CESIS Centre of Excellence for Science and Innovation Studies

CESIS Electronic Working Paper Series

Paper No. 452

An empirical study of firms' absorptive capacity and export diversification

Tina Wallin

April, 2017

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

An empirical study of firms' absorptive capacity and export diversification

Tina Wallin

Centre for Entrepreneurship and Spatial Economics (CEnSE), Jönköping International Business School, Jönköping, Sweden

E-mail: tina.wallin@ju.se

Abstract: The purpose of this paper is to study the firms' internal knowledge in combination with the external knowledge diversity in their region to examine their joint relation to export diversification. Using a data set of the full population of Swedish manufacturing exporters for the period 2003-2013, allows for identifying when firms introduce new products on the export market. The results indicate that firms in the medium-high tech and the medium-low tech manufacturing sectors only benefit from a larger external knowledge diversity if they themselves have some internal knowledge increasing their absorptive capacity. Changing spatial scale or increasing the time lag yields mostly the same results, but extending the external knowledge diversity to include all types of education subjects does not. This further supports the suggested importance of an absorptive capacity to facilitate the acquisition, assimilation and usage of related external knowledge in producing new products.

Keywords: new product; export diversification; absorptive capacity; related knowledge

JEL codes: C33, D22, D83, F14, J24, O31

Introduction

Economic growth is central to any economy and the economic literature has long discussed the role of diversity in economic growth, mainly by increasing the innovative behaviour of firms, see e.g. Jacobs (1969). Evolutionary theorists, such as Nelson and Winter (1982), explain these innovations as being new combinations of old knowledge. Hence, interaction between people with different knowledge should facilitate the creation of new combinations of knowledge because they can exchange information and learn from each other. However, people with too similar knowledge bases do not learn anything new from such interaction (Nooteboom 2000a, Boschma 2004). Cohen and Levinthal (1990) argue in a similar fashion that people's absorptive capacity, i.e. ability to understand and make use of information, decreases as the knowledge bases become too different as the ability to understand each other decreases. The best situation is when the knowledge bases are partially overlapping. Then there is a basic understanding, but still new knowledge to be shared.

Not all relevant knowledge for innovating firms is found within the firms themselves. A large field of literature emphasises that knowledge spillovers from external sources generate positive externalities for firms and are important for the economy at large (see e.g. Glaeser et al. (1992), Glaeser (1999), Jacobs (1969)) Consequently the milieu surrounding firms is also a potential knowledge source for innovative activities. However, as highlighted by e.g. Nooteboom (2000a) the possibility of using this external knowledge is dependent on what absorptive capacity the firms already have. Meaning that a firm with access to external knowledge not necessarily benefits from it – it depends on their possibility to understand and utilise that knowledge.

Many studies have been done examining the role of different aspects of diversity on innovations, either at regional level (Niebuhr 2010, Tavassoli and Carbonara 2014) or at the firm level (Audretsch and Feldman 1999, Sampson 2007, Antonietti and Cainelli 2011), and some have even focused particularly on education (Wixe 2016, Parrotta, Pozzoli, and Pytlikova 2012, Østergaard, Timmermans, and Kristinsson 2011). However, what these studies have missed is that the internal knowledge may serve as a gatekeeper to the external knowledge, which has been highlighted in other studies examining slightly different issues (refs!!). I.e. depending on your internal competence your absorptive capacity may or may not be enough to fruitfully assimilate the external knowledge into your firm. Therefore, the purpose of this paper is to study firms' internal knowledge in combination with the external knowledge diversity in their region, and examine their joint relation to firms' ability to do diversify their export portfolio.

The contribution of this study is threefold. Firstly, the measure of new products introduced on the export market could apart from being a measure of export diversification also be considered a proxy for innovative behaviour. As such the measure is based on actual commercialised products instead of other commonly used proxies such as R&D expenditures or patents. This is more in line with the definition of innovation as suggested by e.g. the Oslo Manual (OECD/Eurostat 2005). Secondly, the main interest is the interaction between the internal and external knowledge bases based on the individuals' education. If the absorptive capacity of a firm determines the possibility to interpret external knowledge and utilise it in the internal production, their multiplicative effect is relevant and not only their separate contributions. Thirdly, the external knowledge diversity is measured in terms of related diversity in technologically related subjects instead of overall diversity. As the maximum value of the measure is an even distribution of related knowledge it reduces the problem with too much diversity being disadvantageous for the firms as highlighted by e.g. Boschma (2004).

A detailed data set with the full population of Swedish exporters is used to examine this topic and fulfil the purpose. For each firm, it is possible to identify which goods they have exported. Meaning that over time one can identify when new products are introduced and this is used as the dependent variable of this study. To ensure that new products are new and not recoded old ones, the codes that changes over the years have been recoded. The panel covers the period 2003-2013 and a negative binomial model is estimated with maximum likelihood for four groups of manufacturing firms, according to Eurostat's division based on technology intensity. As the firms in the data set are operating in the manufacturing sector the focus is on diversity of technological knowledge bases in the regions where the firms are located.

The results indicate that firms within the medium-high tech and the medium-low tech groups only benefit from the external knowledge diversity in their geographical proximity if they themselves have a certain level of internal knowledge. This, provides empirical support for the importance of having an absorptive capacity internally for fruitful knowledge exchanges to occur. With minor alterations, these results hold when changing the regional level to labour market regions instead of smaller municipalities, and when increasing the lag of the independent variable beyond one year, but not when expanding the external knowledge measure to include all subjects.

The rest of the paper is organised as follows: the next section presents the theoretical framework, previous studies, and the hypothesis of this study. Section 3 presents the data, variables and model together with descriptive statistics. Section 4 presents the empirical results and robustness checks. Finally, section 5 concludes.

Knowledge, absorptive capacity, and cognitive proximity

'Knowledge is critical to the process of innovation. Whether the knowledge is internally generated or externally acquired, what an organization knows determines what it can do.' (Thornhill 2006, 691). Important to remember is that knowledge is embodied in people – not in the firms themselves. There are several ways for the firms to acquire knowledge and all of them have the individuals as bearers of the knowledge: firstly, firms can hire people with suitable knowledge; secondly, they can cooperate with such people in other firms, universities, or research institutes; thirdly, they can take advantage of external knowledge through spillovers.

The first refer to the knowledge capacity that is internal to the firms, as the employees are hired to use their knowledge there. The remaining two are referring to external sources of knowledge, where cooperation projects are targeted towards specific actors and the spillovers are general positive externalities spilling over from your surrounding environment.

Many studies discuss agglomeration economies and how one can benefit from knowledge spillovers in the local environment (Glaeser et al. 1992, Glaeser 1999, Teece 1992, Rosenthal and Strange 2001, Gertler 2003, Malmberg and Maskell 2002). Which type of knowledge that is relevant for the spillovers to increase performance in other firms is another question, either specialised knowledge or diverse knowledge. The former refers to the Marshall-Arrow-Romer (MAR) externalities, and Porter's externalities, whereas the latter refers to Jacobs externalities. The MAR externalities, and the Porter externalities, highlight the importance of same industry concentration for productive knowledge spillovers (Marshall 1890, Arrow 1962, Romer 1986, Porter 1990). Jacobs on the other hand believed that the most productive knowledge spillovers were those coming from other industries – hence she favoured a diverse industry structure (Jacobs 1969).

Plenty of studies have attempted to examine which of these are the most important and the results are mixed. Some favour diversity (c.f. Desrochers and Leppälä (2011), Audretsch and Feldman (1999), and Usai and Paci (2003)) and some specialisation (c.f. Baldwin et al. (2008), and Wixe (2015)). This lack of consensus matches the general picture described by e.g. Duranton and Puga (2000), Beaudry and Schiffauerova (2009), and De Groot, Poot, and Smit (2009).

An alternative is that the beneficial knowledge is neither completely different nor very similar but something in-between. Nelson and Winter (1982) suggested that the distance between the knowledge already embedded in the firm and the knowledge required to carry out a particular innovation determine whether it will be discovered or not by that firm. Hence, products that depends on knowledge that is close to the firm's current knowledge are more likely to be discovered. This was empirically confirmed by Breschi, Lissoni, and Malerba (2003) who found that firms tend to patent in technological fields that are related to those in which they have already patented. This implicitly narrows the range of knowledge that is relevant for the development of new products to *related* knowledge. This importance of related diversity has been widely supported in the literature as well (Frenken, Van Oort, and Verburg 2007, Bishop and Gripaios 2010, Wixe and Andersson 2016, Castaldi, Frenken, and Los 2015).

The reasoning behind the relatedness is that some overlap of the knowledge is required for it to be successfully transmitted between individuals. As Nooteboom (2000a) points out, if the knowledge bases are too different the individuals will not be able to understand each other – but if they are too similar nothing new will be learnt. However, too much cognitive proximity can have negative lock-in effects in the long run (Boschma 2005). Hence the cognitive distance can be thought of as an inverse u-shaped function, where the maximum is found approximately in the middle (Wuyts et al. 2005).

A similar argument has been put forward by (Cohen and Levinthal 1990, 1989) who argue that a firm's capability to understand and make use of information, i.e. its absorptive capacity, depends on the prior knowledge held by the firm's employees. Implying that their absorptive capacity also depends on previous experiences and differences in background characteristics. A firm that does not recognise this and invests in its absorptive capacity may risk not detecting profitable innovations and technologies in its surroundings (Cohen and Levinthal 1994).

Other dimensions of proximity between actors, which would increase the absorptive capacity, have been highlighted in the literature. E.g. Boschma (2005) discusses five different types, Kirat and Lung (1999) discuss three types, and Rallet and Torre (1999) discuss two types. What they have in common is the emphasis on geographical and organisational proximity, which Torre and Rallet (2005) mean partially work as substitutes. If you have an organisational proximity you don't need a geographical one, but the opposite is not true. This has been empirically shown by Fitjar and Rodríguez-Pose (2017) who show that most activities resulting in Norwegian innovations are purpose-built with particular partners and not occurring due to simply being close to others.

Previous studies and hypothesis

Many empirical studies have examined the relation between diversity, either related or unrelated, and innovative behaviour. Some have focused on characteristics of the region, such as: diversity of patents on state-level (Castaldi, Frenken, and Los 2015), which was found to be positive for innovation; and diversity of the industrial composition in functional regions, which was also found to be positive for innovation (Tavassoli and Carbonara 2014). Niebuhr (2010) examined cultural diversity, measured as the unrelated variety of country of origin, on NUTS 3 regions in Germany. She found that the cultural diversity in the regions had a positive effect of on the region's probability to patent.

Others have used the level of the firm for similar studies. E.g. Sampson (2007) who measured diversity by classification of patents, and van Beers and Zand (2014) who used a Herfindahl index based on different actors and geographical areas. Industrial diversity is another aspect covered by e.g. Audretsch and Feldman (1999), (Antonietti and Cainelli 2011), and Paci and Usai (1999).

Among studies that do consider individuals, a common aspect of diversity has been the cultural dimension in terms of the country of birth. E.g. a study by Qian (2013), which looks both at the effects on entrepreneurship and innovations. Diversity is measured by a Herfindahl index of the employees' country of birth and the results indicate a positive effect on entrepreneurship but not on innovations. Ozgen, Nijkamp, and Poot (2013) utilise Dutch firm-level data in combination with the Community Innovation Survey (CIS) and find that firms with larger diversity in their employees' country of birth are more prone to innovate. Nathan and Lee (2013) makes another take the topic by examining diversity of firms’ management teams, which has a positive effect on the propensity to introduce product innovations.

Some studies have included diversity of education in their studies of the relationship between innovation and diversity. Parrotta, Pozzoli, and Pytlikova (2012) measure educational diversity with six education categories, mixing subject and level. This measure of diversity has no effect on innovation output. A similar study made by Østergaard, Timmermans, and Kristinsson (2011) make use of a matched employer-employee data set together with CIS data and find that an internal workforce with a diverse educational background increases the probability to innovate. Wixe (2016) uses Swedish CIS data to examine firm innovativeness and related neighbourhood diversity in terms of both industries and education. She finds that education matters more in rural areas and industries more in urban areas.

Hence, it seems that most previous studies in this area seem to focus on the internal knowledge, or the external knowledge, but not the combination of them. Although that should be important considering that a cognitive proximity is needed for firms to be able to absorb the knowledge flows outside the firm.

Some studies do take this interaction into account, but examines slightly different things than those aforementioned. One example is Johansson, Johansson, and Wallin (2015) where it is concluded that the conjunction of internal and external knowledge matters for firm innovation apart from their separate impact. Knowledge intensive business services (KIBS) are used as a measure of external knowledge and level of education as a measure of internal knowledge. Grimpe and Kaiser (2010) interact the expenditure on R&D outsourcing with the internal R&D expenditures and find that it is positively related to the share of sales stemming from products that are new to the market. Antonelli and Colombelli (2015) measure internal knowledge as R&D input and external knowledge as number of patents and found that the interaction of these increased the firm’s probability to patent. Berchicci (2013) interacted internal R&D capacity and external R&D activities and found that firms with higher capacity needed less external R&D activities than others to increase their share of turnover stemming from improved products. Others have focused on the relatedness between the external and internal knowledge without interacting them, e.g. Fitjar, Huber, and Rodríguez-Pose (2016) who conclude that a medium cognitive distance between a firm and its partners is optimal for the probability to innovate.

This paper attempts to elaborate on that interaction further, but with a larger focus on the knowledge types and on knowledge bases relevant for the firm. Hence the hypothesis to be tested is: the interaction between firms' internal knowledge and external knowledge diversity has a positive relationship with their introduction of new export products.

Data and method

The data set used in this study is built upon several data bases of micro data provided by Statistics Sweden. These contain detailed information about all firms, establishments, and individuals in Sweden and identification numbers that allows for matching them. Due to this level of disaggregation the data have restricted public access. The final data set only includes private, single establishment firms within the manufacturing sector yielding 16768 firms in the panel, and a total of 76904 observations for the period 2003-2013.

The advantage of using trade data for this study is that the trade data set provides detailed information about which products have been exported by each firm. Hence the number of new products that have been exported can be identified. One obvious limitation with this data set is the exclusion of firms with only domestic sales. However, Sweden is a small open economy where trade is common which means that the number of firms included in the data set is still very large. Another limitation is that only tangible products are registered in an identifiable way. To minimise the consequences of that this study is focused on the manufacturing sector only.

Bishop and Gripaios (2010) argue that looking at the whole manufacturing sector jointly is not suitable as the separate industries are rather different, Consequently, I divide the manufacturing sector into four different groups that are estimated separately; (1) high-tech, (2) medium-high tech (3) medium-low tech, and (4) low-tech. These sectors are based on the Eurostat classification of the manufacturing industries' technological intensity, which is done for the NACE rev 2. This division is based on the technological intensity observed within the respective industries, where e.g. the high-tech sector includes industries producing pharmaceuticals, computer products, and aircrafts and the low-tech sector includes industries producing food, textiles and wood products.

The dependent variable

The export diversification in terms of new products introduced on export markets is constructed based on Swedish export data. The exported products are classified according to the combined nomenclature (CN), which is used within the whole EU. The classification is made on an 8-digit level, which is very detailed and making it difficult to claim that products from neighbouring codes are different products in practice. Another issue that could arise with the coding is the difficulties for those reporting the export to choose the correct code. To overcome these problems the codes are aggregated to a 6-digit level, which creates more differentiated product groups that should also be easier to identify and separate.

Apart from the issues of identifying the correct codes at one point in time some of these CN codes change yearly. This induces problems with identifying new products as they may actually be old products with new codes. To overcome this issue, the codes in the data set have been recoded so that the whole data set is based on the 2013 product codes.

The dependent variable used in the model is the number of new products exported per firm, which is defined as the number of products exported by firm i in period t, which were not exported in period t-1.

The independent variables

The independent variables are of two kinds, firm specific and region specific. The 290 Swedish municipalities are used as regional unit of analysis to capture the external knowledge diversity and population density in the firms' locations. A summary of the variables is presented in Table A1.

The firm specific variables are based on two data sets. One with establishment information, such as the NACE code and the year of establishment. This data set is also possible to match with the employees to acquire information about their educational background. The other data set contains information about the financial statements and assets, which is available on firm level only. Implying that issues could arise regarding in which establishment the production occurs. About 80% of the establishments in the data set are single establishment firms and for them there is no ambiguity. Although, it is unknown for the remaining 20 percent of the establishments. One solution is to approximate the values based on size of the establishment etc. However, as the geographical context of the firm is important for this study the multi establishment firms are completely excluded.

Firm specific measures

To control for the size of the individual firms a variable with the number of employees is added to the model. This is relevant because a firm with more employees is larger and can more easily manage to produce a larger number of different products (Cohen and Klepper 1996). In addition, dummies are included for the age of the firm (0-4, 5-9, 10-14, and 15+ years) to control for routines and experiences they may have acquired over the years[1]. Their internal knowledge capital is measured as the employees' average number of years of education.

It is relevant to consider financial information of firms when examining their innovative capability (see e.g. Schumpeter (1939), Nelson and Winter (1982), Lundvall (1992), and Borrás and Edquist (2013)). Considering the difficulties of obtaining finance due to asymmetric information, one should expect firms to use internal resources before looking for external financiers. This is controlled for by taking earnings before interest and tax (EBIT) and normalising it with the turnover to create a variable called operating margin. Measures of the firm's tangible (consisting of machinery of different kinds) and intangible (patents, trademarks etc.) capital are measured separately in millions of SEKs. The intangible capital could be seen, to some extent, as an indicator of previous innovative success and may help explaining path dependency of firms. Unfortunately, no direct measures of R&D firms' investments are available in the data set.

The ownership of the firm is another relevant factor to control for (Cook et al. 2013, Johansson and Lööf 2008). This is done with two dummies (*foreign corporate groups* and *Swedish corporate groups*). The base is private Swedish firms.

Region specific measures

The importance of diversified, but not too diverse, knowledge among the workers was highlighted previously. Most common measures like the melting pot index, or an inversed Herfindahl-index are potentially problematic in that sense as they measure on a scale

[1] Unfortunately, the data is truncated to 1986 for all older firms making a continuous variable unsuitable.

where the extremes are specialisation or diversity.

The external knowledge diversity in this study is therefore measured with the entropy measure related variety (RV), which is a measure of within variation. It was originally developed by Shannon (1948) to measure diversity of information, but has also been used to measure diversity of economic variables (c.f. Attaran (1986) and Frenken, Van Oort, and Verburg (2007))

This measure can be used for any classification with a hierarchical structure, which is the case for the codes used to classify type of education by Statistics Sweden. Assuming that each four-digit educational code belongs to only one two-digit educational code in a hierarchical fashion there can be a large diversity on either the two-digit level or on the four-digit level. A region with many different two-digit level subject groups would have a high unrelated diversity, which is closely related to the frequently used Herfindahl-index. A region with many different four-digit level codes present, within the same, 2-digit level subject group has a high related diversity.

In total 351 different codes on the four-digit level, were aggregated into 23 groups on the two-digit level. As these classifications are based on type of education and not level, they also include educations below university degrees[2].

$$E_{gr} = \sum_{i \in g} E_{ir} \qquad (1)$$

Equation (1) shows how the two-digit level shares E_{gr}, based on employment in region r, is calculated as the sum of the shares for the four-digit level employment E_{ir}, where each subgroup i belongs to a group g. The related diversity is then calculated as shown in equation (2) and (3). The difference is that E_{igr} is now introduced to represent

[2] A similar measure on the firm level would have been preferable to more specifically control for the compatibility of the knowledge bases. However, this type of measurement is not stable with very few observations, which is the case for many firms in the data set where half of the firms have less than 10 employees. As certain valuable skills are also found in many university programmes that are not engineering oriented I chose to control for university education in general as a threshold generating a better absorptive capacity.

the share of employees in the subgroup i as a share of the employees in sector g[3]. The H_{gr}, is then multiplied with the share of the two-digit sector, achieved in equation (1), and summed over all sectors g.

$$H_{gr} = -\sum_{i=1}^{I} E_{igr} \ln E_{igr} \qquad (2)$$

$$\mathrm{RD}_r = \sum_{g=1}^{G} E_{gr} H_{gr} \qquad (3)$$

If all employees are concentrated to one of the subject groups the measure takes its minimum value, 0. An equal distribution between all subject groups yield the maximum value ln (I), where I is the number of subgroups, in equation (2). If the distribution is even also between the groups g, the maximum of equation (3) becomes the average value of ln (I) among all groups G.

The overall external knowledge diversity measure could in theory be boosted by a high value for some subject groups that are not relevant for the manufacturing sector. This is remedied by limiting the measure to engineering related subjects. Mathematically, it is equivalent to equation 3, but the sectors G are restricted to fewer, engineering related, subjects.

Apart from the measures of external knowledge diversity, measures of the population density for the municipalities are added to control for the degree of urbanisation and possibilities of general knowledge spillovers that are not based on education (Ciccone and Hall 1996)[4].

---

[3] In the formula by Attaran (1986) Hg = $\sum \frac{E_i}{E_g} \ln \frac{E_i}{E_g}$. However, both Ei and Eg are defined as the share of employment in subgroup i or group g divided by total employment so: $\frac{Emp\ ir}{Total\ emp\ r} \Big/ \frac{Emp\ gr}{Total\ emp\ r} = \frac{Emp\ ir}{Emp\ gr} = E_{igr}$

[4] A measure for regional prosperity has also been attempted to be included in the model and the regression results were almost completely unaffected by the inclusion of this measure. Additionally,

Descriptive statistics

Table 1 displays the descriptive statistics for all variables used for the empirical study, and for two additional variables that only serves a descriptive purpose. The descriptives are divided according to the four sectors mentioned previously, as well as displayed for the manufacturing sector as a whole.

-- Table 1 about here--

The dependent variable, the number of new products exported, has an overall value of approximately 2.9 for the manufacturing sector. Examining the sectors more closely reveals differences between the groups, where the high-tech manufacturers on average have twice the number of new products as the low-tech manufacturers. Looking at their size, measured as the number of employees, it appears that low tech firms on average have less employees than high tech firms. This could indicate that the production in low tech firms is less labour intensive, or simply that they are smaller in their production size.

The average operating margin is negative to a varying degree for all subsectors, indicating that the EBIT on average has been negative for the firms in the data set – as the turnover is a positive value. One reason for this is could be the Swedish tax system, which through their high tax levels discourages firms from generating high profits in their financial statements. But the main reason is likely the definition of EBIT, which means that interest etc. earned that could make the result positive one has not yet been added. It should also be noted that the recent financial crisis is captured in this panel, which could add to the number of negative values observed.

The remaining financial variables indicates large differences between the groups. The tangible capital is largest in the low-tech firms and lowest for the high-tech firms, whereas the intangible capital shows the opposite pattern. Suggesting that on average high-tech firms use more intangible capital in their production than others, perhaps to produce smaller quantities of specialised goods. The opposite is observed for low-tech

---

the measure is correlated to 91% with the population density, resulting in increased variance inflation factors and was therefore excluded.

firms, where little intangible capital is used. Suggesting that they are more likely to produce less advanced goods in large quantities by an intensive use of machinery.

The ownership dummies indicate that about 50% of the firms in the data set are individual private firms. However, the remaining 50 % are either owned by a Swedish corporate group or a foreign corporate group. About 13% of the high-tech and medium-high tech firms are owned by a foreign corporate group, whereas the figure for the remaining groups are approximately 8 %.

The dummies indicating the age of the firms show that most the firms are at least 15 years old. However, there is a higher proportion of high-tech firms than other types of firms that are younger than 15. The high-tech firms are also those with the highest internal knowledge, with approximately 12.6 years of education on average. Completing the compulsory school and upper secondary school takes 12 years in Sweden, which indicates that the average employee in high-tech firms has some university education. For the remaining sectors this value ranges from 11.14 to 11.58 indicating that the average employee has an education from the upper secondary school.

The variable for external knowledge diversity, differ very little between the sectors – although a pattern can be seen where the value is highest for high-tech and becomes smaller for the less technology intensive firms. On the other hand, large differences are seen between the sectors regarding the population density. The high-tech firms are located in areas that are more than twice as dense the firms in the medium-high tech sector. The medium-low tech firms are generally located in the most sparsely populated areas.

The additional variables in Table 1 reinforce the proposition that high-tech firms are producing more specialised products in small quantities compared to the medium- and low tech firms. The export value in SEK per kg of weight is almost ten times higher for high-tech firms compared to low tech firms and almost five times higher than the value for the medium-high tech sector. The export value per employee is roughly the same for the high-tech and the low-tech subsectors, but lower for the remaining ones. Likely a result of the low-tech firms producing large quantities with a high capital intensity, which yields a high output per employee amounting to a high value even if each product is rather cheap. The high-tech firms produce lower quantities, but more valuable output. Summarised, the descriptives indicate that these sectors are indeed different from each other in many aspects. Thereby, supporting the decision not to group them together in the empirical part.

Model and estimation method

The logical choice of a model for this study follows a Poisson process, as the introduction of new products the firms supply to the foreign market is a count variable. However, the property of that distribution is that the variance equals the mean. Table 1 indicates that the variances are much higher than the mean and this would lead to an over dispersion of the model[5]. A negative binomial model is a more flexible version of a Poisson model since it allows for the mean to differ from the variance, and in this case it is more appropriate to use. For both the Poisson and the negative binomial models the rate of counts, μ, follows the functional form in equation (4).

$$\mu_i = \exp(\boldsymbol{x}_i'\boldsymbol{\beta}), i = 1, \ldots, N \tag{4}$$

As the interesting point is how firms' performances differ in different geographical contexts an analysis of the between variation is the most interesting – hence a random-effects model is preferable[6]. This choice is further supported by the nature of the data set. Many firms do not export products continuously, but rather they are moving in and out of that data set, which makes the panel unbalanced. Firms are present on average 4 times in the panel, which drastically lowers the amount of within variation possible to observe, as well as resulting in inconsistent fixed-effects (Greene 2008) and thus renders fixed-effects model also technically unsuitable in this study. In addition, the between variation is clearly lower than the between variation for most variables in this study limiting the variation that could be explained.

Consequently, a random-effects model is used and the specific model, estimated by maximum likelihood in this study, is shown in equation (5).

$$\ln(Y_{\boldsymbol{i,t}}) = \boldsymbol{\beta}'\boldsymbol{X}_{\boldsymbol{i,t-1}} + \boldsymbol{\gamma}'^{\boldsymbol{Z}_{r,t-1}} + \delta * interaction\ effect_{i,t-1} + \rho_t + (\alpha + u_i) + \varepsilon_{i,t} \tag{5}$$

---

[5] This is supported by formal LR tests which reject that the dispersion parameter alpha = 0.

[6] The LR tests for all regressions reject the null hypothesis of a pooled regression being more suitable for this data set.

The Y_i, represents the dependent variable – the number of new products exported. The $\boldsymbol{X_i}$ represents a vector of firm specific variables, such as the size and capital. The $\boldsymbol{Z_r}$ represents a vector of regional specific variables, including the measure of external knowledge diversity in engineering subjects. There is also one interaction term between the internal knowledge capital and the external knowledge diversity measure. Lastly, ρ_t indicates year controls.

When including interaction variables in the model there is a risk of inducing multicollinearity among the interaction term and the original variables, which was indeed the case here with a correlation of 98 % between the interaction variable and the original variables. This problem was remedied by centering the variables at their means before interacting them, which reduced the correlation to more manageable levels[7].

Empirical results

Regression results

To simplify the interpretation of the results, the coefficients presented in Table 2 are also displayed as incidence rate ratios (IRRs). A value of one indicates no relationship, whereas values above one indicates positive relationships and values between zero and one indicates negative relationships.

By using centered variables for the interaction term the results indicate relationships for the average values and not zero. E.g. for a medium-high tech firm located in a municipality with an average external knowledge diversity, having a higher level of education internally is statistically significant from zero and indicates a positive relationship with the number of new products.

--Table 2 about here--

However, the variable of main interest is the interaction effect, which indicates a positive relationship with the number of new products for the medium high tech and medium low tech firms. For these firms, there is no benefit of the external knowledge diversity unless they have internal knowledge to increase the absorptive capacity between

[7] For more details regarding the current correlation please consult Table A2.

the employees and the other workers in the region. This is not found for the high-tech firms where the external knowledge diversity is beneficial for firms with the average value of internal knowledge. For the low-tech firms, the opposite pattern is found – the external knowledge diversity is negatively related to the number of new products.

The coefficients of the interaction effect are not possible to interpret as they are, but must be calculated for different levels of education and external knowledge diversity. Using the medium-high tech results as an example the calculation is displayed in equation (6).

$$IRR_{IKC*RED} = \exp[0.0374 + (1.312 * Internal\ knowledge\ capital)] \tag{6}$$

If one is interested in the result for a firm with employees that on average have completed upper secondary school, i.e. spent 12 years in school the internal knowledge capital would be approximately 0.4 as the variable is centered around the mean, which is approximately 11.6 for this group. Inserting 0.4 in the above equation yields an IRR of approximately 1.75, suggesting that a one unit increase in external knowledge diversity would result in on average 75 % more new exported products. Bearing in mind that the measure of external knowledge diversity is rather small for most firms an increase of one is unlikely. Rather an increase of 0.1 units is more reasonable and they are associated with a 7.5 % increase in the number of new exported products. The equivalent IRR for the medium-low tech firms is approximately 1.85, which corresponds to an 8.5 % increase as the external knowledge diversity increases with 0.1. Overall, this provides some empirical support for the reasoning put forward by e.g. Cohen and Levinthal (1990), Nooteboom (2000a), and Boschma (2005).

The mixed results for the external knowledge diversity here match the previous studies that also find mixed results regarding the impact of diversity. The negative relation found for the low-tech sector indicates that they benefit more from specialisation, whereas the others benefit more from diversity. This may not be that surprising if one considers the type of industries that are included in the low-tech sector, such as manufacturing of food, textiles, and wood. These are all rather basic industries where we would expect specialisation and economics of scale to be important. The descriptive statistics also indicate that the firms in this sector are the least innovative, comparing the number of new products introduced yearly on the export market.

The positive relation found for the high-tech firms for the external knowledge diversity, combined with the insignificant interaction effect, suggests that they benefit from being located in municipalities with a large diversity of engineering education individuals without the internal knowledge working as a gatekeeper. One possible explanation for that is that the high-tech sector as a default contains a large amount of highly educated and skilled individuals. Hence, they do not increase their ability to absorb the external knowledge by increasing the share of highly educated employees further. This is somewhat confirmed by the descriptive statistics, which indicate that this sector on average do have the highest educated employees. The different results found for the different sectors also support the claim by Bishop and Gripaios (2010) that one should not aggregate the whole manufacturing sector as it is heterogeneous.

The coefficients of the capital variables are very small or not significantly different from zero, indicating that the capital of the firm has no meaningful relationship with the number of new products. Regarding the size, one more employee is associated with an increase of 0.2 to 0.3 percent in the number of new products exported. Larger relationships can be seen for the age and ownership of the firms. For most sectors, at least one coefficient is significantly different from zero, and generally they indicate that older firms are associated with more new exported products. E.g. a firm that is at least 15 years old is associated with 15-19% more new products than a firm that is younger than 5.

The coefficients of the ownership variables are significantly different from zero for all sectors. Looking at the high tech sector the IRRs show that a firm that is part of a Swedish corporate group, instead of being a regular private firm on average has 28.2% more new exported products. The equivalence for being part of a foreign corporate group is even higher – 36.8%. The same pattern is seen for the other sectors. This is in line with previous studies highlighting the importance of knowledge spillovers within MNEs for innovations (Cook et al. 2013, Johansson and Lööf 2008).

The population density is either insignificant or too small to be of any economic relevance. This suggests that simply increasing the density of people may not be associated with more new products in manufacturing firms – it matters who those people are and what they know.

In summary, the results indicate that the size of the firms and their capital are not very important for the generation of new products in these firms. The largest relationships can be observed for the age of the firms and their ownership status. This suggest that much knowledge comes with experience and organisational spillovers, which is in line with Torre and Rallet (2005), and previous studies by e.g. Cook et al. (2013). It also seems that for some of the sectors the internal knowledge capital, in terms of number of years of education, serves as a conductor for knowledge generation by lowering the cognitive distance to the external knowledge diversity in the geographical proximity of the firms, which has been suggested by e.g. Boschma (2005), Nooteboom (2000b), and Cohen and Levinthal (1989, 1990).

Sensitivity analysis

To ensure stability of the results several robustness checks were conducted. The regional level was changed to functional regions, the number of years used to lag the independent variables was increased and the measure of external knowledge diversity was extended. These results are all displayed in Table 3 for the variables of interest – the internal knowledge, the external knowledge diversity, and their interaction effect.

--Table 3 about here--

To make sure that the results found are not dependent on the geographical scale, or suffer from spatial autocorrelation between the municipalities due to commuting, the regional variables are also calculated for larger functional regions[8]. There are 72 of these and they are constructed based on actual commuting patterns between Swedish municipalities. Consequently, the borders of the functional regions exactly follow those of the municipalities.

The new estimations yield largely the same results as for the municipalities, although two changes can be observed: (1) the external knowledge diversity no longer has a positive relation with the number of new products for the high-tech sector, whereas the opposite is seen for the medium-high tech sector; (2) the IRR for the interaction effect is no longer significantly different from one for the medium-low tech firms, due to larger standard errors. The change for the high-tech firms could possibly indicate a larger

[8] Smaller units in form of 250 by 250 meter squares, were initially attempted as well but they were too small to capture diversity in most regions due to too few observations.

sensitivity to distance for these firms, which would be in line with e.g. Andersson, Klaesson, and Larsson (2014).

Innovation processes can take long time, and the one-year lag used might be too short. Hence the estimations are also run with longer lags. The results do not change much – in fact most coefficients have the same significance levels, signs, and even magnitudes as those for the one-year lag in the main model.

The only difference observed for the two-year lag is that the IRR for the external knowledge diversity in the medium-high tech sector becomes significantly lower than one, whereas the coefficient for the high-tech sector becomes insignificantly different from one. Additionally, the magnitude of the IRR for the interaction effects are larger than previously, reinforcing the importance of the absorptive capacity for the medium-high and medium-low tech sectors.

Increasing the lags to three years induce some larger changes than previously for the knowledge- and interaction variables. The internal knowledge is no longer significantly different from one, and the positive relationship between the interaction effect and the new products for the medium-high tech and the medium-low tech sectors have been replaced by positive relationships for the high-tech and the low-tech sectors.

The measure of external knowledge diversity was limited to engineering subjects. However, if one is to stay true to the spirit of e.g. Jacobs (1969) and Glaeser et al. (1992) diversity in general is positive for growth. Hence the estimations are also done for a general measure of external knowledge diversity, which is not limited to any subject fields.

The results for most variables are the same as previously, the major change can be seen for the main effect of the external knowledge diversity. The IRR for this variable has become significantly lower than zero, which suggests that for the firm with an average internal knowledge capital the relationship is negative with the number of new products. Possibly indicating that the knowledge diversity being measured is too broad to be useful for the firms in the manufacturing sector. This would further support the hypothesis that related knowledge is what counts for the firms (Cohen and Levinthal 1990, 1989).

Conclusions

This paper set out to analyse firms' internal knowledge in combination with the external knowledge diversity in their regions, and their joint relation to export diversification. I.e.

to examine if the interaction between internal and external knowledge had any added benefit for the manufacturing sector, where the external knowledge was measured as the related diversity of engineering subjects the inhabitants in each region studied in school. Based on the notion of the firms' absorptive capacity and cognitive proximity the hypothesis was that there would be a positive relation between the interaction term and the innovative ability of the firms – measured as number of new products introduced on the export market.

The results indicate that firms in the medium-high and medium-low tech sectors do benefit from this interaction, thus confirming the hypothesis for two of the subsectors. For a firm with an internal capital of 12 years of education for the employees, on average, a one unit increase of the external knowledge diversity is associated with 7.5-8.5 percent more new products. Hence a higher diversity, in terms of technological knowledge, seems to foster a larger diversity of exported products for these firms.

For the remaining two sectors the coefficients are not significantly different from zero. This result is robust to a change of spatial scale and longer time lag, but not a more general diversity measure, which further supports the importance of cognitive proximity for productive spillovers.

Important to bear in mind is that the study is conducted only on exporting manufacturing firms and should not be extrapolated to other firms. If more detailed information about domestic firms' production was available a more inclusive and general study could be made for this measure as a proxy of innovation. Also, the study only captures potential for spillovers, as the spillovers themselves are unmeasurable.

References

Andersson, Martin, Johan Klaesson, and Johan P. Larsson. 2014. "How Local are Spatial Density Externalities? Neighbourhood Effects in Agglomeration Economies." *Regional Studies*:1-14.

Antonelli, Cristiano, and Alessandra Colombelli. 2015. "External and Internal Knowledge in the Knowledge Generation Function." *Industry and Innovation* 22 (4):273-298. doi: 10.1080/13662716.2015.1049864.

Antonietti, Roberto, and Giulio Cainelli. 2011. "The role of spatial agglomeration in a structural model of innovation, productivity and export: a firm-level analysis." *The Annals of Regional Science* 46 (3):577-600.

Arrow, Kenneth J. 1962. "The Economic Implications of Learning by Doing." *The Review of Economic Studies* 29 (3):155-173.

Attaran, Mohsen. 1986. "Industrial Diversity and Economic Pperformance in U.S. Areas." *Annals of Regional Science* 20 (2):44.

Audretsch, David B., and Maryann P. Feldman. 1999. "Innovation in cities: Science-based diversity, specialization and localized competition." *European Economic Review* 43 (2):409-429.

Baldwin, John R., Desmond Beckstead, W. Mark Brown, and David L. Rigby. 2008. "Agglomeration and the Geography of Localization Economies in Canada." *Regional Studies* 42 (1):117-132.

Beaudry, Catherine, and Andrea Schiffauerova. 2009. "Who's right, Marshall or Jacobs? The localization versus urbanization debate." *Research Policy* 38 (2):318-337.

Berchicci, Luca. 2013. "Towards an open R&D system: Internal R&D investment, external knowledge acquisition and innovative performance." *Research Policy* 42 (1):117-127.

Bishop, Paul, and Peter Gripaios. 2010. "Spatial Externalities, Relatedness and Sector Employment Growth in Great Britain." *Regional Studies* 44 (4):443-454.

Borrás, Susana, and Charles Edquist. 2013. "The choice of innovation policy instruments." *Technological Forecasting and Social Change* 80 (8):1513-1522.

Boschma, Ron. 2004. "Competitiveness of Regions from an Evolutionary Perspective." *Regional Studies* 38 (9):1001-1014.

Boschma, Ron. 2005. "Proximity and Innovation: A Critical Assessment." *Regional Studies* 39 (1):61-74.

Breschi, Stefano, Francesco Lissoni, and Franco Malerba. 2003. "Knowledge-relatedness in firm technological diversification." *Research Policy* 32 (1):69-87.

Castaldi, Carolina, Koen Frenken, and Bart Los. 2015. "Related Variety, Unrelated Variety and Technological Breakthroughs: An analysis of US State-Level Patenting." *Regional Studies* 49 (5):767-781.

Ciccone, Antonio, and Robert E. Hall. 1996. "Productivity and the Density of Economic Activity." *The American Economic Review* 86 (1):54-70.

Cohen, Wesley M., and Steven Klepper. 1996. "Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D." *The Review of Economics and Statistics* 78 (2):232-243. doi: 10.2307/2109925.

Cohen, Wesley M., and Daniel A. Levinthal. 1989. "Innovation and Learning: The Two Faces of R & D." *The Economic Journal* 99 (397):569-596.

Cohen, Wesley M., and Daniel A. Levinthal. 1990. "Absorptive Capacity: A New Perspective on Learning and Innovation." *Administrative Science Quarterly* 35 (1):128-152.

Cohen, Wesley M., and Daniel A. Levinthal. 1994. "Fortune Favors the Prepared Firm." *Management Science* 40 (2):227-251. doi: 10.1287/mnsc.40.2.227.

Cook, Gary A. S., Naresh R. Pandit, Hans Lööf, and Börje Johansson. 2013. "Clustering, MNEs, and Innovation: Who Benefits and How?" *International Journal of the Economics of Business* 20 (2):203-227.

De Groot, H.L, J. Poot, and M. Smit. 2009. "Agglomeration externalities, innovation and regional growth: Theoretical perspectives, and meta-analysis." In *Handbook of regional growth and development theories*, edited by Roberta Capello and Peter Nijkamp. Northampton MA: Edward Elgar Publishing

Desrochers, Pierre, and Samuli Leppälä. 2011. "Opening up the 'Jacobs Spillovers' black box: local diversity, creativity and the processes underlying new combinations." *Journal of Economic Geography* 11 (5):843-863.

Duranton, Gilles, and Diego Puga. 2000. "Diversity and Specialisation in Cities: Why, Where and When Does it Matter?" *Urban Studies* 37 (3):533-555.

Fitjar, Rune Dahl, Franz Huber, and Andrés Rodríguez-Pose. 2016. "Not too close, not too far: testing the Goldilocks principle of 'optimal' distance in innovation networks." *Industry and Innovation* 23 (6):465-487. doi: 10.1080/13662716.2016.1184562.

Fitjar, Rune Dahl, and Andrés Rodríguez-Pose. 2017. "Nothing is in the Air." *Growth and Change* 48 (1):22-39. doi: 10.1111/grow.12161.

Frenken, Koen, Frank Van Oort, and Thijs Verburg. 2007. "Related Variety, Unrelated Variety and Regional Economic Growth." *Regional Studies* 41 (5):685-697.

Gertler, Meric 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, Edward L. 1999. "Learning in Cities." *Journal of Urban Economics* 46 (2):254-277.

Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman, and Andrei Shleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100 (6):1126-1152.

Greene, William H. 2008. *Econometric Analysis*. Upper Saddle River: Pearson Prentice Hall

Grimpe, Christoph, and Ulrich Kaiser. 2010. "Balancing Internal and External Knowledge Acquisition: The Gains and Pains from R&D Outsourcing." *Journal of Management Studies* 47 (8):1483-1509.

Jacobs, Jane. 1969. *The economy of cities*. New York: Random House.

Johansson, Börje, Sara Johansson, and Tina Wallin. 2015. "Internal and External Knowledge and Introduction of Export Varieties." *The World Economy* 38 (4):629-654.

Johansson, Börje, and Hans Lööf. 2008. "Innovation Activities Explained by Firm Attributes and Location." *Economics of Innovation and New Technology* 17 (6):533-552.

Kirat, Thierry, and Yannick Lung. 1999. "Innovation and Proximity: Territories as Loci of Collective Learning Processes." *European Urban and Regional Studies* 6 (1):27-38.

Lundvall, Bengt-Åke, ed. 1992. *National systems of innovation - towards a theory of innovation and interactive learning*. London: Pinter.

Malmberg, A., and P. Maskell. 2002. "The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering." *Environment and Planning A* 34 (3):429-449.

Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan.

Nathan, Max, and Neil Lee. 2013. "Cultural Diversity, Innovation, and Entrepreneurship: Firm-level Evidence from London." *Economic Geography* 89 (4):367-394.

Nelson, Richard R., and Sidney G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Niebuhr, Annekatrin. 2010. "Migration and innovation: Does cultural diversity matter for regional R&D activity?" *Papers in Regional Science* 89 (3):563-585.

Nooteboom, Bart. 2000a. *Learning and Innovation in Organizations and Economies* Oxford: Oxford University Press.

Nooteboom, Bart. 2000b. "Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance." *Journal of Management and Governance* 4 (1):69-92.

OECD/Eurostat. 2005. *Oslo Manual*: OECD Publishing.

Ozgen, Ceren, Peter Nijkamp, and Jacques Poot. 2013. "The impact of cultural diversity on firm innovation: evidence from Dutch micro-data." *IZA Journal of Migration* 2 (18).

Paci, Raffaele, and Stefano Usai. 1999. "Externalities, knowledge spillovers and the spatial distribution of innovation." *GeoJournal* 49 (4):381-390.

Parrotta, Pierpaolo, Dario Pozzoli, and Mariola Pytlikova. 2012. "The nexus between labor diversity and firm's innovation." *IZA Discussion Paper Series* (No. 6972).

Porter, Michael E. 1990. *The competitive advantage of nations*. Nw York: Free Press.

Qian, Haifeng. 2013. "Diversity Versus Tolerance: The Social Drivers of Innovation and Entrepreneurship in US Cities." *Urban Studies* 50 (13):2718-2735.

Rallet, Alain, and André Torre. 1999. "Is geographical proximity necessary in the innovation networks in the era of global economy?" *GeoJournal* 49 (4):373-380.

Romer, Paul M. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94 (5):1002-1037.

Rosenthal, Stuart S., and William C. Strange. 2001. "The Determinants of Agglomeration." *Journal of Urban Economics* 50 (2):191-229.

Sampson, Rachelle C. 2007. "R&D Alliances and Firm Performance: The Impact of Technological Diversity and Alliance Organization on Innovation." *Academy of Management Journal* 50 (2):364-386.

Schumpeter, Joseph. 1939. *Business Cycles* Vol. 1. New York: McGraw Hill.

Shannon, C. E. 1948. "A mathematical theory of communication." *Bell System Technical Journal* 27.

Tavassoli, Sam, and Nunzia Carbonara. 2014. "The role of knowledge variety and intensity for regional innovation." *Small Business Economics* 43 (2):493-509.

Teece, David J. 1992. "Competition, cooperation, and innovation: Organizational arrangements for regimes of rapid technological progress." *Journal of Economic Behavior & Organization* 18 (1):1-25.

Thornhill, Stewart. 2006. "Knowledge, innovation and firm performance in high- and low-technology regimes." *Journal of Business Venturing* 21 (5):687-703.

Torre, Andre, and Alain Rallet. 2005. "Proximity and Localization." *Regional Studies* 39 (1):47-59.

Usai, Stefano, and Raffaele Paci. 2003. "Externalities and Local Economic Growth in Manufacturing Industries." In *European Regional Growth*, edited by Bernard Fingleton, 293-321. Berlin, Heidelberg: Springer Berlin Heidelberg.

van Beers, Cees, and Fardad Zand. 2014. "R&D Cooperation, Partner Diversity, and Innovation Performance: An Empirical Analysis." *Journal of Product Innovation Management* 31 (2):292-312.

Wixe, Sofia. 2015. "The Impact of Spatial Externalities: Skills, Education and Plant Productivity." *Regional Studies* 49 (12):2053-2069.

Wixe, Sofia. 2016. "Neighbourhood related diversity, human capital and firm innovation." *Papers in Regional Science* Forthcoming.

Wixe, Sofia, and Martin Andersson. 2016. "Which Types of Relatedness Matter in Regional Growth? Industry, Occupation and Education." *Regional Studies*:1-14.

Wuyts, Stefan, Massimo G. Colombo, Shantanu Dutta, and Bart Nooteboom. 2005. "Empirical tests of optimal cognitive distance." *Journal of Economic Behavior & Organization* 58 (2):277-302.

Østergaard, Christian R., Bram Timmermans, and Kari Kristinsson. 2011. "Does a different view create something new? The effect of employee diversity on innovation." *Research Policy* 40 (3):500-509.

Table 1: Descriptive statistics

	High-tech			Medium-high tech			Medium-low tech			Low-tech			All firms		
VARIABLES	mean	median	sd	mean	median	sd	mean	median	sd	mean	median	sd	mean	median	sd
Number of new products	4.224	2	6.425	3.927	2	6.183	2.628	1	3.999	2.098	1	3.461	2.948	1	4.863
Independent variables															
Employees per firm	31.16	8	99.81	35.47	12	80.28	27.89	12	100.43	25.95	10	63.29	29.67	11	83.93
Operating margin	-0.708	0.044	20.55	-0.241	0.046	15.28	-0.106	0.051	10.694	-0.316	0.039	19.919	-0.260	0.045	16.28
Tangible capital	8,278	393.23	55,636	13,989	1178.18	141,988	13,093	2378.07	138,088	17,226	1261.89	131,068	14,337	1401.50	132,436
Intangible capital	4,684	0	56,502	2,004	0	24,975	443.6	0	11,333	689.5	0	15,169	1,296	0	23,236
Private Swedish firm	0.486	0	0.500	0.448	0	0.497	0.495	0	0.500	0.569	0	0.495	0.505	1	0.500
Swedish corporate group	0.390	0	0.488	0.417	0	0.493	0.422	0	0.494	0.354	0	0.478	0.396	0	0.489
Foreign corporate group	0.125	0	0.330	0.134	0	0.341	0.083	0	0.275	0.077	0	0.266	0.099	0	0.298
Less than 5 years' old	0.206	0	0.404	0.145	0	0.352	0.120	0	0.325	0.166	0	0.372	0.149	0	0.356
5 to 9 years' old	0.171	0	0.376	0.135	0	0.341	0.121	0	0.326	0.139	0	0.346	0.135	0	0.341
10 to 14 years' old	0.155	0	0.362	0.144	0	0.351	0.135	0	0.342	0.134	0	0.341	0.139	0	0.346
At least 15 years' old	0.469	0	0.499	0.577	1	0.494	0.623	1	0.485	0.561	0	0.496	0.577	1	0.494
Internal knowledge	12.58	12.23	1.660	11.58	11.42	1.205	11.14	11.05	0.983	11.20	11	1.225	11.40	11.21	1.253
External knowledge div.	0.508	0.492	0.073	0.485	0.474	0.074	0.477	0.463	0.077	0.468	0.456	0.074	0.479	0.466	0.076
Population density	954.2	133.95	1,431	379.2	60.15	899.5	235.2	34.97	675.7	521.6	47.71	1,152	426.6	48.54	1,002
Export value per kg*	5,666	1,797	38,737	1,385	208	42,730	608	105	3,766	616	76	10,266	1,225	139	26,048
Export value per employee*	681,933	151,041	3.4e+06	614,198	142,197	2.8e+06	405,600	56,107	1.9e+06	710,301	38,431	2.6e+07	587,146	70001	1.5e+07

* For descriptive purposes only.

Table 2: Regression results for the number of new products on municipal level, both displayed as coefficients and incidence rate ratios (IRR).

	High-tech		Medium-high tech		Medium-low tech		Low-tech	
VARIABLES	Coefficients	IRR	Coefficients	IRR	Coefficients	IRR	Coefficients	IRR
Employees per firm	0.00335***	1.003***	0.00248***	1.002***	0.00187***	1.002***	0.00335***	1.003***
	(0.000236)	(0.000236)	(0.000119)	(0.000120)	(0.000184)	(0.000184)	(0.000216)	(0.000217)
Operating margin	0.000826	1.001	-0.000438	1.000	-0.000506	0.999	-0.000361	1.000
	(0.000561)	(0.000561)	(0.000933)	(0.000932)	(0.000483)	(0.000483)	(0.000246)	(0.000246)
Tangible capital	-7.16e-07*	1.000*	-7.57e-08	1.000	1.46e-07***	1.000***	-3.79e-07***	1.000***
	(3.78e-07)	(3.78e-07)	(4.89e-08)	(4.89e-08)	(3.89e-08)	(3.89e-08)	(7.44e-08)	(7.44e-08)
Intangible capital	-7.13e-07***	1.000***	8.44e-07***	1.000***	3.82e-07	1.000	6.51e-07	1.000
	(2.63e-07)	(2.63e-07)	(1.83e-07)	(1.83e-07)	(3.74e-07)	(3.74e-07)	(6.73e-07)	(6.73e-07)
Swedish corporate group	0.248***	1.282***	0.220***	1.246***	0.230***	1.259***	0.107***	1.112***
	(0.0376)	(0.0482)	(0.0196)	(0.0245)	(0.0193)	(0.0242)	(0.0204)	(0.0227)
Foreign corporate group	0.313***	1.368***	0.280***	1.323***	0.346***	1.414***	0.199***	1.221***
	(0.0531)	(0.0727)	(0.0274)	(0.0362)	(0.0323)	(0.0457)	(0.0361)	(0.0441)
5 to 9 years' old	0.124***	1.132***	0.00663	1.007	0.0400	1.041	-0.0374	0.963
	(0.0436)	(0.0494)	(0.0271)	(0.0273)	(0.0291)	(0.0303)	(0.0268)	(0.0258)
10 to 14 years' old	0.100***	1.105**	0.0898***	1.094***	0.0712**	1.074**	-0.0260	0.974
	(0.0501)	(0.0554)	(0.0293)	(0.0320)	(0.0309)	(0.0332)	(0.0295)	(0.0287)
At least 15 years' old	0.176***	1.192***	0.140***	1.150***	0.146***	1.158***	0.0192	1.019
	(0.0504)	(0.0601)	(0.0289)	(0.0332)	(0.0297)	(0.0343)	(0.0279)	(0.0284)
Internal knowledge	-0.00870	0.991	0.0374**	1.038**	0.00116	1.001	0.0181	1.018
	(0.0238)	(0.0236)	(0.0166)	(0.0172)	(0.0182)	(0.0182)	(0.0166)	(0.0169)
External knowledge diversity	1.181*	3.257*	-0.623	0.536	-0.0290	0.971	-1.536***	0.215***
	(0.702)	(2.285)	(0.445)	(0.239)	(0.471)	(0.458)	(0.502)	(0.108)
Interaction effect	-0.933	0.393	1.312*	3.713*	1.533*	4.631*	0.905	2.472
	(1.089)	(0.428)	(0.723)	(2.684)	(0.819)	(3.791)	(0.784)	(1.937)
Population density	4.42e-05***	1.000***	5.91e-06	1.000	-6.51e-06	1.000	-2.43e-06	1.000
	(1.44e-05)	(1.44e-05)	(1.18e-05)	(1.18e-05)	(1.56e-05)	(1.56e-05)	(9.13e-06)	(9.13e-06)
Number of observations	5,853		21,929		23,883		25,239	
Number of firms	1,334		4,516		4,989		5,929	
Log likelihood	-13005.98		-47634.41		-45527.06		-44048.63	

Notes: The constant was suppressed in the table to save space, all models contain year FE, standard errors in parentheses, and *** p<0.01, ** p<0.05, * p<0.1

Table 3: Regression results for the main model as well as four alternative models, displayed as incidence rate ratios (IRR).

SECTOR	VARIABLES	Main model	Functional regions	2 years' lag	3 years' lag	General educational diversity
High-tech firms	Internal knowledge	0.991	0.992	1.000	1.009	0.996
		(0.0236)	(0.0236)	(0.0264)	(0.0290)	(0.0238)
	External knowledge diversity	3.257*	0.799	1.001	1.567	0.562*
		(2.285)	(1.089)	(0.771)	(1.336)	(0.195)
	Interaction effect	0.393	0.207	0.459	13.78*	0.618
		(0.428)	(0.328)	(0.618)	(20.52)	(0.227)
Medium-high tech firms	Internal knowledge	1.038**	1.037**	1.044**	1.023	1.039**
		(0.0172)	(0.0172)	(0.0192)	(0.0209)	(0.0172)
	External knowledge diversity	0.536	0.305*	0.310**	0.357*	0.530***
		(0.239)	(0.204)	(0.158)	(0.201)	(0.0932)
	Interaction effect	3.713*	8.688**	6.480**	0.627	1.784**
		(2.684)	(7.664)	(5.972)	(0.654)	(0.403)
Medium-low tech firms	Internal knowledge	1.001	1.004	1.007	1.003	1.006
		(0.0182)	(0.0183)	(0.0204)	(0.0223)	(0.0184)
	External knowledge diversity	0.971	0.367	0.617	0.842	0.586***
		(0.458)	(0.242)	(0.335)	(0.526)	(0.101)
	Interaction effect	4.631*	4.570	12.95**	4.838	1.481
		(3.791)	(4.266)	(13.82)	(5.943)	(0.355)
Low-tech firms	Internal knowledge	1.018	1.021	0.993	0.980	1.025
		(0.0169)	(0.0169)	(0.0186)	(0.0206)	(0.0170)
	External knowledge diversity	0.215***	0.0413***	0.177***	0.132***	0.361***
		(0.108)	(0.0280)	(0.104)	(0.0877)	(0.0623)
	Interaction effect	2.472	4.180	4.635	6.629*	1.183
		(1.937)	(3.667)	(4.547)	(7.476)	(0.258)

Notes: All models contain year FE, standard errors in parentheses, and *** p<0.01, ** p<0.05, * p<0.1

Table A1: Variable summary

VARIABLES	Description
Number of new products	Number of 6-digit CN codes exported by each firm.
Firm specific variables	
Employees per firm	The number of employees.
Operating margin	The EBIT expressed in SEK divided by the turnover expressed in SEK.
Tangible capital	The value of tangible capital in millions of SEK.
Intangible capital	The value of intangible capital in millions of SEK.
Ownership	1: Private Swedish firm 2: Swedish corporate group 3: Foreign corporate group Private Swedish firm is the base.
Age	1: Less than 5 years' old 2: 5 to 9 years' old 3: 10 to 14 years' old 4: At least 15 years' old Less than 5 years' old is the base.
Average years of education	The average number of years spent in school among the employees.
Regional specific variables	
External knowledge diversity	An entropy measure of within variation for all types of education within in engineering subjects only. This has been calculated for both functional regions and municipalities.
Interaction effect	An interaction variable between the average years of education and the entropy measure for the variation of engineering education in the region (RED engineering). This has been calculated for both functional regions and municipalities.
Population density	The number of inhabitants per square kilometre. This has been calculated for both functional regions and municipalities.

Table A2: Correlation matrix

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1)	Internal knowledge (cent.)	1.00													
(2)	Employees per plant		1.00												
(3)	Operating margin			1.00											
(4)	Tangible capital		0.49		1.00										
(5)	Intangible capital		0.22		0.16	1.00									
(6)	Foreign corporate group		0.24		0.12	0.06	1.00								
(7)	Domestic corporate group		0.09				-0.27	1.00							
(8)	Younger than 5 years		-0.10					-0.17	1.00						
(9)	5 to 9 years' old		-0.06					-0.08	-0.17	1.00					
(10)	10 to 14 years' old								-0.17	-0.16	1.00				
(11)	Older than 15 years		0.14		0.05		0.05	0.20	-0.49	-0.46	-0.47	1.00			
(12)	External knowledge diversity (cent.)	0.20						0.05	-0.12	-0.05		0.14	1.00		
(13)	Population density								0.07			-0.07		1.00	
(14)	Interaction effect								-0.05			0.06			1.00

Notes: Only correlations above 5% are displayed. All displayed correlations have p-values below 0.01