ORIGINAL PAPER

Heterogeneous relatedness and firm productivity

S. Stavropoulos[1,2] · F. G. van Oort[3,4] · M. J. Burger[1,2]

Abstract

In this manuscript, we relate regional structural composition—related and unrelated variety—to firm-level productivity in European regions, applying a Cobb–Douglas production function framework and using firm-, industry- and regional-level mixed hierarchical (multilevel) models. Our analyses indicate that regional-related variety has a positive impact on firm productivity in European regions, especially for firms in high-tech and medium-tech regions. These outcomes have implications for European policies on competitiveness as firms embedded in regions without these technological and institutional circumstances are systematically worse off in terms of productivity, and catching-up is not obvious for such regional economies.

JEL Classification R11 · O18 · C31

1 Introduction

In economic geography, development economics and management studies, cognitive relatedness (also known as technological proximity, cognitive proximity, technological closeness, non-tradable capabilities similarity or skill relatedness) of firms, industries and regions is increasingly suggested to be an important precondition for

✉ S. Stavropoulos
 stavropoulos@ese.eur.nl

 F. G. van Oort
 vanoort@ese.eur.nl

 M. J. Burger
 mburger@ese.eur.nl

1 Department of Applied Economics, Tinbergen Institute, Erasmus University Rotterdam, Rotterdam, The Netherlands

2 Erasmus Happiness Economics Research Organization, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

3 Department of Applied Economics and Institute for Housing and Urban Studies (IHS), Erasmus University Rotterdam, Rotterdam, The Netherlands

4 Department of Economic Geography, Utrecht University, Utrecht, The Netherlands

    ∅ Springer

economic diversification, industry branching and economic development in countries, cities and regions (Boschma 2005; Frenken et al. 2007; Hidalgo et al. 2007; Neffke and Henning 2013). As was already noted by Smith (1776, p. 430), more developed societies tend to be more diversified in their economies: *Though in a rude society there is a good deal of variety in the occupations of every individual, there is not a great deal in those of the whole society (…). In a civilized state, on the contrary, though there is little variety in the occupations of the greater part of individuals, there is an almost infinite variety in those of the whole society.*

In this vain, the literature on economic development of countries has focused on the impact of changing patterns of branching and economic complexity on economic growth (Hausmann and Hidalgo 2011; Hidalgo and Hausmann 2009). However, as pointed out by Saviotti and Pyka (2004, 2010), diversification is not only a determinant of growth, but also a result of growth. Hence, diversification and economic growth are co-dependent, where diversification is a prerequisite for growth, while simultaneously growth is needed to stimulate demand and changes in demand. Examining the causality of the relationship between diversification and economic growth, Dietrich (2012) found that growth structural change (irrespective of being measured by employment and value added) induces economic growth, while economic growth slows down structural change in the short run but accelerates change in the longer run, where the aggregate effect depends on whether structural change is measured by employment or value added.

Along these lines, the economic geography literature on related and unrelated variety has studied the impact of cognitive relatedness across industries to predict both economic diversification and growth opportunities of regional economies (related variety) and regional portfolio circumstances that mitigate economic shocks (unrelated variety) (Frenken et al. 2007). The regional and regional-sector levels of analyses comply most naturally to the competitive cluster conditions for specialized or diversified growth. Castaldi et al. (2015) extend the latter discussion to the question under what conditions regions and countries with unrelated variety can also yield product innovation (especially radical ones), and why some regions and countries manage to diversify into (cognitively) unrelated industries.

The broader benefits of diversification may be captured by either the accessibility of knowledge in other than localized networks, especially via international relations of migration, cooperation and multinational enterprises (Content and Frenken 2016; Boschma and Iammarino 2009), or in the absorptive, learning and entrepreneurial capacities of firms that are beyond path-dependence able to internalize previously unknown knowledge from outside the firm (Fritsch and Kublina 2018). While the former may still be researched by aggregate data (on MNE's, migrants and cooperation), the latter actually requires studies to focus on firm-level performance and mechanisms of interaction (of agglomerated regions, sectors and clusters, e.g., transfer mechanisms of knowledge and knowledge spillovers)—something that is still underdeveloped in the literature (Burger et al. 2011; Van Oort et al. 2012; Knoben et al. 2016). According to Van Oort (2015), such a shift in focus from the regional to the firm level is also needed to capture the sectoral and firm-, entrepreneurial- and skilled-worker-level heterogeneity that actually causes differences in regional development. More specifically, it helps us to better understand for whom related

and unrelated variety matter and under which circumstances. This paper explores under which heterogeneous contextual conditions cognitive relatedness ("related and unrelated variety") is associated with firm-level productivity. Here, it is believed that related variety is positively related to firm productivity through facilitating spillovers, knowledge spillovers and local outsourcing (Castaldi et al. 2015), while high levels of unrelated variety are expected to be negatively associated with firm productivity because firms operating in very different industries lack complementarities in factor inputs and there are low levels of knowledge spillovers (Aarstad et al. 2016).

From the firm-context literature, it is clear that it is important to take multilevel heterogeneity into account when analyzing productivity (Syverson 2011). Both the relatedness (leading to externalities) literature and the firm performance literature hint at their potential interplay, although this has not been substantiated yet in empirical analyses. Despite the complex and sophisticated methods of conceptually linking region and sector-related spillovers with growth and cities, an ever-growing body of empirical literature on external economies of scale remains inconclusive on the exact agglomeration circumstances that optimally enhance growth (Melo et al. 2009; De Groot et al. 2015). Van Oort et al. (2012) argue that the missing link that leads to this ambiguity in this research results on agglomeration economies may be the hypothesized, yet often missing relationship between heterogeneous agglomeration economies and individual firm performance. The related and unrelated variety hypotheses were originally embedded in this agglomeration literature, suggesting that related variety as a measure of cognitive proximity captures growth and productivity circumstances (externalities) of regions and cities better than other measures of localization and urbanization (see Content and Frenken 2016 for an overview of the burgeoning literature confirming this).

Although the measurement of relatedness across regions advanced over the last decade from sectoral branching (Frenken et al. 2007; Hidalgo et al. 2007) to include co-location (Ellison et al. 2010; Neffke et al. 2011), skill relatedness (Neffke and Henning 2013), technological cooperation (Basile et al. 2012) and institutional relatedness (Boschma and Capone 2015; Cortinovis et al. 2017), the exact benefits for firms of whatever relatedness remain unknown. This gap is remarkable because the theories that underlie spatial externalities and agglomeration economies are microeconomic in nature. Heterogeneity at the firm level (Baldwin and Okubo 2006; Cingano and Schivardi 2004), the inter-regional level (Acemoglu and Dell 2010), various firm-context interplay levels (Syverson 2011) and the level of mechanisms of agglomeration economies (Faggio et al. 2017) are introduced in the literature as important for understanding variations in firm-level productivity. A link to—by agglomeration externalities inspired—cognitive relatedness, however defined, then again is missing, although Neffke and Henning (2014) in a framework of agents of change and entrepreneurship explore the role of the firm in diversification processes (not productivity).

In this study, we use data and methods that allow us to disentangle firm-level variation in productivity from that in heterogeneous composed regions, sectors and size types of firms, while identifying relationships between agglomeration and relatedness variables. Previous work on variety and firm productivity has in this regard has argued related variety positively influences productivity through enabling knowledge

spillovers and increasing possibilities of outsourcing locally (Cainelli et al. 2016), while unrelated variety negatively impacts productivity of firms because very different industries do not have complementarities in factor inputs and knowledge spillovers (Aarstad et al. 2016). Building on the firm-level studies (Cainelli and Iachobuchi 2012; Aarstad et al. 2016; Cainelli et al. 2016; Innocenti and Zampi 2019) that have examined the relationship between variety and firm performance, we add to the literature by (a) applying firm-level production functions with capital and labor, (b) estimating firm-level total factor productivity, (c) introducing relatedness as moderating agglomeration externalities mechanism for productivity (controlling for other determinants) and (d) introducing heterogeneity at firm, regional, sector and region–sector (cluster) levels (including business services).

Applying this to Europe (in 2010), we find that regional and sectoral bounded externalities of relatedness foster firm productivity to a large extent (region more so than sector). More specifically, we find that related variety has a positive impact on firm-level productivity, especially for firms in high-tech sectors and regions, for knowledge-intensive business services and for firms of all sizes except micro-firms. Robustness analyses reveal that firm productivity is not affected by both regional and sectoral high-tech characterizations in the same degree, but by regional contexts more than by sectoral ones. Similarly, larger firms systematically have higher relatedness impacts than those that are only sectorally advanced. Unrelated variety is positively associated with productivity of medium-tech firms and micro-firms, and negatively to productivity of low-tech firms, indicating that cognitively unrelated activities can still work out positive for productivity in European firms in specific entrepreneurial circumstances. This heterogeneity has important implications for European and national policies.

The remainder of this paper is organized as follows. Section 2 introduces related and unrelated variety in agglomeration and strategic management studies. Section 3 sets out our data, research methodology and model specification. Section 4 presents modeling results, and Sect. 5 concludes and discusses policy implications.

2 Related literature

In economic geography and regional economics, there is now a long tradition of research that has examined which static and dynamic agglomeration conditions give rise to economic growth. Melo et al. (2009) and De Groot et al. (2015) document accurately the stream of literature that started with the seminal contributions of Glaeser et al. (1992) and Henderson et al. (1995) on localized agglomeration externalities and economic growth. The notion of agglomeration economies refers to external economies of scale arising from the concentration of firms and individuals in space. The effects of agglomeration economies on firm productivity have been researched extensively. Puga (2010) summarizes the evidence of such agglomeration economies based on (1) a clustering of production beyond what can be explained by chance or comparative advantage, (2) on spatial patterns in wages and rents and (3) on systematic variations in productivity with the urban environment. Productivity gains from agglomeration economies have been documented unequivocally in the

literature. Melo et al. (2009) in a meta-analysis of 729 elasticities taken from 34 different studies found that the productivity gains of urban agglomeration economies are positive, but vary greatly in their reported estimates. The authors attribute this significant variability in the magnitude of agglomeration effects to country-specific effects, the industrial coverage, the specification of agglomeration economies and the presence of controls for both unobserved cross-sectional heterogeneity and differences in time-variant labor quality.

Concurrently, there is mounting evidence that the question whether sectoral diversity or specialization is most conducive to regional growth, is actually not very informative from the start as it ignores the fundamental heterogeneity on multiple levels involved. Studies on the effects of agglomeration externalities on regional growth have indeed produced mixed results to such a large extent, that the dichotomy between specialization and diversification cannot fully capture the complexity and heterogeneity of agglomeration externalities. Instead, there is by now also a burgeoning literature stressing that specialization and diversification should be seen as complementary over industry life cycles, contracted to diversification opportunities and capabilities of firms and institutions in economies, with entrepreneurial and cognitive skills functioning as catalysts for emerging and renewed economic diversification (Cortinovis et al. 2017; Van Oort et al. 2012; Faggio et al. 2017). Yet, diversification or (in later stages) specialization in itself does not guarantee economic growth.

Featuring the idea of cognitive proximity, Frenken et al. (2007) distinguish between related and unrelated variety: it is not just the presence of different technological or industrial sectors that stimulates regional development; industries need complementarities in terms of shared competences or untraded capabilities and interdependencies (Storper 1997). Related variety is distinguished from unrelated variety, since information and knowledge will not transfer to all different industries evenly due to varying cognitive distances between each industry pairs (Ellison et al. 2010). It is argued that industries are more related when they are closer to each other in an industry classification system—measured by co-occurrence and hierarchical industry classification membership. Recent studies have shown that especially related variety contributes to productivity and employment growth at the regional level (see for an overview Content and Frenken 2016). Whereas the early studies on related variety mainly examined the overall effect of varieties on regional performance, more recently scholars have turned attention to the issues that (a) relatedness can be measured more directly linked to mechanisms of knowledge transfer (Hartog et al. 2012), for example input–output relations (Fan and Lang 2002; Feser 2003), labor flows (Neffke and Henning 2013, 2013; Diodato and Weterings 2014), innovation based co-operations (Neffke and Henning 2013; Rigby 2012), and (b) that related variety is perhaps not a *conditio sine qua non* either for regional performance, but that the impact of related variety (or agglomeration externalities) on growth is dependent on the characteristics of the sector, size, age and evolution phases of firms (Faggio et al. 2017), institutional settings in nations and regions (Boschma et al. 2015; Cortinovis et al. 2017) and absorptive capacities of firms in regional economic systems (Fritsch and Kublina 2018). In studies that focus on sectoral heterogeneity and diversification, Hartog et al. (2012), Timmermans and

Boschma (2013), Marrocu and Paci (2013), Bishop and Gripaios (2010) and Cortinovis and Van Oort (2015) found that related variety is more important for high-tech industries, where acquisition and internalizing of new knowledge is among the key issues for firm survival and growth. Yet, there is a legion of possibilities of the impact of heterogeneity—ranging from sectors to regions to sector-industry combinations to firms level capabilities and characteristics. This calls for careful quantification and modeling of multilevel processes and characteristics (Knoben et al. 2016).

Increasingly, more empirical studies on agglomeration use disaggregated and firm- and plant-level data, with cities or city industries as a reference unit that is built-up by heterogeneous attributes (Wixe 2015). Studies using aggregated data provide only limited insights and weak support for the effects of agglomeration economies on firm performance (Van Oort et al. 2012). Regional-level relationships are not necessarily reproduced at the firm level because information on the variance between firms is lost when aggregated regional-level data are used. Hence, even if regions endowed with a greater number of agglomeration economies grow faster, this conclusion cannot be generalized to firms.

In addition, agglomeration effects found in area-based studies may be purely compositional (Neumark and Simpson 2015). Combes et al. (2012) and Behrens et al. (2014) show how spatial sorting influences productivity measures. Similarly, Baldwin and Okubu (2006) show that the agglomeration of productive firms may simply be the result of a spatial selection process in which more productive firms are drawn to dense economic areas. For this reason, it remains unclear whether geographical differences are an artifact of location characteristics (e.g., agglomeration economies) or are simply caused by differences in business and economic composition (heterogeneity). Urban economic and spatial econometrics can identify urban and regional-level phenomena while controlling for sorting effects of firms and workers. Alternatively, following social sciences, hierarchical random effects or multilevel modeling, which allows the micro-level and macro-level (agglomerated contexts) to be modeled simultaneously, is becoming an increasingly common practice in strategic management and organization studies (Corrado and Fingleton 2012; Knoben et al. 2016). In this paper, addressing the micro–macro-level heterogeneity and interrelationships—basically, questioning which types of firm's profit from which types of agglomeration and externality contexts—our analysis is served by multilevel modeling.

The multilevel character of firm-level productivity has been set out by Cingano and Schivardi (2004) and Syverson (2011), who make a plea for analyzing firm-level total factor productivity in relation to simultaneously important contexts. Especially Syverson (2011) surveys and evaluates recent empirical work addressing the question of why businesses differ in their measured productivity levels. The causes turn out to be manifold and differ depending on the particular setting. He includes elements sourced in production practices—and therefore over which producers have some direct control, at least in theory—as well as from producers' external operating environments (contexts). Building on papers by Maksimovic and Phillips (2002) and Schoar (2002) that both investigate the productivity of plants within conglomerate firms (in their setting, those that operate in multiple two- or three-digit SIC

industries), and that leverage US manufacturer micro-data to convincingly argue that a diversification leads firms to over perform in productivity because of learning effects inducing selection, the optimal allocation of resources within a business and capable management. The relatedness literature and its absorptive capacity interpretation uses similar arguments, meaning that relatedness (variety) does not directly foster regional development, but only indirectly through its effect on firm performance.

Following the discussed literature, our models, that to the best of our knowledge are one of the first to capture firm-level productivity in various contexts simultaneously, should incorporate heterogeneity at firm level (degree of high-tech and innovativeness), at regional level (core-periphery distinction of technologically advanced regional production complexes), sectoral level (including high-tech manufacturing and knowledge-intensive business services) and sector by region level (clusters). Previous literature that has linked variety to productivity has argued that

Our study comes closest to the work of Aarstad et al. (2016), who conducted a multilevel study on the relationship between related and unrelated variety and enterprise productivity. Using Norwegian data, the authors concluded that unrelated variety is negatively associated with enterprise productivity, when related variety increases the innovation of the enterprises. Cainelli et al. (2016) examined the relation between related variety and total factor productivity of over 12,000 Italian manufacturing companies. They found that there is a positive effect of related variety on firms' TFP, although this effect was less strong than the effect of the specialization-related localization externalities. In related work, Cainelli and Iacobucci (2012) focused on Italian business groups to analyze the effect of the variety of industries in the local area on the vertical integration of firms. They found that especially a higher level of vertically related variety (vis-à-vis unrelated variety) is associated with a lower need of firms to integrate activities along the production chain because they can obtain intermediate goods and services locally. In addition, the authors showed that in areas with many small firms variety significantly influences the degree of vertical integration, while areas characterized by many small firms related variety significantly influence activities along the production chain. Conversely, Innocenti and Zampi (2019) did not find a significant association between related and unrelated variety and the growth of innovative start-ups at an early stage of their life.

3 Data and methodology

Using a Cobb–Douglas production function framework and multilevel (mixed) models, we control for firm inputs such as employment and capital, but also control for other regional-level indicators like localization, education level, population density and economic openness. In this, we are also explicitly interested to what extent firms in different industries are differently affected by related and unrelated variety. To analyze the relationship between related variety, unrelated variety and firm productivity, we use data from the ORBIS–Amadeus database, a commercial database provided by Bureau van Dijk (Moody's Analytics). The ORBIS–Amadeus database provides financial and economic information on firms in almost all European

countries, including balance sheets and income statement items, and provides a wide range of performance indices. ORBIS–Amadeus also includes information on a firm's location, ownership and age and assigns NACE codes to companies—the European standard of industry classification after the French *Nomenclature statistique des Activités économiques dans la Communauté Européenne*—which can be used to classify firms according to sector. The effort to standardize and harmonize financial information included in the ORBIS–Amadeus database and make it comparable across countries has resulted in its wide-spread use in research (Kalemli-Özcan 2015). We employ a cross-sectional database of firms for 13 European Union Countries for year 2010. Only firms with known information of value added for the year included in the model. Our sample consists of 200,463 firms, located across 13 European countries.

We distinguish between several classifications in our models. First, we categorized firms into 184 NUTS-2 regions, second in 9 sectors and third in 1525 sectors by regions (clusters). The countries included in the sample are Belgium, Czech Republic, Germany, Spain, Finland, France, Italy, Portugal, Sweden, UK, Slovakia, Poland and Hungary. The 9 sectors include (1) Manufacturing 1 (food and textiles), (2) Manufacturing 2 (chemicals and plastic), (3) Manufacturing 3 (metals), (4) Manufacturing 4 (transportation and storage), (5) information and communication, (6) real estate services, (7) professional, scientific and technical activities, (8) administrative and support service activities and 9) other service activities. Following Wintjes and Hollanders (2010) and Cortinovis and Van Oort (2015) regions were ranked according to their capacities in terms of knowledge accessibility, knowledge absorption and knowledge diffusion (see "Appendix 1"). We categorized our regions into three technological regimes (high technological regime, medium technological regime and low technological regime). Furthermore, firms were categorized according to their technological and knowledge intensity following Faggio et al. 2017 into four categories, (1) high–medium-technology manufacturing, (2) medium–low technology manufacturing, (3) knowledge-intensive business services and (4) less knowledge-intensive business services. Although service industries are frequently studied in agglomeration analyses (Henderson et al. 1995; Kolko 2010; Jacobs et al. 2013), they are mostly absent in relatedness studies that resort data from input–output methodologies or manufacturing surveys. We further categorized firms according to their size into four categories, micro-sized firms, small-sized firms, medium-sized firms, large-sized firms, according to the European Union categorization.[1]

Our study employs three (composed) firm-level variables of the ORBIS–Amadeus database. The dependent variable in this study is firm labor productivity and is measured as the logarithm of value added on employees for each firm *i* for year 2010. Employment is measured as the logarithm of the number of employees per firm. Capital is measured as the logarithm of the sum of tangible fixed assets and depreciation and is measured in thousands of euro (firm specific).

---

[1] Micro-sized firms (1–9 employees), small-sized firms (10–50 employees), medium-sized firms (51–250 employees), large-sized firms (251 and more). http://ec.europa.eu/enterprise/policies/sme/facts-figuresanalysis/sme-definition/.

Agglomeration variables are measured at the regional level, and besides related and unrelated variety, two other variables are introduced. Localization economies or sectoral specialization is measured as the concentration (location quotient) of employees per region (measured at the region-by-sector level). Localization economies concern sector-specific externalities that are thought important for productivity, especially in later stages of firm and industry life cycles (Kemeny and Storper 2015). Next, density is measured as population per km^2. This dimension of agglomeration is not directly related to localization economies (specialization) and diversity economies, but to pure urban size effects (Puga 2002; Ciccone 2002; Ciccone and Hall 1996). In general, the literature suggests that higher density enables better interaction, enhancing growth (Rice et al. 2006), although recent research that controls for composition effects also moderates the suggested impact (Henderson 2003).

Several variables are introduced on the regional level to control for further determinants of regional and cluster heterogeneity (Kim 1997). The degree of openness of European regions (a measure related to firm competition and regional competitiveness) is calculated as the total value of imports and exports in a region divided by the regions GDP. This volume of trade indicator is based on a make and use tables (IO-Table) for 2000 on NUTS-2 level concerning 14 sectors and 59 product categories, including services. This database is developed by the Netherlands Environmental Assessment Agency (PBL), see Thissen et al. (2017). The volume of trade goes up with the size of the region at a declining rate. It is strongly dependent on global economic development with competition on global markets, driving up productivity and attracting new investments and collaborations. High potential may also spill over to nearby regions or in the regional network of specialized and subcontracting industries and regions. The average educational level of regions is measured by the percentage of tertiary and higher educated in the total population. The relationship with (employment and productivity) growth thought to be is positive, as more skilled people can be more productive, and agglomeration may attract more of these people (Moretti 2004; Rauch 1993, for a more critical interpretation see Shapiro 2006).

To measure related and unrelated variety, we use an entropy measure proposed by Frenken et al. (2007) at the NUTS-2 level. When measuring regional variety to study the effects on firm productivity, decomposition is useful as it is expected that variety at a high level of sector aggregation reflects the possibility to switch between input substitutes (unrelated variety), while one expects variety at a low level of sector aggregation to be an indication of possible knowledge spillovers because of cognitive similarity (related variety). We also use the ORBIS–Amadeus micro-data as source for the calculation of related and unrelated variety at the NUTS 1 Level. Since small firms are underrepresented in this database, firm-level data are weighted by turnover values. In this fashion, we capture the large and sectorial heterogeneous regional economies best. We compute entropy using employment data, which are available for the four-digit level from the ORBIS–Amadeus database. Unrelated variety per region is indicated by the entropy of the two-digit distribution; related variety is indicated by the weighted sum of the entropy at the four-digit level within each two-digit class.

Marginal variety can be computed at all four-digit SIC levels in the dataset, meaning the increase in variety when moving from one digit level to the next. Formally, let P_i be the four-digit SIC levels share of employment in Section G over total employment in region, by summing the four-digit share one can measure variety

$$V = \sum_{g=1}^{G} P_i \log_2 \left(\frac{1}{P_i} \right).$$

This measure is variety in a general form. The higher its value is, the higher the industrial diversification of a region is. This measure can be split into related and unrelated variety. Firstly by summing the four-digit shares P_i, one can derive the two digits shares, P_g

$$P_g = \sum_{i \in S_g} P_i.$$

Then, unrelated variety is measured as the two-digit level entropy and is given by:

$$UV = \sum_{g=1}^{G} P_g \log_2 \left(\frac{1}{P_g} \right).$$

To measure related variety we estimate the marginal increase when moving from the two-digit to the four-digit level. Formally related variety is the weighted sum of entropy within each two-digit sector and is given by:

$$RV = \sum_{g=1}^{G} P_g H_g$$

where[2]

$$H_g = \sum_{i \in S_g} \frac{P_i}{P_g} \log_2 \left(\frac{1}{P_i/P_j} \right)$$

The method of hierarchical or multilevel modeling allows the micro-level and macro-level to be modeled simultaneously. There are two distinct advantages to multilevel models (Van Oort et al. 2012). First, multilevel models offer a natural way to assess contextuality, or the extent to which a link exists between the macro-level and the micro-level. Applying multilevel analysis to empirical work on agglomeration begins from the simple observation that firms sharing the same external environment are more similar in their performance than firms that do not share the same external environment because of shared agglomeration or other externalities. Hence, we can assess the extent to which variance in firm-level

---

[2] To avoid multicollinearity in our models, we tested for high correlations among all explanatory variables (Tables 1 and 2), and we analyzed variance inflation factors for each variable added in the model. No signs of multicollinearity were detected in all of our models.

productivity can be attributed to between-firm variance, between-area variance, between-sector variance or between cluster (sector-region). With multilevel analysis, we are able to assign variability to the appropriate context. Second, multilevel analysis allows us to incorporate unobserved heterogeneity into the model by including random intercepts and allowing relationships to vary across contexts through the inclusion of random coefficients. Whereas "standard" regression models are designed to model the mean, multilevel analyses focus on modeling variances explicitly. This kind of complexity can be captured in a multilevel framework through the inclusion of random coefficients.

A Cobb–Douglas production function is expressed as follows:

$$Y_{ijk} = A(T_{ijk})^{\beta_n} K_{ijk}^{\beta_1} L_{ijk}^{\beta_2} e^{e_{ijk}} \tag{1}$$

where y_{ijk} is labor productivity (value added per worker) of the ith firm nested within jth group of sector by region, which is nested within the kth region, $A(T_{ijk})$ is TFP, K_{ijk} the capital stock and L_{ijk} the labor force. TFP of plant i depends on its regional and sectorial environment (see Fig. 1), and it is expressed in terms of related and unrelated variety (sectorial), localization, and urbanization economies, openness of the region and level of education of the inhabitants (regional). Taking logs of Eq. (1) and applying a three-level model which involves regions at the higher-level, sectors by regions at the meso-level and firms within regions at the lower level, we obtain Eq. (2):
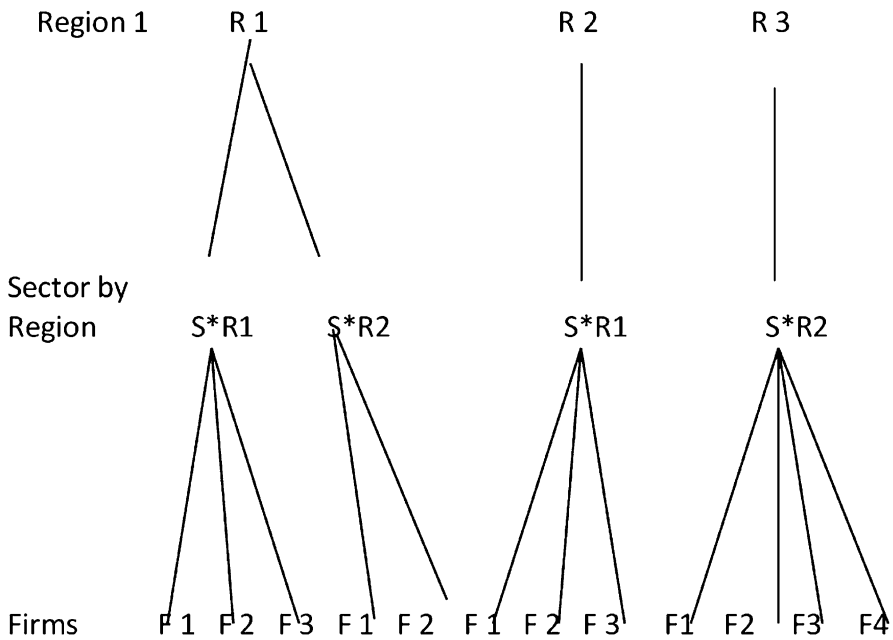

Fig. 1 Multilevel structure of analyses

$$y_{ijk} = \beta_0 + \beta_1 \text{CAP}_{ijk} + \beta_2 \text{EMPL}_{ijk} + \beta_3 \text{REL}_{ijk} + \beta_4 \text{UNREL}_{ijk} + \beta_5 \text{LOC}_{ijk}$$
$$+ \beta_6 \text{URB}_{ijk} + \beta_7 \text{OPEN}_{ijk} + \beta_8 \text{EDUC}_{ijk} + u_{00k} + u_{0jk} + e_{0ijk} \qquad (2)$$

where u_{0k} is the partitioning of the total variance in variance between regions, u_{0jk} is the variance between sectors by regions but within regions, and e_{0ijk} is the variance within sectors by regions. Figure 1 presents the multilevel structure of the analyses.

In Eq. 2, we assume that the firm-level predictor variables are uncorrelated with the sector and regional-level error terms and that the sector-level predictor variables are uncorrelated with the regional-level error terms. However, both theoretical and empirically, such an assumption is difficult to meet. Not correcting for this would lead to inconsistent parameter estimates. However following Snijders and Berkhof (2008), we remove the correlation between the lower-level predictor variables and higher-level error terms, by including sector and regional means of the firm-level predictor variables in the regression model, a procedure known as the Mundlak (1978) correction.

Please note that although the Mundlak correction addresses part of the endogeneity problem, the multilevel modeling framework does not control for endogeneity arising from reverse causality between firms' productivity and variety. Specifically, firms' location choices resulting in lower or higher levels of related and unrelated variety in some places could be induced by the migration of firms between places of low productivity to places of high productivity (Melo et al. 2009). One solution to this problem would be to instrument related and unrelated variety, but unfortunately, finding credible instruments is hard. Therefore, the results should be interpreted as conditional associations, rather than causal relationships.

4 Results

In Table 3, the results of the main model are reported. Although there is no positive association between related variety in the model without control variables, related variety is positively and significantly related to firm's productivity in the last, preferred specification including sector and country fixed effects (dummies). Furthermore in that specification, employment and capital (firm level), and education

	Mean	SD	Minimum	Maximum
Added value	6.33	1.84	0	17
Employment	2.62	1.64	0	13
Capital	4.22	2.14	0	17
Rvar	0.80	0.23	0	1
Uvar	2.78	0.28	1	3
Localization	−0.30	1.51	−9	6
Density	5.09	1.06	2	9
Openness	−0.49	0.22	−1	1
Education	2.94	0.54	2	4
Observations	200,463			

Table 1 Descriptive statistics of explaining variables

Table 2 Correlation matrix

Variables	Related variety	Unrelated variety	Openness	Education	Density	Employment	Capital	Localization
Related variety	1.0000							
Unrelated variety	0.5489	1.0000						
Openness	0.0107	−0.2647	1.0000					
Education	0.2417	0.2348	−0.0943	1.0000				
Density	0.5276	0.4972	−0.2852	0.0517	1.0000			
Employment	−0.0031	−0.0331	−0.0412	−0.0440	0.1973	1.0000		
Capital	0.0231	−0.0173	−0.0792	−0.0800	0.1944	0.6238	1.0000	
Localization	0.4902	0.3670	−0.1265	0.0017	0.7263	0.1757	0.1390	1.0000

Table 3 Regression analyses of firm productivity

	Null model	+RV	+UV	RV and UV	+Firm characteristics	+Regional controls	+Dummies
Fixed part							
Intercept	7.577***	7.617***	7.158***	7.186***	2.429***	0.962***	2.568***
	(0.109)	(0.305)	(0.706)	(0.706)	(0.283)	(0.259)	(0.224)
Employment					0.928***	0.932***	0.932***
					(0.015)	(0.015)	(0.015)
Capital					0.120***	0.117***	0.117***
					(0.006)	(0.006)	(0.006)
Related variety		−0.069		−0.140	0.015	−0.009	0.133**
		(0.482)		(0.512)	(0.162)	(0.157)	(0.049)
Unrelated variety			0.159	0.180	0.389***	0.103	0.038
			(0.270)	(0.287)	(0.105)	(0.086)	(0.031)
Localization						−0.008	−0.007
						(0.005)	(0.005)
Density						0.044	0.023**
						(0.035)	(0.012)
Openness						−0.460***	−0.028
						(0.092)	(0.071)
Education						0.575***	0.167**
						(0.077)	(0.052)
Sector dummies	No	No	No	No	No	No	Yes
Country dummies	No	No	No	No	No	No	Yes
Random part							
e_{0jk}	1.459	1.459	1.458	1.457	0.443	0.336	0.054
	(0.052)	(0.052)	(0.053)	(0.054)	(0.024)	(0.019)	(0.010)
u_{0jk}	0.459	0.459	0.459	0.459	0.238	0.155	0.154
	(0.018)	(0.018)	(0.018)	(0.018)	(0.013)	(0.010)	(0.010)
u_{00k}	1.524	1.524	1.524	1.524	0.647	0.647	0.647
	(0.035)	(0.035)	(0.035)	(0.035)	(0.014)	(0.014)	(0.014)

Table 3 (continued)

	Null model	+RV	+UV	RV and UV	+Firm characteristics	+Regional controls	+Dummies
Statistics							
Sample size	196,330	196,330	196,330	196,330	152,378	147,162	147,162
Number of regions	184	184	184	184	184	184	184
Number of regions * sectors	1523	1523	1523	1523	1510	1443	1443

Robust standard errors in parentheses; **$p<0.05$; ***$p<0.01$

Table 4 Regression analyses of firm productivity and related variety across different spatial regimes

	High technological regime	Medium technological regime	Low technological regime
Fixed part			
Intercept	3.692***	3.414***	1.649***
	(0.789)	(0.231)	(0.283)
Employment	0.880***	0.916***	1.003***
	(0.038)	(0.020)	(0.087)
Capital	0.135***	0.111***	0.119***
	(0.022)	(0.006)	(0.010)
Related variety	0.256**	0.106	−0.108*
	(0.107)	(0.066)	(0.055)
Unrelated variety	−0.054	0.020	−0.043
	(0.077)	(0.037)	(0.039)
Localization	−0.003	0.002	−0.025**
	(0.008)	(0.006)	(0.011)
Density	−0.001	0.017	0.108***
	(0.024)	(0.016)	(0.019)
Openness	−0.244	−0.119	0.049
	(0.220)	(0.101)	(0.119)
Education	−0.074	0.052	0.254***
	(0.157)	(0.060)	(0.063)
Random part			
e_{0ijk}	0.000	0.041	0.000
	(0.000)	(0.010)	(0.000)
u_{0jk}	0.179	0.121	0.149
	(0.128)	(0.014)	(0.024)
u_{00k}	0.685	0.602	0.691
	(0.029)	(0.024)	(0.015)
Statistics			
Sample size	22,184	78,355	46,611
Number of regions	42	86	55
Number of regions * sectors	342	694	401

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

and population density (regional level) are positive and significantly related to firm productivity. Localization and unrelated variety do not have a significant impact on labor productivity in the main models presented in Table 3.

Table 4 tests for the effect of related variety on firm labor productivity in a spatial regime setting, where the regimes reflect different degrees of regional high-tech characteristics. Related variety is only positive significant related to productivity of firms in the high-technology regime, insignificant for medium-technology regime regions and negatively significant for low technology regimes. All other significant results are as expected (with localization impacting negatively on firm productivity, as often encountered in the literature).

Table 5 focuses on the relationship between related variety and firm productivity in different sectors, identified according to knowledge intensity (compare Henderson

Table 5 Regression analyses of firm productivity and related variety across different sectors

	High–medium-technology manufacturing	Medium–low technology manufacturing	Knowledge-intensive services	Less knowledge-intensive services
Fixed part				
Intercept	3.145***	2.692***	2.370***	2.714***
	(0.249)	(0.215)	(0.316)	(0.353)
Employment	0.989***	0.988***	0.912***	0.857***
	(0.012)	(0.013)	(0.022)	(0.016)
Capital	0.070***	0.084***	0.126***	0.171***
	(0.004)	(0.005)	(0.011)	(0.009)
Related variety	0.130**	0.046	0.119*	0.040
	(0.058)	(0.037)	(0.067)	(0.059)
Unrelated variety	0.107**	0.046*	0.023	0.024
	(0.038)	(0.028)	(0.047)	(0.041)
Localization	0.020**	0.007	0.008	0.019
	(0.006)	(0.007)	(0.007)	(0.012)
Density	− 0.013	0.009	0.012	− 0.022
	(0.014)	(0.010)	(0.161)	(0.016)
Openness	0.052	0.124**	− 0.067	0.059
	(0.080)	(0.063)	(0.097)	(0.083)
Education	0.011	0.097**	0.159**	0.274***
	(0.050)	(0.044)	(0.077)	(0.067)
Random part				
e_{0ijk}	0.000	0.048	0.081	0.000
	(0.000)	(0.007)	(0.012)	(0.000)
u_{0jk}	0.107	0.056	0.073	0.149
	(0.079)	(0.005)	(0.013)	(0.014)
u_{00k}	0.542	0.568	0.733	0.742
	(0.011)	(0.009)	(0.016)	(0.023)
Statistics				
Sample size	19,895	66,638	30,758	29,871
Number of regions	179	182	174	178
Number of regions * sectors	324	522	553	576

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

et al. 1995; Faggio et al. 2017; Jacobs et al. 2013). Related variety is positive and significant related to firm productivity in sectors with higher level of technological manufacturing and knowledge-intensive services. Our findings suggest that related variety as an agglomeration externality impacts most in high-tech and high-service firms, confirming earlier findings by inter alia Bishop and Gripaios (2010); Hartog et al. (2012), Timmermans and Boschma (2013) and Cortinovis and Van Oort (2015). In the sectoral models, unrelated variety has similar significant impacts on firm productivity than related variety (compare Castaldi et al. 2015), indicating that also diversification opportunities in cognitively unrelated industries (breakthrough technologies) correlate with firm productivity in high-tech and medium-tech manufacturing sectors.

Table 6 Regression analyses of firm productivity and related variety across different firm sizes

	Micro-firms	Small firms	Medium firms	Large firms
Fixed part				
Intercept	3.923***	3.319***	3.002***	2.735***
	(0.485)	(0.226)	(0.209)	(0.218)
Employment	0.902***	0.967***	0.917***	0.874***
	(0.020)	(0.015)	(0.010)	(0.009)
Capital	0.147***	0.098***	0.098***	0.117***
	(0.012)	(0.006)	(0.004)	(0.005)
Related variety	0.108	0.152**	0.161**	0.132**
	(0.098)	(0.050)	(0.061)	(0.064)
Unrelated variety	−0.099	−0.010	0.027	0.053
	(0.101)	(0.038)	(0.034)	(0.043)
Localization	0.020**	−0.019*	−0.031***	−0.051***
	(0.006)	(0.007)	(0.006)	(0.010)
Density	0.035	0.054***	0.069***	0.078***
	(0.028)	(0.014)	(0.014)	(0.018)
Openness	0.043	−0.012	−0.069	−0.018
	(0.159)	(0.078)	(0.071)	(0.087)
Education	0.126	0.122*	0.141**	0.192**
	(0.131)	(0.065)	(0.056)	(0.060)
Random part				
e_{0ijk}	0.136	0.020	0.035	0.000
	(0.080)	(0.052)	(0.023)	(0.000)
u_{0jk}	0.229	0.217	0.189	0.207
	(0.027)	(0.023)	(0.012)	(0.254)
u_{00k}	0.750	0.508	0.531	0.572
	(0.015)	(0.010)	(0.013)	(0.027)
Statistics				
Sample size	64,697	53,641	21,351	7473
Number of regions	159	175	184	180
Number of regions * sectors	905	1158	1245	1036

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

Model 6 presents models where the sample is split according to firm sizes. There is a positive relationship between related variety and firms of all sizes except the micro-firms. The results indicate that knowledge spillovers are more likely to be transmitted between firms that are not too small, e.g., have larger absorptive capacities (Fritsch and Kublina 2018). Very small firms cannot adapt easily to new techniques and do not allow diverse forms of knowledge to be absorbed easily. The positive and significant sign of localization for micro-firms indicates that specialization externalities are actually those contributing to firm growth. For the other firm size groups, localization is negative significant related to firm productivity (Table 6).

Tables 8, 9 and 10 ("Appendix 2") show results of robustness checks, where further analyses were conducted on the interaction of spatial and sector heterogeneity. From these additional analyses, it can be concluded that the significance and the sign of related variety changes when combining the level of technological regimes and

Table 7 Technological regimes and types of regions (Wintjes and Hollanders 2010)

Technological regimes	Types of regions	Features
High-tech. regime	Metropolitan knowledge-intensive services regions Public knowledge centers High-tech regions	High absorption capacity High accessibility High diffusion, accessibility and absorption capacity
Medium-tech. regime	Knowledge-absorbing regions Skilled technology regions	Average performance in diffusion, accessibility and absorption capacity
Low-tech. regime	Traditional Southern regions Skilled industrial Eastern EU regions	Below average in diffusion, accessibility and absorption capacity Below average in diffusion and absorption capacity

level of sectors' knowledge intensity. The main sensitivity found is that related variety is foremost positively and significantly related to productivity of firms when heterogeneity in the technological level of regions is addressed, while it is less related to the knowledge intensity of sectors. The effect of sectors diminishes when a firm operates in a high-technology regional regime, which implies that the positive and significant coefficient of related variety among high-tech sectors is a result of interdependencies with the regimes; technological level, and not the result of separating it from low-and-medium-sector knowledge intensity variety. This suggests that that related variety has an impact on firm productivity mainly through regionally embedded agglomeration economies.

Similarly, additional analyses are presented on the interaction of sectors (high-tech manufacturing, medium- and low-tech manufacturing, knowledge-intensive services, less knowledge-intensive services) and firm size classes in Tables 11, 12, 13 and 14 in "Appendix 2". These tables show that the firms are, the larger the impact of related variety and localization economies on firm-level productivity. The knowledge-intensive services show identical patterns. Contrary to this, medium- and low-tech firms and less knowledge-intensive firms show hardly any productivity impact from related variety.

5 Conclusions and discussion

In this paper, we use a model to take multilevel heterogeneity into account when analyzing productivity in a firm context. Both the relatedness/externalities literature and the firm performance literature hint at their potential interplay, yet this is not substantiated yet in detailed analyses at the European level. From our analyses, it became clear that differences in regional, sectoral and firm size compositions of economic activities condition the degree to which firms are more productive in Europe. These insights are important for analyzing the growth opportunities and competitive advantage of regions, and for drafting place-based policies to improve these positions. Given the hypothesized importance of European national and regional integration, the testing of the effectiveness of sectoral policies, and the identification of entrepreneurial and firm-level opportunities and niches, the absence of insights into these policies so far is remarkable.

We used data (for 2010) and methods that allowed us to disentangle firm-level variation in productivity from that in heterogeneous composed regions, sectors, region–sector combinations (clusters) and size types of firms, while identifying relationships between agglomeration and relatedness variables. Applying this to Europe, we found that regional and sectoral bounded externalities of relatedness fostered firm productivity to a large extent, suggesting that the regional level of absorptive capacity of firms plays an important role. We more specifically found that related variety impacts positively on firm-level productivity, especially for firms in high-tech sectors and regions, for knowledge-intensive business services and for firms of all sizes except micro-firms.

Robustness analyses revealed that firm productivity was not affected by sectoral and regional high-tech characterizations to the same degree, but by regionally mediated

spillovers more than by sectoral ones. Similarly, larger firms systematically had higher relatedness impacts than those that were only sectorally advanced. These findings back the general notion of place-based (regional) development strategies in the European Union being an essential ingredient next to people-based or firm-based policies (Barca et al. 2012), as localized policies are best used to address specific bottlenecks and diversification opportunities given path-dependent development trajectories locally.

This research also has limitations that may be addressed in future research. The analysis was based on a cross-sectional dataset of firms in European regions and sectors for 2010. Dynamic data would allow better identification, and addition insights into growth opportunities. As relatedness becomes more and advanced in measurement, staying increasingly close to mechanisms of knowledge transfer (labor mobility, innovative cooperation, detailed input–output linkages), future analysis could couple to these measurement improvements to test the robustness of their outcomes. Another problem that has remained unaddressed in the paper is that the effects of related and unrelated variety could be biased as a result of possible endogeneity, arising from reverse causality between firms' productivity and variety: firms' location choices inducing related and unrelated variety could be induced by high levels of productivity present in some places (Melo et al. 2009). In other words, differences in productivity between places might result in migration of firms. Finally, better grained data on firm level (on absorptive capacity for instance) would give us more grip on the impacts of heterogeneity.

Appendix 1: Technological regional regimes

The variables used by Wintjes and Hollanders (2010) are grouped into five sets:

1. Employment: including shares in specific NACE classes for high-tech manufacturing, medium–high-tech manufacturing, high-tech services and market services, as well as employment shares in industry sectors (NACE from C to E), service sectors (from G to K) and Government sectors (from L to P);
2. Human resources: including the share of employment in science and technology occupations and the share of the workforce with secondary and tertiary education;
3. Activity rates: activity rates for females, activity rates for individuals with tertiary education, and the share of long-term unemployed over total employment;
4. Technology: R&D as percentage of GDP, share of university R&D over total R&D, share of government R&D over total R&D, and EPO applications per million population;
5. Economy: capital formation as a percentage of GDP, labor productivity in industry sectors and labor productivity in Service sectors.

These indicators are then used in a factor analysis and reduced to eight factors related to knowledge economies. Wintjes and Hollanders studied these factors through a cluster analysis, allowing them to identify the seven typologies of knowledge economies into which they classify European regions (see Table 7).

Appendix 2: Results from regime/sector models

See Tables 8, 9, 10, 11, 12, 13 and 14.

Table 8 High technological regime/sectors

	High–medium-technology manufacturing	Medium–low technology manufacturing	Knowledge-intensive services	Less knowledge-intensive services
Fixed part				
Intercept	4.315***	3.997***	2.383***	4.613***
	(0.655)	(0.385)	(0.870)	(1.221)
Employment	0.972***	0.984***	0.874***	0.779***
	(0.026)	(0.034)	(0.047)	(0.026)
Capital	0.065***	0.066***	0.137***	0.200***
	(0.014)	(0.012)	(0.027)	(0.018)
Related variety	0.239*	0.144**	0.313**	0.295**
	(0.130)	(0.068)	(0.152)	(0.145)
Unrelated variety	0.064	0.015	0.012	−0.080
	(0.080)	(0.040)	(0.087)	(0.095)
Localization	−0.007	0.008	−0.011	−0.002
	(0.014)	(0.010)	(0.012)	(0.026)
Density	0.016	0.016	0.000	−0.001
	(0.031)	(0.015)	(0.023)	(0.033)
Openness	−0.131	−0.088	0.017	−0.303
	(0.212)	(0.151)	(0.282)	(0.306)
Education	−0.354**	−0.279***	0.090	−0.158
	(0.123)	(0.063)	(0.145)	(0.227)
Random part				
e_{0ijk}	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
u_{0jk}	0.071	0.034	0.050	0.176
	(0.025)	(0.251)	(0.028)	(0.067)
u_{00k}	0.526	0.510	0.826	0.693
	(0.018)	(0.025)	(0.021)	(0.049)
Statistics				
Sample size	2918	7057	7305	4904
Number of regions	42	42	42	42
Number of regions * sectors	78	122	136	138

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

Table 9 Medium technological regime/sectors

	High–medium-technology manufacturing	Medium–low technology manufacturing	Knowledge-intensive services	Less knowledge-intensive services
Fixed part				
Intercept	3.380***	3.782***	2.991***	2.739***
	(0.276)	(0.259)	(0.553)	(0.404)
Employment	0.979***	0.961***	0.910***	0.829***
	(0.016)	(0.022)	(0.023)	(0.024)
Capital	0.070***	0.081***	0.119***	0.160***
	(0.005)	(0.006)	(0.010)	(0.008)
Related variety	0.132	0.010	0.144	0.075
	(0.081)	(0.053)	(0.099)	(0.087)
Unrelated variety	0.134**	0.036	0.017	−0.008
	(0.053)	(0.034)	(0.072)	(0.055)
Localization	0.026***	0.017*	0.005	0.060***
	(0.007)	(0.008)	(0.009)	(0.010)
Density	−0.013	0.004	0.020	−0.063**
	(0.019)	(0.014)	(0.032)	(0.033)
Openness	0.022	0.097	−0.265*	−0.016
	(0.110)	(0.099)	(0.139)	(0.136)
Education	−0.058	0.008	−0.038	0.342***
	(0.085)	(0.067)	(0.130)	(0.096)
Random part				
e_{0ijk}	0.000	0.038	0.079	0.015
	(0.000)	(0.008)	(0.022)	(0.046)
u_{0jk}	0.111	0.048	0.084	0.100
	(0.034)	(0.007)	(0.060)	(0.013)
u_{00k}	0.517	0.533	0.683	0.700
	(0.014)	(0.014)	(0.024)	(0.044)
Statistics				
Sample size	13,166	35,393	15,329	14,467
Number of regions	83	86	85	86
Number of regions * sectors	153	246	277	282

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

Table 10 Low technological regime/sectors

	High–medium-technology manufacturing	Medium–low technology manufacturing	Knowledge-intensive services	Less knowledge-intensive services
Fixed part				
Intercept	2.657***	1.768***	2.045***	0.844
	(0.285)	(0.287)	(0.397)	(0.553)
Employment	1.045***	1.033***	0.991***	0.955***
	(0.017)	(0.012)	(0.012)	(0.011)
Capital	0.072***	0.094***	0.131***	0.173***
	(0.007)	(0.008)	(0.017)	(0.019)
Related variety	−0.090	−0.150	−0.225***	−0.119
	(0.099)	(0.040)	(0.056)	(0.091)
Unrelated variety	−0.117**	−0.040	−0.044	0.032
	(0.053)	(0.030)	(0.069)	(0.076)
Localization	0.011	−0.010	0.036**	−0.028
	(0.015)	(0.017)	(0.013)	(0.018)
Density	0.027	0.055***	0.059**	0.084**
	(0.022)	(0.014)	(0.027)	(0.029)
Openness	−0.202	0.019	0.137	0.399**
	(0.167)	(0.100)	(0.194)	(0.195)
Education	0.036	0.203***	0.251**	0.377**
	(0.087)	(0.056)	(0.080)	(0.110)
Random part				
e_{0ijk}	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
u_{0jk}	0.064	0.067	0.054	0.144
	(0.023)	(0.011)	(0.198)	(0.039)
u_{00k}	0.628	0.625	0.725	0.802
	(0.023)	(0.012)	(0.021)	(0.028)
Statistics				
Sample size	3811	24,185	8120	10,495
Number of regions	54	53	46	49
Number of regions * sectors	93	152	137	154

Robust standard errors in parentheses; **p<0.05; ***p<0.01

Table 11 High–medium-technology manufacturing sector for the 4 firm size groups

High–medium-technology manufacturing	High–medium-technology manufacturing micro-firms	High–medium-technology manufacturing small firms	High–medium-technology manufacturing medium firms	High–medium-technology manufacturing large firms
Fixed part				
Intercept	3.767***	3.493***	4.020***	3.555***
	(0.433)	(0.230)	(0.302)	(0.383)
Employment	0.938***	1.016***	1.023***	0.945***
	(0.035)	(0.015)	(0.021)	(0.013)
Capital	0.094***	0.073***	0.053***	0.076***
	(0.009)	(0.005)	(0.005)	(0.008)
Related variety	−0.073	0.106*	0.185*	0.254**
	(0.107)	(0.061)	(0.096)	(0.127)
Unrelated variety	0.175*	0.084*	0.099*	0.025
	(0.102)	(0.048)	(0.053)	(0.080)
Localization	0.037**	0.015*	0.017	−0.019
	(0.014)	(0.007)	(0.011)	(0.015)
Density	−0.016	0.013	0.000	0.049
	(0.027)	(0.018)	(0.021)	(0.032)
Openness	0.382**	0.057	−0.108	−0.022
	(0.160)	(0.075)	(0.131)	(0.178)
Education	0.133	0.033	−0.110	0.032
	(0.116)	(0.069)	(0.077)	(0.089)
Random part				
e_{0ijk}	0.027	0.000	0.000	0.000
	(0.071)	(0.000)	(0.000)	(0.000)
u_{0jk}	0.114	0.107	0.170	0.177
	(0.042)	(0.013)	(0.030)	(0.016)
u_{00k}	0.634	0.469	0.513	0.497
	(0.015)	(0.008)	(0.023)	(0.022)

Table 11 (continued)

High–medium-technology manufacturing	High–medium-technology manufacturing micro-firms	High–medium-technology manufacturing small firms	High–medium-technology manufacturing medium firms	High–medium-technology manufacturing large firms
Statistics				
Sample size	5499	8082	4547	1767
Number of regions	93	148	172	160
Number of region * sectors	154	245	294	248

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

Table 12 Medium–low technology manufacturing sector for the 4 firm size groups

Medium–low technology manufacturing	Medium–low technology manufacturing micro-firms	Medium–low technology manufacturing small firms	Medium–low technology manufacturing medium firms	Medium–low technology manufacturing large firms
Fixed part				
Intercept	3.381***	3.367***	3.197***	3.087***
	(0.234)	(0.202)	(0.194)	(0.258)
Employment	0.943***	1.012***	0.966***	0.917***
	(0.019)	(0.014)	(0.014)	(0.016)
Capital	0.096***	0.077***	0.078***	0.102***
	(0.007)	(0.005)	(0.005)	(0.006)
Related variety	−0.030	0.084*	0.085*	−0.024
	(0.046)	(0.044)	(0.049)	(0.093)
Unrelated variety	0.091**	0.020	0.011	0.105
	(0.044)	(0.040)	(0.032)	(0.065)
Localization	0.021**	0.001	−0.017*	−0.024*
	(0.009)	(0.007)	(0.009)	(0.014)
Density	−0.004	0.000	0.037**	0.063**
	(0.011)	(0.012)	(0.014)	(0.023)
Openness	0.389***	0.099	0.098	0.023
	(0.069)	(0.071)	(0.065)	(0.104)
Education	0.213**	0.114**	0.100*	0.066
	(0.067)	(0.055)	(0.054)	(0.076)
Random part				
e_{0ijk}	0.032	0.056	0.024	0.000
	(0.020)	(0.010)	(0.024)	(0.000)
u_{0jk}	0.071	0.054	0.080	0.101
	(0.009)	(0.006)	(0.012)	(0.040)
u_{00k}	0.657	0.473	0.490	0.536
	(0.010)	(0.008)	(0.009)	(0.015)

Table 12 (continued)

Medium–low technology manufacturing	Medium–low technology manufacturing micro-firms	Medium–low technology manufacturing small firms	Medium–low technology manufacturing medium firms	Medium–low technology manufacturing large firms
Statistics				
Sample size	27,238	27,303	9870	2227
Number of regions	111	165	178	169
Number of regions * sectors	270	409	497	435

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

Table 13 Knowledge-intensive services sector for the 4 firm size groups

Knowledge-intensive services	Knowledge-intensive services micro-firms	Knowledge-intensive services small firms	Knowledge-intensive services medium firms	Knowledge-intensive services large firms
Fixed part				
Intercept	4.287***	3.802***	3.495***	2.855***
	(0.781)	(0.308)	(0.373)	(0.440)
Employment	0.916***	0.938***	0.856***	0.852***
	(0.025)	(0.019)	(0.024)	(0.015)
Capital	0.157***	0.110***	0.111***	0.108***
	(0.021)	(0.011)	(0.007)	(0.010)
Related variety	0.005	0.215**	0.190*	0.269**
	(0.178)	(0.074)	(0.100)	(0.117)
Unrelated Variety	−0.137	−0.083	0.024	−0.023
	(0.145)	(0.055)	(0.064)	(0.088)
Localization	0.037**	0.003	−0.028**	−0.046**
	(0.012)	(0.009)	(0.012)	(0.018)
Density	0.070	0.050**	0.092***	0.034
	(0.052)	(0.022)	(0.025)	(0.031)
Openness	−0.047	−0.107	−0.075	−0.204
	(0.241)	(0.097)	(0.126)	(0.154)
Education	0.148	0.088	0.040	0.334**
	(0.217)	(0.089)	(0.086)	(0.127)
Random part				
e_{0jk}	0.254	0.068	0.037	0.000
	(0.289)	(0.019)	(0.059)	(0.000)
u_{0jk}	0.185	0.103	0.117	0.179
	(0.059)	(0.041)	(0.024)	(0.137)
u_{00k}	0.800	0.586	0.599	0.634
	(0.022)	(0.022)	(0.018)	(0.066)

Table 13 (continued)

Knowledge-intensive services	Knowledge-intensive services micro-firms	Knowledge-intensive services small firms	Knowledge-intensive services medium firms	Knowledge-intensive services large firms
Statistics				
Sample size	16,177	8883	3592	2106
Number of regions	130	152	170	147
Number of regions * sectors	328	403	444	342

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

Table 14 Less knowledge-intensive services sector for the 4 firm size groups

Less knowledge-intensive services	Less knowledge-intensive services micro-firms	Less knowledge-intensive services small firms	Less knowledge-intensive services medium firms	Less knowledge-intensive services large firms
Fixed part				
Intercept	4.413***	3.212***	2.769***	2.662***
	(0.420)	(0.381)	(0.333)	(0.381)
Employment	0.834***	0.880***	0.777***	0.787***
	(0.023)	(0.021)	(0.025)	(0.017)
Capital	0.207***	0.143***	0.159***	0.162***
	(0.012)	(0.010)	(0.008)	(0.011)
Related variety	0.039	0.148**	0.127	0.023
	(0.082)	(0.073)	(0.080)	(0.106)
Unrelated variety	−0.114	0.012	−0.008	0.210**
	(0.077)	(0.066)	(0.049)	(0.068)
Localization	0.063***	−0.031**	−0.024	−0.075**
	(0.012)	(0.013)	(0.022)	(0.026)
Density	−0.028	0.041*	0.040	0.120***
	(0.024)	(0.024)	(0.030)	(0.034)
Openness	0.092	0.132	−0.022	−0.050
	(0.140)	(0.099)	(0.116)	(0.176)
Education	0.130	0.217**	0.358***	0.080
	(0.102)	(0.109)	(0.085)	(0.098)
Random part				
e_{0jk}	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
u_{0jk}	0.208	0.261	0.220	0.210
	(0.027)	(0.033)	(0.036)	(0.026)
u_{00k}	0.854	0.529	0.541	0.566
	(0.023)	(0.029)	(0.030)	(0.025)

Table 14 (continued)

Less knowledge-intensive services	Less knowledge-intensive services micro-firms	Less knowledge-intensive services small firms	Less knowledge-intensive services medium firms	Less knowledge-intensive services large firms
Statistics				
Sample size	15,783	9373	3342	1373
Number of regions	147	166	177	157
Number of regions * sectors	398	454	434	328

Robust standard errors in parentheses; *p<0.01; **p<0.05; ***p<0.01

References

Aarstad J, Kvitastein OA, Jakobsen SE (2016) Related and unrelated variety as regional drivers of enterprise productivity and innovation: a multilevel study. Res Policy 45(4):844–856

Acemoglu D, Dell M (2010) Productivity differences between and within countries. Am Econ J Macroecon 2:169–188

Baldwin R, Okubo T (2006) Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. J Econ Geogr 6:323–346

Barca F, McCann P, Rodriguez-Pose A (2012) The case for regional development intervention: place-based versus place-neutral approaches. J Reg Sci 52(1):134–152

Basile R, Capello R, Caragliu A (2012) Technological interdependence and regional growth in Europe: proximity and synergy in knowledge spillovers. Pap Reg Sci 91:697–722

Behrens K, Duranton G, Robert-Nicoud F (2014) Productive cities: sorting, selection and agglomeration. J Polit Econ 122:507–553

Bishop P, Gripaios P (2010) Spatial externalities, relatedness and sector employment growth in Great Britain. Reg Stud 44:443–454

Boschma R (2005) Proximity and innovation: a critical assessment. Reg Stud 39:61–74

Boschma R, Capone G (2015) Institutions and diversification: related versus unrelated diversification in a varieties of capitalism framework. Res Policy 44:1902–1914

Boschma RA, Iammarino S (2009) Related variety, trade linkages, and regional growth in Italy. Econ Geogr 85:289–311

Boschma RA, Balland P-A, Kogler DF (2015) Relatedness and technological change in cities: the rise and fall of technological knowledge in us metropolitan areas from 1981 to 2010. Industrial Corporate Change 24(1):223–250

Burger MJ, Van Oort FG, Raspe O (2011) Agglomeration and new establishment survival: a mixed hierarchical and cross-classified model. In: Kourtit K, Nijkamp P, Stough RR (eds) Drivers of innovation, entrepreneurship and regional dynamics. Springer, Berlin, pp 45–63

Cainelli G, Iacobucci D (2012) Agglomeration, related variety, and vertical integration. Econ Geogr 88(3):255–277

Cainelli G, Ganau R, Iacobucci D (2016) Do geographic concentration and vertically related variety foster firm productivity? Micro-evidence from Italy. Growth Change 47(2):197–217

Castaldi C, Frenken K, Los B (2015) Related variety, unrelated variety and technological breakthroughs. An analysis of US state-level patenting. Reg Stud 49:767–781

Ciccone A (2002) Agglomeration effects in Europe. Eur Econ Rev 46:213–228

Ciccone A, Hall R (1996) Productivity and the density of economic activity. Am Econ Rev 86:54–70

Cingano F, Schivardi F (2004) Identifying the sources of local productivity growth. J Eur Econ Assoc 2:720–742

Combes PP, Duranton G, Gobillon L, Puga D, Roux S (2012) The productivity advantages of large cities: distinguishing agglomeration from firm selection. Econometrica 80:2543–2594

Content J, Frenken K (2016) Related variety and economic development: a literature review. Eur Plan Stud 24:2097–2112

Corrado L, Fingleton B (2012) Where is the economics in spatial econometrics? J Reg Sci 52:2010–2239

Cortinovis N, van Oort F (2015) Variety, economic growth and knowledge intensity of European regions: a spatial panel analysis. Ann Reg Sci 55:7–32

Cortinovis N, Xiao J, Boschma R, van Oort F (2017) Quality of government and social capital as drivers of regional diversification in Europe. J Econ Geogr. https://doi.org/10.1093/jeg/lbx001

De Groot HLF, Poot J, Smit M (2015) Which agglomeration externalities matter most and why? J Econ Surv 30:756–782

Dietrich A (2012) Does growth cause structural change, or is it the other way around? A dynamic panel data analysis for seven OECD countries. Empir Econ 43(3):915–944

Diodato D, Weterings ABR (2014) The resilience of regional labour markets to economic shocks: exploring the role of interactions among firms and workers. J Econ Geogr 15:723–742

Ellison G, Glaeser E, Kerr W (2010) What causes industry agglomeration? Evidence from coagglomeration patterns. Am Econ Rev 100:1195–1213

Faggio G, Silva O, Strange WC (2017) Heterogeneous agglomeration. Rev Econ Stat 99:80–94

Fan JPH, Lang LHP (2002) The measurement of relatedness: an application to corporate diversification. J Bus 73:629–660

Feser EJ (2003) What regions do rather than make: a proposed set of knowledge-based occupation clusters. Urban Stud 40(10):1937–1958

Frenken K, van Oort F, Verburg T (2007) Related variety, unrelated variety and regional economic growth. Reg Stud 41:685–697

Fritsch M, Kublina S (2018) Related variety, unrelated variety and regional growth: the role of absorptive capacity and entrepreneurship. Reg Stud 52(10):1360–1371. https://doi.org/10.1080/00343404.2017.1388914

Glaeser E, Kallal H, Sheinkman J, Schleifer A (1992) Growth in cities. J Polit Econ 100:1126–1152

Hartog M, Boschma R, Sotarauta M (2012) The impact of related variety on regional employment growth in Finland 1993–2006: high-tech versus medium/low-tech. Ind Innov 19:459–476

Hausmann R, Hidalgo CA (2011) The network structure of economic output. J Econ Growth 16:309–342

Henderson V (2003) The urbanization process and economic growth: the so-what question. J Econ Growth 8:47–71

Henderson JV, Kuncoro A, Turner M (1995) Industrial development in cities. J Polit Econ 103:1067–1090

Hidalgo C, Hausmann R (2009) The building blocks of economic complexity. Proc Natl Acad Sci 106:10570–10575

Hidalgo CA, Klinger B, Barabàsi AL, Hausmann R (2007) The product space conditions the development of nations. Science 317:482–487

Innocenti N, Zampi V (2019) What does a start-up need to grow? An empirical approach for Italian innovative start-ups. Int J Entrepr Behav Res 25(2):376–393

Jacobs W, Koster H, van Oort F (2013) Co-agglomeration of knowledge intensive business services and multinational enterprises. J Econ Geogr 14:443–475

Kalemli-Özcan S, Sørensen B, Villegas-Sanchez C, Volosovych V, Yeşiltaş S (2015) How to construct nationally representative firm level data from the ORBID Global Databse. Tinbergen Discussion Paper 2015-110

Kemeny T, Storper M (2015) Is specialization good for regional economic development? Reg Stud 49:1003–1018

Kim SJ (1997) Productivity of cities. Ashgate, Alderhot

Knoben J, Arikan A, van Oort F, Raspe O (2016) Agglomeration and firm performance: one firm's medicine is another firm's poison. Environ Plan A 48:132–153

Kolko J (2010) Urbanization, agglomeration, and co-agglomeration of service industries. In: Gleaser E (ed) The economics of agglomeration. National Bureau of Economic Research Press, Cambridge, pp 151–180

Maksimovic Vojislav, Phillips Gordon (2002) Do conglomerate firms allocate resources inefficiently? J Finance 57:721–767. https://doi.org/10.1111/1540-6261.00440

Marrocu E, Paci R, Usai S (2013) Productivity growth in the old and new Europe: the role of agglomeration externalities. J Reg Sci 53:418–442

Melo P, Graham DJ, Noland RB (2009) A meta-analysis of estimates of urban agglomeration economies. Reg Sci Urban Econ 39:332–342

Moretti E (2004) Human capital externalities in cities. In: Henderson JV, Thisse J (eds) Handbook of regional and urban economics, vol 4. North-Holland, Amsterdam, pp 2243–2291

Mundlak Y (1978) On the pooling of time series and cross-sectional data. Econometrica 46:69–86

Neffke F, Henning M (2013) Skill relatedness and firm diversification. Strateg Manag J 34(3):297–316. https://doi.org/10.1002/smj.2014

Neffke F, Henning M, Boschma R (2011) How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. Econ Geogr 87:237–265

Neffke F, Hartog M, Boschma R, Henning M, (2014) Agents of structural change. The role of firms and entrepreneurs in regional diversification. Papers in evolutionary economic geography 14.10, Utrecht University

Neumark D, Simpson H (2015) Place-based policies. In: Duranton G, Henderson JV (eds) Handbook of regional and urban economics, vol 5. Elsevier, Amsterdam

Puga D (2002) European regional policies in light of recent location theories. J Econ Geogr 2(4):373–406

Puga D (2010) The magnitude and causes of agglomeration economies. J Reg Sci 50:203–219

Rauch JM (1993) Productivity gains from geographic concentration of human capital: evidence from the cities. J Urban Econ 34:380–400

Rice P, Venables AJ, Patacchine E (2006) Spatial determinants of productivity: analysis for the regions of Great Britain. Reg Sci Urban Econ 36:727–752

Rigby D (2012) The geography of knowledge relatedness and technological diversification in U.S. cities. Papers in evolutionary economic geography no. 12.18. Utrecht University, Utrecht

Saviotti P, Pyka A (2004) Economic development by the creation of new sectors. J Evolut Econ 14:1–35

Saviotti PP, Pyka A (2010) The Co-evolution of Innovation, demand and growth, paper presented at conference 'Technical Change, History, Economics and Policy' to be held in SPRU, Freeman Centre, University of Sussex, p 2010

Schoar A (2002) Effects of corporate diversification on productivity. J Finance 57:2379–2403

Shapiro JM (2006) Smart cities: quality of life, productivity, and the growth effects of human capital. Rev Econ Stat 88:324–335

Smith A (1776) An inquiry into the nature and causes of the wealth of nations. Oxford University Press, Oxford **(reprint 2008)**

Snijders TA, Berkhof J (2008) Diagnostic checks for multilevel models, Chapter 3. In: de Leeuw J, Meijer E (eds) Handbook of multilevel analysis. Springer, Berlin, pp 141–175

Storper M (1997) The regional world. Territorial development in a global economy. The Guildford Press, New York

Syverson C (2011) What determines productivity? J Econ Lit 49:326–365

Thissen M, Lankhuizen M, Los B, van Oort F, Diodato D (2017) EUREGIO: the construction of a global input–output database with regional detail for Europe (2000–2010). TI discussion paper 2017

Timmermans B, Boschma R (2013) The effect of intra-and inter-regional labour mobility on plant performance in Denmark: the significance of related labour inflows. J Econ Geogr 9:169–190

Van Oort F (2015) Unity in variety? Agglomeration economics beyond the specialization-diversity controversy. In: Karlsson C, Andersson M, Norman T (eds) Handbook of research methods and applications in economic geography. Edward Elgar, Cheltenham

Van Oort F, Burger M, Knoben J, Raspe O (2012) Multilevel approaches and the firm-agglomeration ambiguity in economic growth studies. J Econ Surv 26:468–491

Wintjes R, Hollanders H (2010) The regional impact of technological change in 2020-synthesis report. European Commission, DG Regional Policy, Brussels

Wixe S (2015) The impact of spatial externalities: skills, education and plant productivity. Reg Stud 49:2053–2069