.

Regional characteristics

Regarding urbanization economies in general, both Models 3 and 6 show that the size of the local market influences productivity positively. Hence the results support the findings by CICCONE and HALL (1996) and CICCONE (2000), that density of economic activity increases average labour productivity. This result seems to be robust in more general terms since urbanization economies are here defined as accessibility rather than employment density and the analysis is conducted at the plant level, rather than at the regional level as in CICCONE and HALL (1996) and CICCONE (2000). Despite these differences, the result goes in the same direction. In addition, the effects of intra- and extra-regional accessibility are insignificant in the fixed-effects estimation and significantly smaller in the random-effects estimation. This implies that geographical proximity is important for externalities to occur. BOSCHMA (2005) discusses the importance of geographical proximity in conjunction with four other dimensions of proximity; cognitive, organizational, social and institutional. Geographical proximity facilitates face-to-face interaction, which increases learning and knowledge spillovers. According to BOSCHMA (2005), this is mostly due to the strengthening of the other dimensions of proximity. A somewhat contradicting result in Model 6 is that the effect of extra-regional accessibility both has a higher level of significance and is significantly larger in magnitude than intra-regional accessibility. However, it could be argued that while the effect of local accessibility concerns interactive face-to-face learning and spatially bounded knowledge spillovers, the effect of extra-regional accessibility concerns productivity benefits from having a potential market that is broader than the own region. This involves reducing the risk of the

problems of lock-in, which implies negative effects from a lack of openness and flexibility (BOSCHMA, 2005).

Models 3 and 6 show that region-wide industrial diversity significantly reduces average labour productivity, using both fixed and random effects. In addition, under random effects, diversity at the industry level has a negative impact on productivity. This implies that a long-run negative effect is found for plants belonging to more diverse industries. Cognitive proximity in terms of industry belonging is hence not enough to give rise to Jacobs externalities. This result is consistent with that of FRENKEN *et al.* (2007), who find that related variety has a negative effect on productivity. This is a robust result, since in the present case productivity is measured in levels at the plant level, while FRENKEN *et al.* (2007) measure productivity *growth* at the regional level. As for general urbanization economies, the results are consistent despite the methodological differences.

Regarding specialization and competition, the coefficients are insignificant in the fixed-effects estimations, which implies that no short-term effects are found. This could be a result of low within variation in these variables, as discussed above. However, when the cross-sectional variation is retained, the effect of specialization is highly significant, and robust to the inclusion of the accessibility variables. Specialization, in relative terms, is hence positively related to labour productivity in the long term. As discussed in the section ‘Variables’, the entropy measures for diversity are used as proxies for inverse absolute specialization. Hence the negative effects from diversity can be regarded as positive effects from both region-wide industrial specialization and within-industry specialization. This implies that all the results regarding specialization point to positive effects, and consequently evidence of MAR externalities is found. This is not unexpected since knowledge spillovers from localization economies occur between similar firms producing similar products, which implies that the innovations are more incremental in nature, affecting mostly productivity (FRENKEN *et al.*,

2007). In addition, previous empirical research on plant-level productivity points to positive effects from specialization (see section ‘Externalities’).

The opposite effects from diversity and specialization may seem intuitive at a first glance. However, the coexistence of diversity and specialization is possible due to these not being mutually exclusive, especially when measuring diversity in absolute terms and specialization in relative terms. A municipality can hence be highly diversified and at the same time incorporate a large part of a specific industry, making the municipality more specialized in that industry than the country average. In addition, a large region has greater possibilities to be both diverse and specialized. Diversity in such a region can imply many different specializations.

The result for competition is not as robust as for relative specialization, since when introducing the accessibility variables it is only weakly significant. There is also a significant decrease in the magnitude of the effect. However, the results still point to a positive impact from an increase in competition, which implies that evidence of Porter externalities is found. What should be kept in mind regarding competition is that it measures competition only at the local level. Many producers, especially in the manufacturing sector, are exposed to competition from regional, national as well as international actors.

The last group of externalities tested for concerns the labour market. In line with ERIKSSON and LINDGREN (2009) the results show that these externalities are important for productivity. The positive effect of an increased labour market matching proves to be the most robust result. There is no significant difference in the coefficient between the fixed- and random-effects estimates and no significant change from the inclusion of the accessibility variables. Hence, an increased match between the plant workforce and the regional workforce significantly enhances productivity, in both the short term and the long term. This result confirms the findings of BALDWIN *et al.* (2008), even though the applied measure is somewhat

different (see discussion in the section ‘Variables’). Hence the importance of labour market matching seems to be robust to changes in methodology. There are both direct and indirect channels for the positive effect of labour market matching on productivity. The direct effect originates from having the right person at the right place, while the indirect effect concerns knowledge spillovers from increased labour mobility. In addition, the employment rate is highly significant and influences productivity positively. The results thus support the theories by AKERLOF and YELLEN (1990) and DARITY and GOLDSMITH (1996) regarding the behaviour and psychology of employees. In addition, a higher employment rate reflects that the plants in the region are doing well, or that the labour market matching works well.

CONCLUSIONS

This paper has analysed the impact of a broader than usual range of spatial externalities on average labour productivity in Swedish manufacturing plants. Short-term as well as long-term effects of four sources of externalities are tested for, urbanization economies with diversity in particular, specialization, competition and labour market matching. Positive effects from general urbanization economies, measured as access to the local market, and labour market matching, measured as the similarity between the plant’s workforce and the regional workforce, are found to be the most robust results. In addition, the results show a negative long-term effect of diversity at the industry level. These effects point in the same direction as previous studies, despite methodological differences, which implies a generalizability of the results. This is the main contribution of the present study, especially considering the general inconclusiveness of previous research on spatial externalities. There are no short-term effects of diversity at the industry level, neither from specialization nor competition, which can be explained by low within variation. On the other hand, long-term positive effects are found for both specialization and competition, showing evidence of both MAR and Porter externalities. Regarding Jacobs

externalities, it has been argued that cognitive proximity is required in order for knowledge spillovers to occur. The results show that belonging to the same industry does not imply enough cognitive proximity to spawn productivity-enhancing Jacobs externalities. An interesting area for further research is thus the untangling of what bodes for cognitive proximity, if not industry belonging.

Finally, even though the results show that regional characteristics have significant effects on productivity, adding these variables does not add much to the explanatory power of the models. Hence, to explain average labour productivity at the plant level, the characteristics of the individual plants are what matters most. This gives weight to the issue of ecological fallacy, since the plant-level variation is not retained with aggregated data. Hence, the importance of plant-specific characteristics, including the characteristics of the employees, should not be forgotten. Externality effects are commonly associated with knowledge spillovers and, since knowledge is embedded in individuals, a further focus on the characteristics of the employees is motivated.

Acknowledgements - Financial support from the Swedish Research Council Formas (grant 2009-1192) is gratefully acknowledged. This paper has greatly benefited from comments received at the Uddevalla Symposium 2011 in Bergamo, as well as at internal seminars at Jönköping International Business School (JIBS). The author would like to acknowledge associate professor Johan Klaesson at JIBS for the measure of labour market matching (Equation 7). Finally, the author would like to thank two anonymous referees for providing valuable comments and suggestions for improvements.

NOTES

1. See BRESCHI and LISSONI (2001) for a discussion on the different channels for externality effects.
2. See WORLD BANK (2009), p.128, for a division of localization and urbanization economies into subgroups.
3. The same procedure is used for distributing physical capital.
4. For some firms the reported capital level is zero. This is likely due to missing information rather than a non-existent capital stock. The plants belonging to these firms are excluded.
5. This makes sense especially for the plant-specific variables since output year t is produced by year t inputs. Regarding the regional variables, empirical estimations show that the results are robust when lagging them one year.
6. In addition, pooled ordinary least-squares (OLS) estimations are run. However, since a Breusch-Pagan Lagrange Multiplier test rejects the hypothesis of no differences between plants, random effects is preferred to pooled OLS.

REFERENCES

- AGHION P. and HOWITT P. (1992) A model of growth through creative destruction, *Econometrica* 60(2), 323-351.
- AKERLOF G. A. and YELLEN J. L. (1990) The fair wage-effort hypothesis and unemployment, *Quarterly Journal of Economics* 105(2), 255-283.
- ANDERSSON F., BURGESS S. and LANE J. I. (2007) Cities, matching and the productivity gains of agglomeration, *Journal of Urban Economics* 61(1), 112-128.
- ANDERSSON M. and GRÅSJÖ U. (2009) Spatial dependence and the representation of space in empirical models, *Annals of Regional Science* 43(1), 159-180.
- ANDERSSON M. and KLAESSON J. (2009) Regional interaction and economic diversity: exploring the role of geographically overlapping markets for a municipality's diversity in retail and durables, in KARLSSON C., JOHANSSON B. and STOUGH R. R. (Eds.) *Innovation, Agglomeration and Regional Competition*, pp. 19-37. Edward Elgar, Cheltenham.
- ANDERSSON M., KLAESSON J. and LARSSON J. P. (2012). How local are spatial density externalities? Evidence from square grid data, *CIRCLE Paper no. 2010/10*.
- ARROW K. J. (1962) The economic implications of learning by doing, *Review of Economic Studies* 29(3), 155-173.
- ATTARAN M. (1986) Industrial diversity and economic performance in U.S. areas, *Annals of Regional Science* 20(2), 44-54.
- BACOLOD M., BLUM B. S. and STRANGE W. C. (2009) Skills in the city, *Journal of Urban Economics* 65(2), 136-153.
- BALDWIN J. R., BECKSTEAD D., BROWN W. M. and RIGBY D. L. (2008) Agglomeration and the geography of localization economies in Canada, *Regional Studies* 42(1), 117-132.
- BEAUDRY C. and SCHIFFAUEROVA A. (2009) Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy* 38(2), 318-337.

- BELOGEY N., LECAT R. and MAURY. T-P. (2006) Determinants of productivity per employee: An empirical estimation using panel data, *Economics Letters* 91(2), 153-157.
- BISHOP P. and GRIPAIO P. (2010) Spatial externalities, relatedness and sector employment growth in Great Britain, *Regional Studies* 44(4), 443-454.
- BOSCHMA R. (2005) Proximity and innovation: A critical assessment, *Regional Studies* 39(1), 61-74.
- BOSCHMA R., ERIKSSON R. and LINDGREN U. (2009) How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity, *Journal of Economic Geography* 9(2), 169-190.
- BOSCHMA R., MINONDO A. and NAVARRO M. (2012) Related variety and regional growth in Spain, *Papers in Regional Science* 91(2), 241-256.
- BRESCHI S. and LISSONI F. (2001) Knowledge spillovers and local innovation systems: A critical survey, *Industrial and Corporate Change* 10(4), 975-1005.
- CAPELLO R. (2002) Entrepreneurship and spatial externalities: Theory and measurement, *Annals of Regional Science* 36(3), 387-402.
- CICCONE A. (2000) Agglomeration effects in Europe, *European Economic Review* 46(2), 213-227.
- CICCONE A. and HALL R. E. (1996) Productivity and the density of economic activity, *American Economic Review* 86(1), 54-70.
- COHEN W. M. and LEVINTHAL D. A. (1990) Absorptive capacity: A new perspective on learning and innovation, *Administrative Science Quarterly* 35(1), 128-152.
- COMBES P.-P. and DURANTON G. (2006) Labour pooling, labour poaching and spatial clustering, *Regional Science and Urban Economics* 36(1), 1-28.
- DARITY W. A. and GOLDSMITH A. H. (1996) Social psychology, unemployment and macroeconomics, *Journal of Economic Perspectives* 10(1), 121-140.

- DE LUCIO J. J., HERCE J. A. and GOICOLEA A. (2002) The effects of externalities on productivity growth in Spanish industry, *Regional Science and Urban Economics* 32(2), 241-258.
- DURANTON G. and PUGA D. (2004) Micro-foundations of urban agglomeration economies, in HENDERSON J. V. and THISSE J.-F. (Eds.) *Handbook of Regional and Urban Economics*, Vol. 4, pp. 2063-2117. Elsevier, Amsterdam.
- DURLAUF S. N. and QUAH D. T. (1999) The new empirics of economic growth, in TAYLOR J. B. and WOODFORD M. (Eds.) *Handbook of Macroeconomics*, Vol. 1, pp. 235-308. Elsevier, Amsterdam.
- ERIKSSON R. and LINDGREN U. (2009) Localized mobility clusters: Impacts of labour market externalities on firm performance, *Journal of Economic Geography* 9(1), 33-53.
- FELDMAN M. P. and AUDRETSCH D. B. (1999) Innovation in cities: Science-based diversity, specialization and localized competition, *European Economic Review* 43(2), 409-429.
- FRENKEN K., VAN OORT F. and VERBURG T. (2007) Related variety, unrelated variety and regional economic growth, *Regional Studies* 41(5), 685-697.
- GERTLER M. S. (2003) Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there), *Journal of Economic Geography* 3(1), 75-99.
- GLAESER E. L., KALLAL H. D., SCHEINKMAN J. A. and SHLEIFER A. (1992) Growth in cities, *Journal of Political Economy* 100(6), 1126-1152.
- HENDERSON V., KUNCORO A. and TURNER M. (1995) Industrial development in cities, *Journal of Political Economy* 103(5), 1067-1090.
- JACOBS J. (1969) *The Economy of Cities*. Vintage, New York, NY.
- JACQUEMIN A. P. and BERRY C. H. (1979) Entropy measure of diversification and corporate growth, *Journal of Industrial Economics* 27(4), 359-369.

- JOHANSSON B. and KLAESSON J. (2011) Creative milieus in the Stockholm region, in ANDERSSON D. E., ANDERSSON, Å. E. and MELLANDER C. (Eds.) *Handbook of Creative Cities*, pp. 456-481, Edward Elgar, Cheltenham.
- JOHANSSON B., KLAESSON J. and OLSSON M. (2002) Time distances and labor market integration, *Papers in Regional Science* 81(3), 305-327.
- JOHANSSON B., KLAESSON J. and OLSSON M. (2003) Commuters' non-linear response to time distances, *Journal of Geographical Systems* 5(3), 315-329.
- KARLSSON C. and PETTERSSON L. (2005) Regional productivity and accessibility to knowledge and dense markets, *CESIS Electronic Working Paper Series No. 32*.
- KÉZDI G. (2004) Robust standard error estimation in fixed-effects panel models, *Hungarian Statistical Review Special*(9), 96-116.
- LUCAS JR. R. E. (1988) On the mechanics of economic development, *Journal of Monetary Economics* 22(1), 3-42.
- MARSHALL A. (1890) *Principles of Economics*. Macmillan, London.
- MARTIN P., MAYER T. and MAYNERIS F. (2011) Spatial concentration and plant-level productivity in France, *Journal of Urban Economics* 69(2), 182-195.
- MASKELL P. (2001) Towards a knowledge-based theory of the geographical cluster, *Industrial and Corporate Change* 10(4), 921-943.
- MASKELL P. and MALMBERG A. (1999) Localised learning and industrial competitiveness, *Cambridge Journal of Economics* 23(2), 167-185.
- PALAN N. (2010). Measurement of specialization - The choice of indices, *FIW Working Paper No. 62*.
- PORTER M. E. (1998) *The Competitive Advantage of Nations*. Free Press, London.

- POWER D. and LUNDMARK M. (2004) Working through knowledge pools: Labour market dynamics, the transference of knowledge and ideas, and industrial clusters, *Urban Studies* 41(5/6), 1025-1044.
- RODRÍGUEZ-POSE A. and TSELIOS V. (2012) Individual earnings and educational externalities in the European Union, *Regional Studies* 46(1), 39-57.
- RODRÍGUEZ-POSE A., TSELIOS V., WINKLER D. and FAROLE T. (2013) Geography and the determinants of firm exports in Indonesia, *World Development* 44, 225-240.
- ROMER P. M. (1986) Increasing returns and long-run growth, *Journal of Political Economy* 94(5), 1002-1037.
- ROMER P. M. (1990) Endogenous technological change, *Journal of Political Economy* 98(5, Part 2), S71-S102.
- ROSENTHAL S. S. and STRANGE W. C. (2008) The attenuation of human capital spillovers, *Journal of Urban Economics* 64(2), 373-389.
- WORLD BANK. (2009) *World development report 2009: reshaping economic geography*. The International Bank for Reconstruction and Development, The World Bank, Washington D.C.
- WORLD ECONOMIC FORUM (2010) Klaus Schwab (Ed.) *The global competitiveness report 2010-2011*. The World Economic Forum, Geneva.

APPENDIX A

Table A1. Description of the industries included.

SIC code ^a	Description ^a	Number of plants ^b	Number of employees ^b	Average firm size ^b
15	Manufacture of food products and beverages	16,599	466,567	28.11
16	Manufacture of tobacco products	78	7,216	92.51
17	Manufacture of textiles	4,275	53,819	12.59
18	Manufacture of wearing apparel	1,447	11,851	8.19
19	Tanning and dressing of leather; manufacture of luggage and footwear	728	8,240	11.32
20	Manufacture of wood and products of wood and cork	18,428	287,128	15.58
21	Manufacture of pulp, paper and paper products	2,943	311,050	105.69
22	Publishing, printing and reproduction of recorded media	25,081	304,360	12.14
23	Manufacture of coke, refined petroleum products and nuclear fuel	342	24,552	71.79
24	Manufacture of chemicals and chemical products	4,551	301,796	66.31
25	Manufacture of rubber and plastic products	8,388	184,639	22.01
26	Manufacture of non-metallic mineral products	7,093	151,384	21.34
27	Manufacture of basic metals	2,768	309,059	111.65
28	Manufacture of fabricated metal products	45,351	579,928	12.79
29	Manufacture of machinery and equipment n.e.c.	24,843	772,612	31.10
30	Manufacture of office machinery	970	24,527	25.29
31	Manufacture of electrical machinery	6,255	197,277	31.54
32	Manufacture of radio, television and communication equipment	2,715	214,432	78.98
33	Manufacture of medical, precision and optical instruments, watches and clocks	8,794	189,290	21.52
34	Manufacture of motor vehicles and trailers	4,532	617,930	136.35
35	Manufacture of other transport equipment	4,855	166,042	34.20
36	Manufacture of furniture	12,261	294,742	24.04
37	Recycling	1,790	18,539	10.36

Notes: ^a Statistics Sweden, www.scb.se.

^b Own calculation.

Appendix B - Table B1. Correlation matrix for all variables, log transformed.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. Productivity	1																				
2. Capital	.32	1																			
3. Labour	.15	.73	1																		
4. Multifirm	.14	.27	.30	1																	
5. Maturity	.14	.26	.36	.00	1																
6. Age	-.05	-.14	-.12	.02	.16	1															
7. Female	.06	.31	.46	.10	.23	-.05	1														
8. Education	.16	.34	.49	.21	.11	-.06	.33	1													
9. Cognitive	.12	.23	.37	.17	.10	.01	.24	.40	1												
10. Management	.13	.35	.46	-.04	.24	.01	.45	.29	.11	1											
11. Social	.12	.28	.40	.22	.10	-.11	.34	.31	.18	.15	1										
12. Motor	.00	.30	.34	-.09	.21	-.13	.04	-.06	-.23	.02	-.06	1									
13. Diversity	.01	-.07	-.02	.02	-.01	-.00	.01	.04	.04	.01	.05	-.08	1								
14. Industry diversity	.04	.19	.23	-.01	.07	-.03	.05	.09	.05	.10	.05	.12	.13	1							
15. Specialization	.06	.21	.25	.03	.06	-.02	.12	.14	.10	.12	.08	.06	.11	.79	1						
16. Competition	.01	-.05	-.04	-.05	-.01	-.01	-.04	-.01	-.01	-.01	-.02	-.02	.20	.43	.23	1					
17. Labour matching	.15	.61	.76	.12	.33	-.15	.36	.31	.19	.39	.27	.56	-.08	.23	.24	-.05	1				
18. Employment rate	.01	.06	-.00	-.07	.04	-.01	.01	-.06	-.05	.01	-.04	.09	-.14	.05	.07	-.10	.08	1			
19. Local accessibility	.05	-.15	-.03	.06	-.05	-.01	.01	.13	.10	.00	.09	-.21	.42	-.06	-.06	.31	-.19	-.44	1		
20. Intra-regional acc.	.03	-.05	-.01	.00	.00	.01	.03	.07	.05	.02	.04	-.08	.00	-.05	-.04	.06	-.08	.04	.21	1	
21. Extra-regional acc.	.05	.04	.05	-.00	.04	.02	.04	.03	.02	.05	.02	.02	.02	-.02	-.03	.06	.02	.00	.07	.33	1

APPENDIX C

Table C1. VIF-values.

Variable	VIF
Plant characteristics	
Capital	2.52
Labour	5.09
Multi-firm	1.25
Maturity	1.29
Age	1.11
Female	1.51
Education	1.54
Cognitive	1.53
Management	1.61
Social	1.38
Motor	2.17
Regional characteristics	
Diversity	1.28
Industry diversity	3.39
Specialization	2.82
Competition	1.52
Labour matching	3.53
Employment rate	1.28
Local accessibility	1.95
Intra-regional accessibility	1.22
Extra-regional accessibility	1.14

APPENDIX D

Table D1. Descriptive statistics.

Variable	N	Mean	Std. Dev.	Min	Max
Plant characteristics					
Value added per employee	205,087	551.88	873.94	0.33	93,313
Capital	205,087	14.20M	164.68M	0.50	14,879M
Labour	205,087	26.80	148.12	1	12,146
Maturity	205,087	13.68	7.34	1	25
Age	205,087	44.29	7.85	17	84
Female	205,087	22.72	26.43	0	100
Education	205,087	7.03	17.19	0	100
Cognitive	205,087	14.43	27.32	0	100
Management	205,087	21.58	25.93	0	100
Social	205,087	10.34	21.50	0	100
Motor	205,087	53.64	36.72	0	100
Regional characteristics					
Diversity	205,087	1.96	0.36	0.38	2.58
Industry diversity	205,087	0.19	0.11	0.00	0.37
Specialization	205,087	1.95	3.91	0.00	211.04
Competition	205,087	1.88	1.12	0	5.11
Labour matching	205,087	15.28	110.99	0.68	21,615.84
Employment rate	205,087	76.79	4.54	54.2	87.3
Local accessibility	205,087	16,018M	32,000M	124.16M	147,466M
Intra-regional accessibility	205,087	14,971M	24,625M	0	157,445M
Extra-regional accessibility	205,087	2,380M	1,898M	87,785	13,506M

APPENDIX E

Table E1. Descriptive statistics for the log transformed data.

Variable		Mean	Std. Dev.	Min	Max	Observations
Value added per employee	overall	6.113845	0.6227951	-1.098612	11.44371	N = 205087
	between		0.6680039	-1.098612	10.63321	n = 39730
	within		0.336661	-0.2803134	10.69341	T = 5.16202
Capital	overall	13.43024	2.327726	-0.6931472	23.42319	N = 205087
	between		2.252686	-0.1386294	22.84318	n = 39730
	within		0.6490242	1.670063	20.29626	T = 5.16202
Labour	overall	1.903533	1.393575	-3.78419	9.404763	N = 205087
	between		1.338902	-1.702049	8.949182	n = 39730
	within		0.2862679	-2.962619	6.139106	T = 5.16202
Multi-firm (dummy)	overall	0.1834441	0.3870311	0	1	N = 205087
	between		0.373569	0	1	n = 39730
	within		0.0980208	-0.7054448	1.072333	T = 5.16202
Maturity	overall	2.364096	0.8446833	0	3.218876	N = 205087
	between		1.031361	0	3.218876	n = 39730
	within		0.2683722	0.7912293	3.597863	T = 5.16202
Age	overall	3.774723	0.1802539	2.833213	4.430817	N = 205087
	between		0.1988873	2.833213	4.424829	n = 39730
	within		0.0746355	2.86664	4.502339	T = 5.16202
Female ^a	overall	-1.50038	6.415983	-10	4.60517	N = 205087
	between		6.249888	-10	4.60517	n = 39730
	within		2.670664	-14.48275	11.48199	T = 5.16202
Education ^a	overall	-5.980937	5.884261	-10	4.60517	N = 205087
	between		5.34163	-10	4.60517	n = 39730
	within		2.538659	-18.73943	7.001436	T = 5.16202
Cognitive ^a	overall	-4.832793	6.433809	-10	4.60517	N = 205087
	between		5.900947	-10	4.60517	n = 39730
	within		2.807902	-17.81517	8.149581	T = 5.16202
Management ^a	overall	-1.261822	6.274989	-10	4.60517	N = 205087
	between		6.094841	-10	4.60517	n = 39730
	within		2.927499	-14.2442	11.72055	T = 5.16202
Social ^a	overall	-4.994108	6.253453	-10	4.60517	N = 205087
	between		5.547373	-10	4.60517	n = 39730
	within		3.26085	-17.97648	7.988265	T = 5.16202
Motor ^a	overall	.9149559	5.940511	-10	4.60517	N = 205087
	between		6.101084	-10	4.60517	n = 39730
	within		2.422313	-12.06742	13.89733	T = 5.16202
Diversity	overall	0.6524614	0.2223098	-0.9636526	0.9487484	N = 205087
	between		0.2160082	-0.8445471	0.9429304	n = 39730
	within		0.0554496	-0.5630672	1.808111	T = 5.16202
Industry diversity	overall	-1.908903	0.8839875	-7.510163	-1	N = 205087
	between		0.9036973	-7.150859	-1	n = 39730
	within		0.2434251	-6.142968	1.601159	T = 5.16202
Specialization	overall	-0.0090473	1.195688	-6.459445	5.352043	N = 205087
	between		1.184613	-6.232182	5.244453	n = 39730
	within		0.3383616	-5.473134	5.293833	T = 5.16202
Competition ^a	overall	-0.011839	2.349464	-10	1.631899	N = 205087
	between		2.203234	-10	1.631899	n = 39730
	within		0.9455777	-9.857934	9.970153	T = 5.16202
Labour matching	overall	1.725863	1.245849	-.3916223	9.981182	N = 205087
	between		1.151309	-.3854015	9.528898	n = 39730
	within		0.4203101	-2.053578	5.126214	T = 5.16202
Employment rate	overall	4.339286	0.0607175	3.992681	4.46935	N = 205087
	between		0.0583784	3.992681	4.46935	n = 39730
	within		0.0184815	4.081994	4.596549	T = 5.16202

Local accessibility	overall	22.21762	1.514171	18.63704	25.71686	N = 205087
	between		1.535648	18.66547	25.71686	n = 39730
	within		.2544272	18.2973	26.39902	T = 5.16202
Intra-regional accessibility	overall	20.91533	5.663244	0	25.78234	N = 205087
	between		5.692286	0	25.78234	n = 39730
	within		0.4560857	0.0971277	42.45578	T = 5.16202
Extra-regional accessibility	overall	21.21946	1.150401	11.38264	23.32637	N = 205087
	between		1.151268	11.38264	23.32637	n = 39730
	within		0.1256452	14.85235	28.32454	T = 5.16202

Notes: ^aOverall and between minimum is -10 due to missing values (the zeros in the non-transformed dataset) being replaced by this. Various values are tested without any changes in the results.