The Review of Regional Studies

The Official Journal of the Southern Regional Science Association

Empirical Assessment of the Smart Specialization Concept on Firm Performance in European Urban and Rural Regions[*]

Pia Nilsson

Department of Economics, Finance and Statistics, Jönköping International Business School, Jönköping University, SWEDEN

Abstract: This paper studies the role played by factors that are considered central to the concept of smart specialization on firm performance in a European context. The focus is on the type of spatial spillovers that occur at the firm level, which are connected to technological relatedness and knowledge externalities. The influence of such externalities is studied using firm-level data on firms located across Europe and unobserved heterogeneity and spatial dependencies are modelled by employing a multilevel model. Diverging patterns across the urban-rural range are studied by applying a regional typology. Findings indicate that measures thought to reflect smart specialization are positively associated with firm performance. It is also found that indicators of smart specialization do not affect firm performance homogenously across the studied regions.

Keywords: firm performance, related variety, smart specialization, urban-rural

JEL Codes: R11, R12, C21

1. INTRODUCTION

There is a growing focus on the role played by smart specialization from the perspective of European regional growth (Asheim, Boschma, and Cooke, 2011; McCann and Ortega-Argilés, 2015; Naldi, Nilsson, Westlund and Wixe, 2015). Smart growth and the related concept of smart specialization are central in the Europe 2020 growth strategy, which comprises a broad set of strategies aimed to spur knowledge transfer, innovation, and investments in research and development (R&D). In this context, smart specialization refers to regional growth strategies that account for differences within and between regions and that build on regional prerequisites of competitive advantage (Tödtling and Trippl, 2005; Asheim, Boschma, and Cooke, 2011; Camagni and Cappello, 2013). The main message is that smart specialization can be achieved by implementing policies in which regions specialize in R&D and innovation activities that have a base in their areas of competitive advantage. Many of the key ideas behind smart specialization are connected to established economic theories that highlight the role played by knowledge and innovation for regional growth (Romer, 1990). The new perspective offered here is a stronger focus on regional preconditions and bottom-up-planning processes and policies that are embedded in the regional knowledge, resources, and entrepreneurial base of the area.

There are several outstanding issues related to the logic of smart specialization as discussed in prior literature. One concerns how to define the concept and its component parts to enable empirical assessments. Another concern is if the logic can provide a solution for rural and peripheral regions of Europe. Specifically, it is argued that the underlying logic of smart specialization can offer only limited possibilities for regions that are remote or very isolated, due to lack of scale and distance decay effects (McCann and Ortega-Argilés, 2015; Naldi et al., 2015). Based on these studies, remote rural regions can be predicted to have limited potential to benefit from the Europe 2020 strategy and its smart specialization policies.

The purpose of this paper is to empirically address the factors that are identified as central to the concept of smart specialization on firm performance in a diverse set of rural regions. The focus is on determinants connected to technological relatedness and knowledge externalities. The empirical analysis uses firm-level data across Europe with the overall purpose of addressing two main perspectives set forth in research on the smart specialization concept. The first is that determinants of smart specialization can be linked to technological relatedness across and within industries, the presence of agglomeration economies, and knowledge externalities, as argued in Naldi et al. (2015). Their argument follows the view that diversity in related economic activities in terms of technology and knowledge base stimulates knowledge spillovers and thereby innovation and growth (Nooteboom, 2000; Frenken, Van Oort, and Verburg, 2007). The second perspective suggests that determinants of smart specialization have different impacts on the intended goals depending on the size of regions. Foray et al. (2012) and McCann and Ortega-Argilés (2015) argue that the underlying logic of smart specialization can offer only limited possibilities for regions that are remote or very isolated, due to lack of scale, distance decay effects, and socioeconomic pre-conditions. To test this proposition and take the analysis further, this paper addresses the influence of technological relatedness and knowledge externalities on firm performance in different types of regions.

The empirical analysis is based on firm-level data across Europe obtained from the ORBIS (Bureau Van Dijck) database.[1] These data hold information on financial and administrative characteristics in terms of firm size, location, and indicators of firm performance. These data are combined with spatial indices that reflect technological relatedness and knowledge interconnections, measured both at the local and regional levels. They are calculated with regards to intra- and inter-industry composition to reflect firms' access to external knowledge (also referred to as related and unrelated variety) as outlined in Frenken, Van Oort, and Verburg (2007). Having access to firm-level microdata enables the study to capture heterogeneity across firms and conduct different and more detailed types of data aggregation (by geographical unit and industry). Prior studies tend to take a regional approach and focus on European NUTS 3 regions in assessing knowledge interconnections and innovation for regional growth (e.g., Frenken, Van Oort, and Verburg, 2007). NUTS 3 is the official division of Europe for regional statistics and divides European countries into 1342 regions for the purpose of data collection and harmonization of European regional statistics.[2] Due to the high degree of regional heterogeneity at the NUTS 3 level, it is necessary to take an empirical analysis to the firm or local level. Having access to microdata across Europe allows for the estimation of a nested multilevel model that addresses different types of locally bounded externalities and that accounts for unobserved heterogeneity by combining determinants at the firm, local, and regional level. To address diverging patterns across the urban-rural range, the analysis employs category-wise estimations using the EDORA structural typology, which

---

[1] The sample include firms from seven European countries and the initially retrieved sample contain around 4.5 million firms.

[2] See http://ec.europa.eu/eurostat/web/nuts/overview for an overview of the NUTS classification system.

divides European NUTS 3 regions into five categories ranging from urban to agrarian regions. This is a regional typology designed to account for patterns of rural differentiation across Europe and that is supportive of the place-based approaches put forward in the Barca Report (Barca, 2009; Coupus and Noguera, 2010; Barca, McCann and Rodríguez-Pose, 2012).

The paper is structured as follows. The second part reviews the most relevant literature on smart specialization, technological relatedness, and knowledge externalities in the context of urban and rural regions. The third section describes the data and the empirical approach to test the outlined hypotheses and summarizes some of the key facts observed in the data. The fourth section present the findings and discusses their relevance in relation to theory and prior literature. Section five concludes.

2. BACKGROUND AND THEORETICAL FRAMEWORK

The Europe 2020 strategy introduces several growth targets formulated to enhance productivity and increase R&D and innovation in Europe (European Commission, 2010). Among other things, these targets are formulated with the goal of reducing regional inequality, particularly between the urban and peripheral regions of Europe. The concept of smart growth and smart specialization are central to the strategy and are highlighted to provide a logic that contribute to long-term growth and convergence.[3] Although the concept of smart growth is foremost policy-oriented, its interpretation and possible application in a regional context has been discussed in several papers (Boschma, 2014; McCann and Ortega-Argilés, 2015; Naldi et al., 2015). Among other things, it is argued that the concept and its underlying logic is not new. Rather, it reflects the well-known fact that one-size-fits-all regional policy models should be reformulated into policies that are both place-based and knowledge-based (Tödtling and Trippl, 2005; Camagni and Cappello, 2013; Boschma, 2014). The idea that knowledge and innovation are key drivers of regional growth is also well-known. The way in which smart growth is defined can thus be related to established conceptual frameworks regarding the importance of innovation, knowledge, and their externalities as key factors in explaining regional economic growth and convergence (Grossman and Helpman, 1990).

This connects to the literature on the spatial determinants of firm innovation and economic performance. Here, it is emphasized that the characteristics of the locational environment of firms influences their economic performance and their ability to innovate (Boschma, 2005; Feldman and Kogler, 2010). A key emphasis in this literature is that agglomeration economies enhance the cross-fertilization of ideas, which give rise to knowledge externalities (Rosenthal and Strange, 2003; Audretsch and Feldman, 2004). The literature on the spatial determinants of firm performance has largely distinguished between two types of externalities, those that arise from industrial specialization (Marshall, 1920) and those that arise from industrial diversification (Jacobs, 1969). In Marshall's view, firms benefit from being located geographically close to other firms within the same industry for several reasons, including minimization of transport costs, the sharing of inputs, and access to a specialized workforce. Hence, this view emphasizes the idea that knowledge externalities occur between firms within the same industry, as well as those in similar industries. In the view of Jacobs, the sources of knowledge spillovers are thought to be external to the industry which the firm operates, meaning that it is the local variety of industries that promotes knowledge spillovers and eventually firm performance and innovation activity. Hence, agglomeration economies result from either the geographic concentration of similar industrial activities linked to the

---

[3] Growth policies in Europe, such as the Europe 2020 strategy, act as an umbrella and include a broad set of strategies to encourage smart, sustainable, and inclusive sustainable growth. The present paper is limited to the perspectives of smart growth that can be connected to knowledge and innovation.

economies of a location or from diversified concentration of activities, which are more strongly linked to the urban economies (Neffke, Henning, and Boschma, (2011).

The concept of smart specialization can be linked to the theories above and the idea that spatial determinants play an important role in enhancing firm innovativeness and economic performance. But it is argued to not simply be understood as a strategy of specialization, but as a strategy through which firms and regions focus on interests that are connected to and embedded in their knowledge, resource, and entrepreneurial base (Foray and Hall, 2011). This points to other central components of smart growth, namely the role of place-specific policies and the need to acknowledge that industries with different knowledge bases innovate in different ways (Jensen et al. 2007; Asheim, Boschma, and Cooke, 2011; McCann and Ortega-Argilés, 2015). Naldi et al. (2015) argue along these lines and stress that the concepts of smart growth and smart specialization involve the connection between knowledge, innovation, and what appears to be the relatable specialization of the region. This implies that there is also a sectoral dimension, meaning this it is not only geographic concentration that matters, but also industrial composition in a more qualitative sense (Jacquemin and Berry, 1979; Duranton and Puga, 2004).

Frenken, Van Oort, and Verburg (2007) take the question of regional diversity one step further and emphasize that there needs to be some sort of cognitive proximity or relatedness among firms for the effects to emerge. Following Nooteboom (2000), they argue that to be able to fully depict the complexity of spatial externalities there is a need for a qualitative approach with regards to the implications of industry membership. Building on such approaches, the concepts of related and unrelated variety were introduced to conceptualize technological relatedness among firms (Frenken, Van Oort, and Verburg, 2007). Following their terminology, related variety is defined as within-industry diversity and unrelated variety as between-industry diversity. What follows from this outline is that the smart specialization logic can be closely linked to these concepts as they point to the importance of relatable specialization to obtain regional growth and development (Naldi et al., 2015).

Prima facie, agglomeration economies may seem relevant only for urban regions. But they also crop up in rural regions, although on a different scale as there are differences within rural regions and areas that are relatively more urbanized than others. Hence, there also exist areas within rural regions that have potential for matching, sharing, and learning processes to take place. As highlighted in Naldi et al. (2015), there are reasons to expect that firms in rural regions may have more to gain from increases in agglomeration as the marginal effect from increased agglomeration may be larger in rural regions than for those that are already urbanized.

Although the concepts of smart growth and smart specialization are not new, their interpretation and application in a regional context has been under recent debate (McCann and Ortega-Argilés, 2015). As discussed, the smart growth concept connects to the role of innovation and technological advances as key factors in explaining economic growth and productivity. This also means that increasing returns to investments, due to technological improvements, can be expected to lead to divergence among regions (Asheim, Boschma, and Cooke, 2011). The underlying theory predicts that R&D investments in core regions are more efficient when it comes to generating innovations, compared to investments in peripheral regions. It is thus likely that regions will respond differently to investments in knowledge, innovation, and relatable specialization, depending on whether they are peripheral or nonperipheral.

Naldi et al. (2015), studied these issues considering regional heterogeneity and argue that there are reasons to believe that many determinants of smart specialization logic have

different impacts across regions of different types. They argue that the logic is well suited for urban and intermediate regions that tend to have a large population and industrial base as they benefit from their size advantage and spillovers from nearby cities, without having to incur the higher costs of sustaining urban core areas (Dijkstra and Poleman, 2008; Noseleit, 2012). Being intermediate also implies improved connectivity and mobility, which may facilitate the inflow (and outflow) of knowledge and skills (McCann and Ortega-Argilés, 2015). In summary, there should exist more possibilities to achieve specialized diversification across related technologies in urban and intermediate rural regions compared to rural peripheral regions. Rodríguez-Pose (2001) and Bilbao-Osorio and Rodríguez-Pose (2004) raise similar issues, such as whether research and development policies can target peripheral rural regions in Europe. They argue that the need to reach a minimum threshold of research and the existence of distance decay effects in the diffusion of technological spillovers are some of the factors that may prevent peripheral rural regions from generating innovations and, hence, growth.

The logic that underlies the concept of smart specialization offers limited possibilities for regions that are remote or very isolated, due to their lack of scale, distance decay effects, and the unavailability of necessary socioeconomic conditions according to McCann and Ortega-Argilés (2014) and Boschma (2014). A general characteristic of rural peripheral regions is a lack of accessibility to markets and services, negative migratory balance, low land values, and low education levels. In total this means that remote regions do not have much potential for endogenous development (Naldi et al. 2015).

3. DATA, VARIABLES, AND EMPIRICAL MODEL

To empirically address the research questions outlined above, the paper uses firm-level data across Europe. These data are compiled from the ORBIS database of Bureau Van Dijck (BvDEP) and contain financial and administrative information of firms across Europe with locational references to both postcodes and NUTS 3 codes. Data for the latest available year (2011) are compiled that include 4.5 million firms in Sweden, France, Italy, Austria, Hungary, Czech Republic, and Poland. Ideally, we would prefer to include all European firms, but the cost of downloading and compiling the full set of 14 million observations (total number of European firms currently in ORBIS) was deemed to be excessive. The particular geographical delimitation was mainly chosen for operational reasons. The sample is restricted to active firms and firms that registered a NACE code (the European standard for industrial classifications).[4] Removing observations with missing values on key variables (turnover and location) results in a sample of 940,184 firms nested into 256 European NUTS 3 regions and 34,964 postcode areas. Ribeiro, Menghinello, and DeBacker (2010) compared the representativeness of the ORBIS data to the national SDBS (Structural Business Demographic Statistics) databases with regards to number of firms, size categories, and industry. Their assessment suggests that these data are comparable in several European countries.[5] We assume this holds for our situation as well.

Two measures of firm performance are used as dependent variables (see Table B1 in Appendix B for variable definitions). The first is turnover per employee defined as firms operating revenue (net sales plus other operating revenues and stock variations) divided by the total number of employees. Turnover per employee and similar revenue measures provides a measure of productivity and how efficiently firms use their resources. The second dependent variable is Return on Equity (ROE), defined as net income divided by shareholder's equity,

---

[4] Nomenclature generale des Activites economiques dans les Communautes Europeennes (NACE) refers to the industrial classification as defined in Revision 2 used by Eurostat.

[5] The industrial composition of the firms in our sample is displayed in Appendix A.

which is a measure of firm profitability and performance (Dewenter and Malatesta, 2001; Johansson, Lööf, and Olsson, 2010). Ideally, a count of patents and trademarks, which is available in the database, could also be used as a dependent variable to proxy for innovation (Acs, Anselin, and Varga, 2002). Many firms do not have any registered patents and trademarks, so a substantial number of observations have either zero or a missing value reported.[6] We are unable to judge the reliability of this variable, e.g., if zero can be treated as a true value, so the variable is not included. Moreover, even though patents and trademarks have been used as proxies for innovation, these measures may not be equally applicable for urban and rural regions. Since both patents and R&D intensity are geographically clustered towards urbanized regions (Feldman and Audretsch, 1999), it may imply that conventional technology-based measures are unable to capture the type of innovation activities that take place in (peripheral) rural regions. For rural regions, there are other dimensions of innovation that may be more relevant, such as improved services, new marketing, branding, and design methods, and new forms of business organizations (Wallet and Torre, 2016). These dimensions of innovation are not captured in the measure.

3.1 Explanatory variables and summary statistics

We focus on explanatory variables that measure related and unrelated variety calculated at the local and regional levels. Based on the approach in Jacquemin and Berry (1979) and Frenken, Van Oort, and Verburg (2007), this paper uses four-digit NACE codes (industrial classification as defined in Revision 2 used by Eurostat) and the entropy approach in constructing the measures. To better account for local effects, we use postcode areas as the reference in calculating the measures. Measures of related and unrelated variety are thus calculated using the local share of employees that work within industries at the lowest four-digit NACE code level that share the same higher two-digit level in the following way; let S_g denote the two-digit NACE level where $g = 1, ..., G$. Moreover, E_g denotes the share of total employment in each local area that belong to the same 2-digit level. Let E_{ig} denote the share of employees working in four-digit industry I, where $i = 1, ..., I$, where E_{ig} is measured as the share of employment in the respective two-digit industry g. Hence, unrelated variety in the industry measures the distribution of employees between two-digit industries for the postcode area. Using the entropy approach, unrelated variety is calculated as follows at the local (postcode) level:

$$(1) \qquad\qquad UV = -\sum_{g=1}^{G} E_g \ln E_g$$

The range of unrelated variety is from 0 to lnG, where zero is reached when all employees are working in the same two-digit industry. Similarly, the distribution of employees between four-digit industries within each two-digit industry is calculated as follows:

$$(2) \qquad\qquad H_g = -\sum_{i=1}^{I} E_{ig} \ln E_{ig}$$

The interpretation is identical to that for Equation 1 except that variety is measured within each two-digit industry instead of between two-digit industries. This implies that there is zero within variety when all employees in the two-digit industry g are working in the same four-digit industry i. Similarly, maximum variety for industry g, lnI, is achieved when there is an equal distribution of employees over all four-digit industries i. The information about the degree of within variety for each two-digit industry g, i.e., Hg, is weighted by the relative size of industry g. Summing over all g gives the entropy measure for related variety, regarding the postcode area and the NUTS 3 region:

---

[6] Out of the total sample of firms only 4 percent have at least one registered patent or trademark.

$$(3) \qquad RV = -\sum_{g=1}^{G} E_g H_g$$

Thus, increases in the values obtained by Equations 1 and 3 imply increases in unrelated and related variety. Location quotients have also been calculated to indicate the influence of technologically related businesses located nearby. Since there is a high correlation between the quotients and the measure of related variety there are multicollinearity issues. Hence, only the latter is used in the estimations. Locational quotients have no significant influence on firm performance when calculated at the more aggregated NUTS 3 level.

Other predictors at the local or regional levels include population density (inhabitants per square kilometer), knowledge context (share of employees that work in high-tech manufacturing firms and knowledge intensive business services), and local competition. Population density is included to control for urbanization economies and to separate between the effects of related variety and the size of regions, as there is naturally a high correlation between the two (Frenken, Van Oort, and Verburg, 2007). Since population is unavailable at the postcode level, regional size is measured using density at the more aggregated NUTS 3 level. Competition is defined using the number of firms per employee (Glaeser et al., 1992) and is calculated at the local level:

$$(4) \qquad Competition = \left(\frac{F_{local}}{E_{local}} \right)$$

where F is the number of firms and E is the number of employees. This measure of competition is thought to reflect Porter (1996, 2008) externalities through which local competition stimulates innovation and, hence, economic performance as firms are pushed to differentiate production and, thereby, enhance product variety for their customers. Since a measure of human capital (share of population with a tertiary education) can only be attained at the NUTS 3 level, knowledge context is indicated by the local share of employees that work in Knowledge Intensive Business Services, KIBS, and High Tech Manufacturing Industries, *HTMF*, (Johansson, Johansson, and Wallin, 2015).[7]

The firm-specific control variables are firm size (number of employees), capital (tangible fixed assets), the number of registered companies in corporate group, and the number of subsidiaries, which follow standard definitions (ORBIS, SBS). Summary statistics are presented in Table 1 and variable definitions can be found in Table B1 in Appendix B. As can be seen from Table 1, some of the firm specific variables have negative and extreme values and skewed distributions (e.g., the dependent variables, firm size and access to capital). To handle skewed distributions and control for outlier presence the variables are log-transformed, and we trim the 1st and 99th percentiles of the dependent variables to avoid extreme leverage points.

3.2 Urban-rural typology

To address regional heterogeneity in the outcomes, this paper applies the EDORA structural typology. This typology is based on the findings of the EDORA project (Copus and Noguera, 2010), which was designed to analyze the process of rural differentiation to better understand how European cohesion policy (EU 2020) can enable rural regions to build on their potentials. EDORA findings support place-based approaches put forward in the Barca Report (Barca, 2009). They show that rather than becoming more uniform in character, rural Europe

---

[7] See Table 1A in Appendix A for a definition of HTMF and KIBS firms in NACE Rev. 2 codes.

Table 1. Summary Statistics							
Variables	Min.	Max.	Mean	Median	Std. Dev.	Skewness	Kurtosis
Dependent variables							
Turnover/employee	-11,308	950,302	289.98	125.77	2,797.38	169.94	42,262
ROE	-1,000	999.00	9.11	12.98	102.23	-2.54	30.78
Firm-level explanatory variables							
Firm size	0.00	376,025	27.00	28.68	1,123.36	200.87	48,185
Capital	-1,859	81,115,001	2525.39	28.68	163,252.04	329.63	136,433
Number of companies in corporate group	0.00	4,411	13.71	0.00	140.78	20.87	546.75
Number of recorded subsidiaries	0.00	1,845	0.41	0.00	4.61	171.83	45,718
Local-level predictors							
Related variety	0.00	1.68	0.50	0.47	0.33	0.35	2.44
Unrelated variety	0.00	2.93	1.08	1.967	0.21	-0.84	3.61
Competition	0.00	1.00	0.10	0.13	0.12	0.16	0.43
Share HTMF, KIBS	0.00	1.00	0.20	0.41	0.06	0.10	0.33
Regional-level predictors							
Related variety	0.06	1.55	1.08	1.09	0.24	0.59	3.46
Unrelated variety	0.35	3.66	3.34	3.45	0.31	-3.23	2.99
Population density	2.00	21,467	843.47	900.55	2,840.12	6.55	1,002.60
Urban-rural typology (EDORA)							
EDORA 1	0.00	1.00	0.43	-	0.497	-	-
EDORA 2	0.00	1.00	0.02	-	0.156	-	-
EDORA 3	0.00	1.00	0.37	-	0.471	-	-
EDORA 4	0.00	1.00	0.01	-	0.115	-	-
EDORA 5	0.00	1.00	0.17	-	0.383	-	-

Note: Financial variables are reported in thousands of €.

is, in many ways, increasingly diversifying. A summary of the characteristics of the five regional types is presented below.[8]

Predominantly urban. Regions (at the NUTS 3 level) that are predominantly urban and located in the capital regions and in the larger city regions (e.g., the area delimited by London-Paris-Milan-Munich-Hamburg-London).

Agrarian regions. Regions in which the primary sector plays a large role in the regional economy and accounts for an above average share of gross value added and employment. These regions are mainly concentrated in the eastern and southern parts of the EU and are classified as low performers, in terms of various performance indicators. They also show many of the characteristics of socioeconomic depletion.

Consumption countryside regions. Includes regions in which the primary sector is less important and in which countryside public goods form the basis for the regional economy. This is reflected in indicators related to tourism, recreation activities, and to the supply of natural assets as well as the role of small-scale diversified forms of farming. These regions are classified as high performers.

Diversified regions with a strong secondary sector. Includes regions which do not fulfil the criteria for the first two types, and in which manufacturing accounts for a higher share of Gross Value added (GVA) than market services. Regions that belong to this category also tend to be low performers because of declining manufacturing industries.

Diversified regions with a strong market service sector. Contains regions which did not fulfil the criteria for the first two categories, and in which market services accounts for a higher share of GVA than manufacturing. Like consumption countryside regions, this category of regions is also classified as high performing. The EDORA typology is illustrated in Copus et al. (2011).

3.3 Empirical model

As discussed, firm performance likely depends on factors besides their inherent quality characteristics in terms of various interdependencies between local and regional factors. To address these interdependencies, we apply a nested multilevel model to mitigate unobserved heterogeneity at the local and regional level. Firms' postcodes and NUTS 3 codes are used to impose the hierarchical structure. Following the multilevel literature (Besag, York, and Mollié, 1991; Langford et al. 1999), the effects of unobserved heterogeneity are incorporated using the following two-level model:

$$(5) \qquad \ln Y_{i(zn)} = \beta_0 + \sum_{z=1}^{r} \beta_z x_{zi(zn)} + \zeta_{(zn)} + \zeta_n + \mu_{iz}$$

where $Y_{i(zn)}$ stands for the dependent variable (either the log of turnover per employee or ROE) of firm i, nested in postcode area z and NUTS 3 region n. Since the method allows the use of covariates at any level in the proposed hierarchy, the r covariates contain the firm characteristics and β_z is the corresponding coefficient of each covariate. The random intercepts are measured at local ($\zeta_{(zn)}$) and regional (ζ_n) levels and capture unobserved heterogeneity at those levels. The random intercepts are assumed independent (given the covariates) and normally distributed with mean zero and constant variance and can therefore be directly estimated (Goldstein, 2005).

Via the multilevel model, the potential influence of each spatial level is measured. This is particularly useful for dealing with any possible bias due to endogeneity (Skrondal and Rabe-

---

[8] The EDORA working papers (e.g. Copus and Noguera 2010) and the ESPON Scientific papers provide a comprehensive background and motivation of the opportunities and constraints that characterize each regional type

Hesketh, 2004) and some potential spatial autocorrelation (Bhat, 2000). Firm performance might be more correlated within a given postcode area or NUTS 3 region than strictly across an industry. Introducing varying intercepts captures unobserved heterogeneity at the different levels. Furthermore, as proposed by Mundlak (1988), among others, bias introduced by possible endogeneity can be controlled for via a clustered sample of means. Since the main interest lies in the variables measuring related and unrelated variety, means of these variables are included at the NUTS 3 level and the random intercept measured at this level provides an estimate of unobserved heterogeneity (Skrondal and Rabe-Hesketh, 2004).

4. REGRESSION RESULTS

Before introducing the explanatory variables into the model, the hierarchical structure of the data is examined by estimating variances for the random intercepts at the two levels, excluding the regressors. This provides information on how the proposed hierarchical structure relates to firm performance and how much of the variance can be attributed to the two geographical levels. The results of estimating the unconditional model, i.e., Equation 5 including only the random components, show that the between-level heterogeneity at a local and regional level is statistically significant and positive, implying that there is a sizeable level of variance across and within the imposed hierarchy. This can be described by the intraclass correlation coefficient (ICC). The ICC measures the degree of correlation among observations within a postcode area or a NUTS 3 region and is a useful tool to evaluate how much of the total variance in turnover can be assigned to these two different geographical levels. The ICC coefficient ranges from 0 to 1, where a value of 0 indicates that the grouping bears no information and 1 indicates that all units in the group are identical. Results show that the ICC is .023 for NUTS 3 regions and .121 for postcode areas nested in NUTS 3 regions. The low variance at the NUTS 3 level indicates that only a small part of the variance is explained by characteristics of NUTS 3 regions and most of the variance is explained by the more disaggregated postcode level. This indicates that the main factors that explain firm performance are firm attributes and attributes measured at the more disaggregated level, which is in line with the theoretical discussion about locally bounded effects. Still, the broader regional context is statistically significant, which suggests that the inclusion of NUTS 3 attributes is important. For brevity, we do not report the results from estimating the unconditional model, but they can be attained from the author on request.

Table 2 presents the results from estimating Equation 5 including the random intercepts that control for unobserved heterogeneity and results are presented in three model specifications. Specification A uses turnover per employee as the dependent variable and Specification B uses ROE, which enables us to address robustness and to strengthen the interpretation. Using the Wald χ^2 as a means for model selection, the specification with turnover per employee produces the best fit and is thus used in the analysis that follow. The third specification (C), with the log of turnover per employee as the dependent variable, enables us to examine whether related and unrelated diversity are robust to the inclusion of regional predictors (population density) and the clustered sample means of related and unrelated variety (Mundlak, 1988).[9]

From Specification C, the coefficient of related variety has a positive and significant coefficient indicating that a high degree of technological relatedness among firms is positively related to firm performance. The result is also robust to the inclusion of population density. As discussed, there is a high correlation between related variety and regional size as variety in firms and products require a certain regional size and market potential. The correlation between

---

[9] Model 3 has also been estimated using ROE as the dependent variable and results are comparable.

Table 2. Regression Results (multilevel model), Total Sample

Independent variable Dependent variables	A. Turnover per employee (ln)	B. ROE	C. Turnover per employee (ln)
Fixed effects			
Firm size (ln)	-0.148*** (0.001)	4.939*** (0.115)	-0.149*** (0.001)
Capital (ln)	0.157*** (0.001)	0.098*** (0.021)	0.158*** (0.001)
Number of companies in corp. group	0.001*** (7.62e-06)	0.002 (0.002)	0.001*** (7.62e-06)
Number of recorded subsidiaries	0.009*** (0.0003)	0.219 (0.200)	0.010*** (0.0003)
Local level predictors			
Related variety (ln)	0.121*** (0.008)	1.297** (0.579)	0.118*** (0.009)
Unrelated variety (ln)	-0.035*** (0.004)	0.048 (0.353)	-0.036*** (0.004)
Competition	0.5e-03*** (1.0e-03)	0.001 (0.001)	1.11e-03*** (1.29e-04)
Share HTMF, KIBS	0.067*** (1.14e-03)	0.030* (1.11e-03)	0.057*** (1.16e-03)
Regional level predictors			
Related variety (ln)	-	-	0.062 (0.053)
Unrelated variety (ln)	-	-	-0.004 (0.037)
Population density (ln)	-	-	0.075*** (0.008)
Industry	YES	YES	YES
Country	YES	YES	YES
Random effects			
Firm	0.942***	10086.96***	0.942***
Postcode	0.026***	13.382***	0.027***
NUTS 3	0.025***	8.116***	0.018***
ICC			
Regional	-	-	0.020*** (0.002)
Local | Regional	-	-	0.040*** (0.001)
Wald test	182,868.53***	4,236.49	182,996.20***
Number of firms	742,489	673,119	742,489

***, ** and * denote statistical significance at the 1, 5 and 10 percent levels. Standard errors are reported in parenthesis below each coefficient. Dependent variables: turnover per employee (Specification A and C) and ROE (Specification B). Condition numbers are used to validate the results and detect the presence of multicollinearity in the estimations. The condition numbers associated with the estimations are all below 30, indicating that multicollinearity does not prevent the interpretation of results (calculated as the square root of the largest eigenvalue divided by the smallest) (Wheeler and Tiefelsdorf, 2005).

population density and related variety is relatively low in our sample (.28), probably because density is measured at the more aggregated NUTS 3 level.[10] Results also show that the coefficient of related variety, when measured at the regional level, is insignificant. A possible interpretation is that knowledge spillovers occur at the local rather than regional level, e.g., that there is attenuation in space (Van Oort 2007; Rosenthal and Strange 2008). There is also a difference in the influence of related and unrelated variety on firm performance, the latter being negative and significant in the fully specified model, which could indicate that important

---

[10] A correlation matrix is presented in Appendix C.

knowledge spillovers come from interactions among firms in related industries rather than from interactions outside the own industry as discussed in the theoretical section. There is also the possibility that the measure of regional size (population density) at the NUTS 3 level is too general and unable to capture size effects. The coefficient reflecting competition is thought to capture the degree to which local competition influences firm performance. Specifically, that a high degree of local competition motivates firms to become more innovative in order to survive and develop, which enhances their productivity. This is shown to be positive and significant only in a statistical sense in the fully specified model. The coefficient reflecting the local share of employees that works in high-tech manufacturing and KIBS firms is positive and significant in the fully specified model. A positive association could indicate that access to external knowledge is important, which arises due to a large share of highly educated individuals. This result would thus lend support to the findings in prior empirical studies on the influence of external knowledge (e.g., Moretti, 2004).

Table 2 also presents the estimated ICC coefficients, which is a tool in the estimation of multilevel models that measure the degree of correlation among observations and how much of the variation that can be assigned to the different geographical levels. From the ICC coefficients, a significant part of firm performance is explained by unobserved heterogeneity at the local and regional level, although the findings also suggest that the local level explains a larger share of the variance in the dependent variable compared to the more aggregated NUTS 3 level. Moreover, since both random intercepts are statistically significant, there appears to be unobserved heterogeneity at the local and regional level that the explanatory variables cannot fully explain, but that is instead embedded in the value of the intercepts.

4.1 Regional heterogeneity

The discussion so far has been focused on the influence of measures of technological relatedness and industrial diversity on firm performance. Another central question addressed in this paper is whether the relative importance of relatedness and diversity vary across the urban-rural range. In accordance with prior research, we hypothesize that indicators of smart specialization may not affect firm performance homogenously across regions of different types. This follows the argumentation in McCann and Ortega-Argilés (2015) among others, which point to lack of scale and critical mass as the main obstacles for these externalities to materialize. To empirically assess this question, the sample is split into five regional categories as described in Section 3.2. Results from these case-wise estimations are presented in Table 3.

Findings indicate that the coefficient of related variety, measured at the local level, is positively and statistically related to firm performance in all regional types except for diversified regions with a strong market sector.[11] These results are interesting as they point to the presence of heterogenous outcomes across regional types and that urban regions and regions with a strong service sector can benefit from a higher variety of related industries. Results indicate no statistical relationship between firm performance and technological relatedness for firms located in regions with a strong secondary sector. This category includes regions in which manufacturing accounts for a higher share of GVA than market services. Regions that belong to this category also tend to be low performers because of declining manufacturing industries. These results could also indicate the presence of industry differences and that there is a need to differentiate the analysis as there may exist differences between manufacturing regions compared to diversified regions with a large service sector. Similar arguments and results are presented in Mameli, Boschma, and Iammarino (2012), who found no evidence of similar externalities in manufacturing and tertiary sectors with regards to employment growth. Similar results are found in Combes (2000), and Bishop and Gripaios (2010).

---

[11] This is confirmed by a t-test of equality of coefficients.

Table 3. Regression Results (multilevel model) across Regional Types
(Dependent Variable: Log of Turnover per Employee)

Variable	Urban	Agrarian	Consumption countryside	Diversified strong secondary sector	Diversified strong market sector
Fixed effects					
Firm size (ln)	-0.162***	-0.217***	-0.113***	-0.339***	-0.151***
	(0.001)	(0.007)	(0.002)	(0.011)	(0.002)
Capital (ln)	0.157***	0.182***	0.152***	0.157***	0.157***
	(0.001)	(0.003)	(0.001)	(0.006)	(0.001)
#. companies in corp. group	0.001***	0.0002**	0.0001***	0.001***	0.0001***
	(1.10e-04)	(0.0001)	(0.00001)	(9.6e-04)	(0.00001)
#. recorded subsidiaries	0.007***	0.091***	0.034***	0.053***	0.028***
	(0.0003)	(0.006)	(0.001)	(0.005)	(0.001)
Local-level predictors					
Related variety (ln)	0.174***	0.123***	0.107***	0.069	0.066***
	(0.015)	(0.035)	(0.013)	(0.059)	(0.015)
Unrelated variety (ln)	-0.093***	-0.012	-0.009	0.006	0.001
	(0.009)	(0.022)	(0.006)	(0.033)	(0.007)
Competition	2.22e-04	-0.001	0.001***	0.004**	0.0001*
	(1.45e-04)	(0.001)	(3.40e-04)	(0.0008)	(7.66e-04)
Share HTMF, KIBS	0.042*	0.001	0.023*	0.003	0.002
	(1.10e-03)	(0.010)	(0.001)	(0.011)	(0.002)
Regional-level predictors					
Related variety (ln)	0.063	0.276	0.082	0.100	-0.107
	(0.122)	(0.210)	(0.082)	(0.437)	(0.095)
Unrelated variety (ln)	-0.089	0.205**	0.001	0.0004	0.020
	(0.092)	(0.104)	(0.043)	(0.407)	(0.073)
Population density (ln)	0.056**	0.050	0.046***	0.059	0.066***
	(0.025)	(0.032)	(0.017)	(0.075)	(0.024)
Industry	YES	YES	YES	YES	YES
Country	YES	YES	YES	YES	YES
Random effects					
Firm	1.041***	1.101***	0.867***	1.237***	0.741***
Postcode	0.041***	0.023***	0.020***	0.008**	0.012***
NUTS 3	0.020***	0.002*	0.012**	0.007	0.010***
ICC					
Regional	0.003**	0.005**	0.010*	0.001*	0.006*
Local | Regional	0.022***	0.016***	0.011***	0.023**	0.028***
Wald test	82,367.01***	7,997.41***	54,514.57***	23,43.09***	42,321.91***
Number of firms	351,424	21,930	217,793	9,249	142,093

***, and ** denote statistical significance at the 1 and five percent levels. The models are estimated using Specification C in Table 2. Standard errors are reported in parenthesis below each coefficient

While related variety is thought to reflect the presence of shared knowledge bases and complementarities that may encourage externalities and knowledge spillovers, unrelated variety is thought to reflect diversity among sectors that are not interrelated in terms of shared competences. Results show that the coefficient of unrelated variety, measured at the local level, is negative and statistically significant in urban regions, but not statistically significant for all other regional types. Nonetheless, unrelated variety, which is measured only at the more aggregated regional level, is significant and positive in agrarian regions. Thus, knowledge spillovers and externalities that arise from a diversified firm environment are positively related to firm performance in rural (agrarian) regions, but negatively in urban regions. These finding suggest the presence of a portfolio effect in agrarian regions, i.e., that more diversified rural areas with unrelated sectors have better performing firms as they are more protected against external shocks in demand (Boschma, 2014).

Urbanization economies, which are proxied by population density, have a statistically significant positive effect on firm performance in urban and diversified regions, but no statistically significant effect in agrarian regions and regions with a strong secondary sector. Again, these results indicate the presence of heterogenous outcomes across regional types and that urban regions and regions with a strong service sector can benefit from urbanization economies. It is important to note that there are relatively few firms in the regions with a strong secondary sector which calls for a careful interpretation of the results.

The coefficients reflecting firm-level control variables are in line with expectations in that access to subsidiaries and capital is positively associated with firm performance. The effects of firm's size on firm performance or productivity is ambiguous. A negative association, as found here, may reflect that small firms have better managerial capacity and generally exhibit a higher profit rate and are more motivated to undertake actions that make them more productive compared to large firms, as discussed in Acs and Audretsch (1990).

4.2 Unobserved heterogeneity and contextual variation

The results show that even after introducing the covariates and estimating the model across the typology, a substantial part of firm performance differentials is explained by unobserved heterogeneity at the firm, local, and regional levels. The estimated variation of the random intercept at the local level is significant in most of the estimations implying that there is local heterogeneity that explanatory variables are unable to fully account for. Similarly, the estimated variation of the random intercepts at regional level are significant across the estimations, implying that there is unobserved heterogeneity at the regional level that the explanatory variables cannot fully explain. There is a restricted possibility to measure spatial dependencies at the firm level since the exact location of each firm is unknown. Bhat (2000) shows that a hierarchical structure with levels of geographical aggregation reduces spatial heterogeneity and spatial autocorrelation. To find out whether spatial autocorrelation exists, firm variables have been aggregated to the local (postcode level) and tested using Moran's I and Equation 5, without the typology. Moran's I is not statistically significant (1.50, *p*-value .05) and the null hypothesis that there is no spatial autocorrelation can be accepted at the five percent level.

5. CONCLUSIONS

The concept of smart specialization is part of the EU 2020 growth strategy and has been introduced as a response to the low growth rates of innovation and productivity among regions in Europe. The purpose of this paper is to empirically address and bring together some of the concepts

that are identified as central behind the smart specialization logic (Naldi et al. 2015). The paper adds to the existing literature by differentiating the analysis for various types of rural regions and by testing the presence of externalities on firm performance across the urban-rural range. The analysis controls for spatial autocorrelation in the dat, and considers different geographical levels of unobserved heterogeneity. The analysis focuses on proximity to both related and unrelated economic activities, alongside other locational and firm specific attributes, which are thought to reflect aspects of the smart specialization logic (McCann and Ortega-Argilés, 2015). A sample of firms across Europe are included in a nested multilevel model that differentiate between firm, local, and regional effects. Having access to these data enables an empirical analysis that address locally bounded effects and the presence of urban-to-rural heterogeneity. Micro-level analyses that include local and regional diversity at the European level are still limited compared to the number of studies the focus on these relationships at the more aggregate regional level. Moreover, while most previous studies with this focus use firm- or employee counts or more general measures of industry composition, this study uses a more detailed approach.

The findings in this paper indicate the presence of heterogenous outcomes across regional types. A positive association between a higher variety of related industries and firm performance is found in urban regions and regions with a strong service sector, that generally have a broad diversification. Results indicate no significant association for firms located in regions with a strong secondary sector. These results may be reflective of industry differences meaning that there is a need to differentiate the analysis to take regional preconditions into account. It is also found that unrelated variety, thought to reflect diversity among sectors that are not interrelated in terms of shared competences, is negatively related to firm performance in urban regions, but positively related in agrarian regions. This can indicate a portfolio effect in agrarian regions, i.e., that more-diversified rural areas are likely to be more protected against external shocks in demand and, thereby, enhances firm performance (Boschma, 2014).

Although this paper contributes to the literature by addressing regional heterogeneity in the studied outcomes considering a diverse set of rural regions, which to the author's knowledge has not been done this way before, the analysis is unable to disentangle several important relationships in relation to the smart growth logic. One important issue that limits the scope of the analysis is the obvious challenges in creating indicators more suitable for measuring innovation and knowledge in rural regions (Torre and Wallet, 2016). Another issue is that rural peripheral regions may have varying potential when it comes to their access to resources and social infrastructure, which may significantly affect their growth potential. Some firms might, for example, have a potential to achieve smart specialization through exploiting local amenities (Dissart and Marcouiller, 2012), developing creative economies (McGranahan, Wojan, and Lambert, 2011), and other resources to build specialized links to urban supply and demand. The implication is that externalities that are not directly related to the smart growth logic can be expected to play a significant role and create different preconditions for regional growth (Naldi et al. 2015), which this analysis is unable to capture. Hence, there is a need for further studies that decompose not only the meaning of space but also the different components that are included in the broad concept of place-specific characteristics and how they can be expected to influence the growth potential for different rural regions. It is important to note that there are relatively few firms in the regions with a strong secondary sector which calls for a careful interpretation of the results. Further studies that can extend the empirical analysis to include the total sample of firms across regional subcategories will thus contribute to extend the validity of these findings to a more general context.

REFERENCES

Acs, Zoltan J. and David B. Audretsch. (1990) "The Determinants of Small-firm Growth in US Manufacturing," *Applied Economics*, 22, 143–153.

Acs, Zoltan J., Luc Anselin, and Attila Varga. (2002) "Patents and Innovation Counts as Measures of Regional Production of New Knowledge," *Research Policy*, 31, 1069–1085.

Asheim, Bjorn T., Ron Boschma, and Philip Cooke. (2011) "Constructing Regional Advantage: Platform Policies Based on Related Variety and Differentiated Knowledge Bases," *Regional Studies*, 45, 893–904.

Audretsch, David B. and Maryann P. Feldman. (2004) "Knowledge Spillovers and the Geography of Innovation," *Handbook of Regional and Urban Economics*, 4, 2713–2739.

Barca, Fabrizio. (2009) "Agenda for a Reformed Cohesion Policy; A Place-based Approach to Meeting European Union Challenges and Expectations," European Communities. Available online in December 2016 at: http://www.europarl.europa.eu/meetdocs/2009_2014/documents/regi/dv/barca_report_/barca_report_en.pdf

Barca, Fabrizio, Philip McCann, and Andrés Rodríguez-Pose. (2012) "The Case for Regional Development Intervention: Place-based Versus Place-neutral Approaches," *Journal of Regional Science*, 52, 134–152.

Besag, Julian, Jeremy York, and Annie Mollié. (1991) "Bayesian Image Restoration, with Two Applications in Spatial Statistics," *Annals of the Institute of Statistical Mathematics*, 43, 1–20.

Bhat, Chandra. R. (2000) "A Multi-level Cross-classified Model for Discrete Response Variables," *Transportation Research Part B: Methodological*, 34, 567–582.

Bilbao-Osorio, Benat and Andrés Rodríguez-Pose. (2004) "From R&D to Innovation and Economic Growth in the EU," *Growth and Change*, 35, 434–455.

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

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

_____. (2014) "Constructing Regional Advantage and Smart Specialisation: Comparison of Two European Policy Concepts," *Science Regionali*, 18, 51–68.

Camagni, Roberto and Roberta Capello. (2013) "Regional Innovation Patterns and the EU Regional Policy Reform: Toward Smart Innovation Policies," *Growth and Change*, 44, 355–389.

Combes, Pierre-Philippe. (2000) "Economic Structure and Local Growth: France, 1984–1993," *Journal of Urban Economics*, 47, 329–355.

Copus, Andrew and Joan Noguera. (2010) "A Typology of Intermediate and Predominantly Rural NUTS 3 Regions." *European Observation Network for Territorial Development and Cohesion (ESPON) Applied Research Project* 2013/1/2: Luxembourg.

Dijkstra, Lewis and Hugo Poelman. (2008) "Remote Rural Regions: How Proximity to a City Influences the Performance of Rural Regions." *Regional Focus Research Paper* 01/2008.

Dissart, Jean-Christophe and David W. Marcouiller. (2012) "Rural Tourism Production and the Experience-scape," *Annals of Tourism*, 17, 691–704.

Duranton, Gilles and Diego Puga. (2004) "Micro-foundations of Urban Agglomeration Economies," in J. Vernon Henderson and Jacques-François Thisse (eds.), *Handbook of Regional and Urban Economics*. Elsevier: Amsterdam.

Dewenter, Kathryn L. and Paul H. Malatesta. (2001) "State-owned and Privately Owned Firms: An Empirical Analysis of Profitability, Leverage, and Labor Intensity," *American Economic Review*, 320–334.

European Commission. (2010) "Europe 2020: A European Strategy for Smart, Sustainable and Inclusive Growth," Publications Office of the European Union: Brussels.

Feldman, Maryann P. and David B. Audretsch. (1999) "Innovation in Cities: Science-based Diversity, Specialization and Localized Competition," *European Economic Review*, 43, 409–429.

Feldman, Maryann P. and Dieter F. Kogler. (2010) "Stylized Facts in the Geography of Innovation," *Handbook of the Economics of Innovation*, 1, 381–410.

Foray, Dominique and Bronwyn H. Hall. (2011) "Smart Specialization: From Academic Idea to Political Instrument, the Surprising Career of a Concept and the Difficulties Involved in its Implementation." *MTEI Working Paper* 2011-001. EPFL: Lausanne, Switzerland.

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

Glaeser, Edward L., H. D. Kallal, Jose A. Scheinkman, and Andrei Shleifer. (1992) "Growth in Cities," *Journal of Political Economy*, 100, 1126–1152.

Goldstein, Harvey. (2005) *Multilevel Statistical Models.* Wiley: Hoboken, NJ.

Grossman, Gene M. and Elhanan Helpman. (1990) "Trade, Innovation, and Growth," *American Economic Review*, 80, 86–91.

Jacobs, Jane. (1969) *The Economy of Cities.* Vintage: New York.

Jacquemin, Alexis P. and Charles H. Berry. (1979) "Entropy Measure of Diversification and Corporate Growth," *Journal of Industrial Economics*, 27, 359–369.

Jensen, Morten B., Börje Johnson, Edward Lorenz, and Bengt Åke Lundvall. (2007) "Forms of Knowledge and Modes of Innovation," *Research Policy*, 36, 680–693.

Johansson, Börje, Hand Lööf, and Amy R. Olsson. (2010) "Firm Location, Corporate Structure, R&D Investment, Innovation and Productivity," in Charlie Karlsson, Roger R. Stough, and Börje Johansson (eds.), *Entrepreneurship and Innovations in Functional Regions*. Edward Elgar Publishing: Cheltenham, pp. 77–106.

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

Langford, Ian H., Alastair H. Leyland, Jon Rasbash, and Harvey Goldstein. (1999) "Multilevel Modelling of the Geographical Distributions of Diseases," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 48, 253–268.

McGranahan, David, Timothy R. Wojan, Dayton M. Lambert. (2011) "The Rural Growth Trifects: Outdoor Amenities, Creative Class and Entrepreneurial Context," *Journal of Economic Geography*, 11, 529–557.

Mameli, Hague F., Simona Iammarino, and Ron Boschma. (2012) "Regional Variety and Employment Growth in Italian Labour Market Areas: Services Versus Manufacturing Industries." *Papers in Evolutionary Economic Geography* 12.03. Utrecht University: Utrecht, Netherlands.

Marshall, Alfred. (1920) *Principles of Economics*. Macmillan: London.

McCann, Philippe, and Raquel Ortega-Argilés. (2015) "Smart Specialization, Regional Growth and Applications to European Union Cohesion Policy," *Regional Studies*, 49, 1291–1302.

Moretti, Enrico. (2004) "Workers' Education, Spillovers, and Productivity: Evidence from Plant-level Production Functions" *American Economic Review*, 94, 656–690.

Mundlak, Yair. (1988) "Endogenous Technical Change and the Measurement of Productivity," Chapter 11 in Susan M. Capalbo and John M. Antle (eds.), *Agricultural Productivity: Measurement and Explanation*. Resources for the Future: Washington, DC., pp. 316–331.

Naldi, Lucia, Pia Nilsson, Hans Westlund, and Sofia Wixe. (2015) "What is Smart Rural Development?," *Journal of Rural Studies*, 40, 90–101.

Neffke, Frank, Martin Henning, and Ron Boschma. (2011) "How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions," *Economic Geography*, 87, 237–265.

Noseleit, Florian. (2015) "The Role of Entry and Market Selection for the Dynamics of Regional Diversity and Specialization," *Regional Studies*, 49, 76–94.

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

Porter, Michael E. (1996) "Competitive Advantage, Agglomeration Economies, and Regional Policy," *International Regional Science Review*, 19, 85–90.

_____. (2008) *On Competition*. Harvard Business Press: Cambridge, Massachusetts.

Romer, Paul M. (1990) "Endogenous Technological Change," *Journal of Political Economy*, 98, 71–102.

Ribeiro, Samuel P., Stefano Menghinello, and Koen DeBacker. (2010). "The OECD ORBIS Database: Responding to the Need for Firm-level Micro-data in the OECD." *OECD Statistics Working Papers Series Working Paper 2010/1*. OECD Statistics Directorate: Paris.

Rodrigues-Pose, Andrés. (2001) "Is R&D Investment in Lagging Areas of Europe Worthwhile? Theory and Empirical Evidence," *Papers in Regional Science*, 80, 275–295.

Rosenthal, Stuart S. and William C. Strange. (2003) "Geography, Industrial Organization, and Agglomeration," *Review of Economics and Statistics*, 85, 377–393.

_____. (2008) "The Attenuation of Human Capital Spillovers," *Journal of Urban Economics*, 64, 373–389.

Skrondal, Anders and Sophia Rabe-Hesketh. (2004) *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*. Chapman & Hall/CRC: Boca Raton, FL.

Torre, Andre and Frederic Wallet. (2016) *Regional Development in Rural Areas. Analytical Tools and Public Policies*. Springer: Berlin, Germany.

Tödtling, Franz and Michaela Trippl. (2005) "One Size Fits All? Towards a Differentiated Regional Innovation Policy Approach," *Research Policy*, 34, 1203–1219.

Van Oort, Frank G. (2007) "Spatial and Sectoral Composition Effects of Agglomeration Economies in the Netherlands," *Papers in Regional Science*, 86, 5–30.

Wheeler, David, and Michael Tiefeldorf. (2005) "Multicollinearity and Correlation among Local Regression Coefficients in Geographically Weighted Regression," *Journal of Geographical Systems*, 7, 161–187.

APPENDIX

Section	Title	Four-digit NACE Code	Mean	Std. Dev.
Table A1. Industrial composition of sample, NACE Rev. 2				
A	Agriculture, forestry and fishing	0111-0322	0.017	0.129
B	Mining and quarrying	0510-0990	0.002	0.045
C	Manufacturing (excluding HTMF)	1011-2599 2711-3320	0.144	0.351
C	Manufacturing, high tech manufacturing firms, HTMF	2611-2680	0.004	0.070
D, E, F	Electricity, water supply and construction	3511-4399	0.184	0.387
G	Wholesale and retail trade	4511-4799	0.266	0.441
H, I	Accommodation, transportation and food services	4910-5320 5510-5630	0.058	0.235
J	Information and communication (excluding KIBS)	5811-6190	0.013	0.114
M	Professional, scientific and technical activities (excluding KIBS)	7410-7500	0.012	0.111
J, M	Knowledge intensive business services, KIBS	6201-6399 6910 -7320	0.133	0.340
K, L	Financial, insurance and real estate activities	6411-6832	0.074	0.262
O-S	Other service activities (Public, education, health and entertainment)	7711-9609	0.099	0.299
T, U	Other (activities of households as employers, activities of extraterritorial organizations and bodies)	9700-9900	0.0001	0.010

Table B1. Variable Definitions

Variable	Definition
Dependent variables	
Turnover/employee (ln)	The natural logarithm of; operating revenue (net sales plus other operating revenues and stock variations) divided by the total number of full-time employees.
ROE	Net income divided by shareholder's equity.
Firm-level explanatory variables	
Firm size (ln)	The natural logarithm of total number of full-time employees.
Capital (ln)	The natural logarithm of the value of the firm's tangible assets (e.g., buildings and machinery)
Nr. companies in corporate group	Total number of companies included in corporate group
Nr. subsidiaries	Total number of subsidiaries of the firm
Local- level explanatory variables	
Related variety (ln)	The natural logarithm of Equation 3 calculated with respect to the number of employees at the local (postcode) level.
Unrelated variety (ln)	The natural logarithm of Equation 1 calculated with respect to the number of employees at the local level.
Competition	Calculated according to Equation 4.
Share HTMF, KIBS	Share of firms in local area that are defined as either High Tech Manufacturing Firms or Knowledge Intensive Business Services.
Regional-level explanatory variables	
Related variety (ln)	The natural logarithm of Equation 3 calculated with respect to number of employees in the region (NUTS 3).
Unrelated variety (ln)	The natural logarithm of Equation 1 calculated with respect to the number of employees in the region.
Population density (ln)	The natural logarithm of number of inhabitants per square kilometer in the region.

Table C1. Correlation matrix, key variables												
Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Turnover/employee (*ln*)	1.000											
2. ROE	.103	1.000										
3. Firm size (*ln*)	.026	.026	1.000									
4. Capital (*ln*)	.243	-.035	.552	1.000								
5. Nr. companies in corp.	.056	.004	.133	.083	1.000							
6. Nr. subsidiaries	.059	-.005	.183	.153	.065	1.000						
7. Unrelated variety (local) (*ln*)	-.025	-.002	-.019	.014	-.030	-.026	1.000					
8. Related variety (local) (*ln*)	.044	.002	.036	.045	-.014	-.007	.574	1.000				
9. Competition	.022	.001	.072	.047	.029	.104	-.112	-.042	1.000			
10. Related variety (regional) (*ln*)	.016	-.008	.032	.105	-.033	-.011	.286	.337	-.034	1.000		
11. Unrelated variety (regional) (*ln*)	-.002	-.008	-.023	.022	-.024	-.004	.194	.152	-.046	.576	1.000	
12. Population density (*ln*)	.057	.012	.069	.001	.046	.032	.147	.288	.040	-.034	-.016	1.000