



Regional knowledge base and firm efficiency: Evidence from start-ups and fast-growing medium-sized firms

European Urban and Regional Studies

1–21

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09697764261451228

journals.sagepub.com/home/eur

Giedrė Dzemydaitė¹ , Patricia Laurens² 
and Laurynas Naruševičius¹ 

Abstract

This study evaluates the production efficiency of start-ups and fast-growing medium-sized firms in advanced technology (AT) sectors and examines how the regional knowledge base contributes to reducing inefficiency. Using micro-level data for the period 2009–2016, we examine production efficiency across European regions (NUTS2), sectors (two-digit NACE) and firm-size groups by applying stochastic frontier analysis. Our findings suggest that firms' fixed assets significantly affect production efficiency across firm types. Consequently, direct investments in these firms significantly shift the production possibility frontier of AT sectors throughout European regions. Among regional conditions, we included an indicator of the local knowledge base using the relatedness density measure on a sector-by-sector basis. This helped us evaluate how the combinatorial potential of regional technological capabilities influences firm performance. Our research reveals that regional conditions are particularly important for the performance of fast-growing medium-sized firms. This suggests that such firms may constitute a “missing middle” in regional innovation strategies, offering greater territorial returns than policies focusing solely on start-ups or large R&D firms. While start-ups can be efficient even in fragmented knowledge environments, medium-sized firms' growth potential is constrained without access to a broader knowledge base. This indicates a structural challenge for place-based policy, enabling knowledge flows and institutional capacity in regions to support the production efficiency of fast-growing medium-sized firms.

Keywords

production efficiency, local knowledge, start-ups, medium-sized firms, technological relatedness, advanced technology sectors, stochastic frontier analysis, European NUTS2 regions

Introduction

In recent years, the European Union (EU) has continued to face challenges in achieving its full productive potential (European Commission, 2024; Glass and Kenjegalieva, 2024). The EU's production efficiency remains below the global production frontier, with a 10–20% gap (Sanchez, 2021). Significant differences in productive performance and a wide technology gap also remain across European regions

¹Faculty of Economics and Business Administration, Vilnius University, Lithuania

²LISIS - Laboratory for Interdisciplinary studies of Science Innovation and Society, Université Gustave Eiffel – CNRS - INRAE, IFRIS, Champs-sur-Marne, France

Corresponding author:

Giedrė Dzemydaitė, Faculty of Economics and Business Administration, Vilnius University, Saulėtekio av. 9, II bld, Vilnius 01513, Lithuania.

Email: giedre.dzemydaite@evaf.vu.lt

(Burger et al., 2021). In the context of intensifying global competition, the ability to develop more complex, technology-intensive products has become increasingly important for both economically advanced and lagging regions (Balland et al., 2019; Hartmann et al., 2021).

From a policy perspective, the European Commission (2009, 2015) has introduced initiatives to strengthen Europe's innovation base and foster industrial transformation, including the Digital Single Market Strategy (COM(2015) 192 final), the Strategy for Key Enabling Technologies (COM/2009/0512 final) and smart specialisation strategies (S3 Platform). These initiatives underline the EU's long-term commitment to enhancing regional innovation capacity and competitiveness. However, despite these sustained efforts, it remains unclear to what extent strengthening regional knowledge bases and industrial capabilities translates into improved firm-level performance across diverse regions, particularly in advanced technology (AT) sectors strategic for European well-being.

Regional and evolutionary economics emphasise that knowledge creation, learning and innovation processes are spatially embedded (Asheim, 2007; Jaffe et al., 1993; Lundvall, 1992). Regions with stronger knowledge bases and institutional capacities are expected to host more productive and innovative firms. However, empirical evidence across Europe remains mixed. Even in regions with extensive R&D infrastructure and patenting activity, firms differ considerably in the extent to which they benefit from local knowledge assets (Colombelli, 2016; Holm et al., 2020). This variation suggests that the structure and nature of regional knowledge, rather than its scale alone, may be decisive for firm efficiency.

Earlier studies of regional knowledge and firm performance, especially within the knowledge spillover theory of entrepreneurship and regional innovation systems research, have mainly relied on indicators such as R&D expenditure, the number of patents or human capital levels (Acs et al., 2009; Audretsch and Lehman, 2005; Fritsch and Slavtchev, 2011). While these measures capture the size or intensity of regional knowledge assets, they reveal little about their internal structure: how technologies are related and recombined to generate innovation

(Colombelli, 2016). Building on evolutionary economic geography and economic complexity perspectives (Boschma et al., 2015; Hidalgo et al., 2009), this study considers the combinatorial view of regional knowledge, where opportunities for recombination and innovation depend on the relatedness among technologies within a region's portfolio.

From this perspective, firm performance depends not only on internal resources and capabilities but also on the combinatorial opportunities provided by the regional knowledge base (in this article, we use terms regional knowledge base and local knowledge base interchangeably). Regions with a high degree of technological relatedness are more likely to facilitate knowledge spillovers, collaborative innovation and efficient production processes, while fragmented or less related knowledge bases may limit such effects, even in R&D-intensive environments (Boschma et al., 2015; Colombelli, 2016).

To capture these dynamics, this study integrates a variable describing the relatedness density (*RD*) of the regional knowledge base into firm-level efficiency analysis. This approach enables the assessment of whether regions with more technologically related knowledge bases provide greater efficiency gains for firms and whether these effects differ across firm types.

Additionally, firm-level heterogeneity may condition how firms benefit from regional knowledge environments. Start-ups, fast-growing medium-sized firms and large R&D performers differ in capabilities, innovation strategies and dependence on local networks (Colombelli, 2016; Luo et al., 2021; Teirlinck and Khoshnevis, 2022). While start-ups and entrepreneurial ecosystems have been widely analysed, less is known about how medium-sized firms engage with regional knowledge structures. Research on this group is often aggregated under the broader small- and medium-sized enterprises category (i.e. Hwang and Oh, 2023; Teirlinck and Khoshnevis, 2022), which tends to overlook distinctions between firm types. By incorporating this dimension, the article provides a more differentiated understanding of firm heterogeneity and how efficiency outcomes vary across firm types within regional contexts.

To the best of our knowledge, this is the first study to incorporate the evolutionary concept of

technological relatedness into a stochastic frontier framework of firm efficiency, directly linking regional knowledge structures with firm-level productive performance. It thus bridges two strands of research that have developed largely separately, particularly the literature on production efficiency and productivity analysis (Coelli et al., 2005; Farrell, 1957; Kumbhakar et al., 2012) and on regional diversification and relatedness (Boschma et al., 2015; Hidalgo et al., 2009).

Theoretically, the article extends evolutionary and regional economics perspectives by positioning RD as a determinant of productive efficiency. Rather than treating productivity differences as purely internal to firms, it highlights how external knowledge structures embedded in regions help reduce inefficiency. Empirically, it applies an integrated approach linking firm-level efficiency evaluation with regional measures of technological relatedness, providing new evidence on how regional knowledge bases affect firms of different types.

The empirical analysis focuses on five economic sectors at the NACE two-digit level that are important from an EU regional innovation policy perspective: (1) pharmaceuticals, (2) chemicals, (3) machinery, (4) automotive and (5) computer, electronic and optical products. These sectors were selected based on the following criteria: (1) the sector was a smart specialisation area in various EU regions (S3 Platform), (2) ATs were developed in that sector and (3) new knowledge was typically patented in that sector. To evaluate firm-level production efficiency in these industries, the analysis applies a stochastic frontier analysis (SFA) for the period 2009–2016, with separate estimations by sector and firm type to capture heterogeneity across regional contexts.

The remainder of this article is organised as follows. The next section discusses the related literature; the third section describes the data, methodological approach and model; the fourth section presents empirical results and the fifth section concludes the study.

Literature review

The production efficiency of firms reflects their ability to transform their inputs into outputs relative to

the most productive, or attainable, frontier (Coelli et al., 2005; Farrell, 1957). According to the efficiency and productivity literature (Kumbhakar et al., 2012; Sickles and Zelenyuk, 2019), efficiency differences depend on how effectively firms use internal resources. Factors such as capital intensity, human capital, firm size, management quality and innovation capacity influence how efficiently firms transform inputs into outputs.

From a theoretical perspective, these factors originate from the neoclassical growth framework (Solow, 1956), in which efficiency differences reflect variations in technology and factor use. Technological progress, treated as exogenous, shifts the production frontier outwards and explains differences in output not accounted for by capital and labour alone. Later, endogenous growth theory (Howitt and Aghion, 1998; Romer, 1986) demonstrated that technological progress can arise within the economic system, linking productivity growth to investment in knowledge, R&D and human capital. In this view, efficiency represents not only optimal input allocation but also firms' capacity to generate, absorb and apply knowledge, which is central to models of cumulative innovation and learning.

Building on these theoretical foundations, more recent approaches in regional and evolutionary economics have shifted attention towards the micro-foundations of productivity growth, recognising that technological progress is spatially and institutionally embedded (Balland et al., 2019; Boschma et al., 2015; Lundvall, 1992; Nelson and Winter, 1982). In technology-intensive industries, production efficiency depends not only on resource allocation but also on knowledge dynamics. Such industries are characterised by high R&D intensity, intangible assets and scientific labour, where efficiency relies on firms' capacity to innovate under uncertainty and manage high innovation costs (Camisón-Haba et al., 2019; Dzemydaitė and Naruševičius, 2023). Consequently, efficiency can be viewed as influenced by internal factors related to firms' resource endowment and productive capacity and external factors, such as access to regional knowledge, collaboration networks and institutional quality.

From a regional economics perspective, firm efficiency is conditioned by the characteristics of the surrounding innovation system (Asheim et al., 2019;

Fritsch and Slavtchev, 2011). Firms do not operate in isolation but benefit from agglomeration economies and the knowledge spillovers among co-located actors. Spatial proximity facilitates interactive learning, informal knowledge exchange and the diffusion of tacit knowledge, reducing search costs and uncertainty in innovation processes (Fritsch and Slavtchev, 2011; Jaffe et al., 1993). By improving access to complementary capabilities and lowering coordination and experimentation costs, such spillovers may allow firms to operate closer to the production frontier.

Building on these insights, several regional factors have been identified as particularly relevant for firm-level efficiency, including agglomeration and market size, which create shared labour and input markets and reduce transaction costs (Amara, 2023; Broekel, 2015; Marshall, 1920), with agglomeration benefits varying between diversified and specialised regional contexts and across stages of firm development (Duranton and Puga, 2001); institutional and human-capital conditions, which enable learning, collaboration and technology diffusion (Capello and Lenzi, 2015); and the regional stock of knowledge, specifically, the regional knowledge base, which provides more opportunities for innovation and firm performance (Acs et al., 2009; Audretsch, 1995). Among these factors, the regional knowledge base merits further discussion. The knowledge spillover theory of entrepreneurship, conceptualised by Audretsch (1995), Audretsch and Lehman (2005) and Acs et al. (2009), implied a relationship between knowledge spillovers and new firm formation, stating that a larger local knowledge base could provide more entrepreneurial opportunities. This literature typically captured the extent of local knowledge using measures such as R&D investments, human capital characteristics, research activities by universities and research centres and intensity of patents (Acs et al., 2009; Bishop, 2012; Bonaccorsi et al., 2014). However, these variables reflect the size of the local knowledge base rather than its structure or nature (Colombelli, 2016).

A second branch of research challenges the assumption that technological knowledge is homogeneous, instead emphasising its heterogeneous and combinatorial nature. According to the

recombinant knowledge approach, new knowledge emerges through the combination of related technologies within a region (Boschma et al., 2015; Hidalgo and Hausmann, 2009). In this view, a local knowledge base can be represented as a network of interconnected technologies, where the degree of relatedness reflects opportunities for recombination and capability upgrading. Regions with a broader and more coherent network of related technologies provide a richer environment for firms to integrate complementary knowledge into new products and processes.

While new technologies build on prior knowledge, regions differ in how their knowledge is structured. Regions characterised by technologically related activities tend to form dense networks of complementary capabilities that facilitate knowledge recombination (Boschma et al., 2015). In contrast, regions with fragmented technological portfolios composed of weakly related or isolated activities offer fewer opportunities for such recombination. The related variety literature emphasises that the degree of technological relatedness within a regional knowledge base shapes the potential for combining existing capabilities into new products and processes (Boschma et al., 2015; Colombelli, 2016). Empirical studies using RD as a proxy for the regional knowledge base confirm that such relatedness enhances total factor productivity and supports innovation and diversification processes (Cieřlik et al., 2025; Laffi and Boschma, 2022; Dzemydaitė et al., 2025). In this study, the local knowledge base is measured through RD, capturing the combinatorial potential of regional technological capabilities. The expectation is that firms located in regions with higher RD benefit from a more favourable environment for innovation and technological upgrading, which reduces inefficiency. To the best of our knowledge, the recombinant knowledge and relatedness frameworks have not yet been sufficiently applied to studies of production efficiency, particularly in assessing how the regional knowledge base contributes to firms' productive performance.

However, the strength of this influence is likely to differ across firm types. Start-ups often rely on local entrepreneurial ecosystems and are sensitive to

immediate knowledge and financial resources (Colombelli, 2016; Luo et al., 2021). Fast-growing medium-sized firms benefit from diversified regional networks that support scale-up and capability accumulation (Teirlinck and Khoshnevis, 2022). Large R&D-intensive firms, while possessing greater absorptive capacity, often operate within global knowledge networks and depend less on local structures (Holm et al., 2020). This differentiation reflects the interaction between internal and external efficiency mechanisms.

Accordingly, firm performance may be influenced not only by the local knowledge base and its structure but also by access to external knowledge flows. The cluster literature emphasises that external pipelines connecting firms to distant knowledge sources may provide complementary capabilities that are not available locally, thereby introducing new technological opportunities (Bathelt et al., 2004). Recent contributions to the relatedness literature suggest that interregional linkages with regions having complementary technological capabilities can facilitate diversification and technological upgrading (Balland and Boschma, 2021). However, firms differ in their ability to access such external pipelines, as the establishment and maintenance of external linkages require substantial time and resources. Larger firms typically maintain broader international networks that facilitate access to external knowledge sources, whereas smaller firms are often more dependent on locally available knowledge structures. Combining these elements, the conceptual framework assumes that production efficiency depends on firm-level resource endowment and productive capacity, regional knowledge structures and the heterogeneous influence of regional conditions across different firm types. This perspective highlights that efficiency reflects not only how firms use their own resources but also how regional conditions contribute to reducing inefficiency and supporting productive performance. Building on these insights, the following hypotheses are proposed:

H1: Technological relatedness within the regional knowledge base positively affects the production efficiency of innovating firms.

H2: The effect of technological relatedness within the regional knowledge base on production efficiency differs across firm types.

Methods and data

Data

This study uses harmonised firm-level, regional and patent data covering the EU, Norway and Switzerland over the period 2009–2016. Firm-level information was derived from two Research Infrastructure for Science and Innovation Policy Studies (RISIS) Core Facility databases: VICO (for start-ups) and Cheetah (for fast-growing medium-sized firms). Regional indicators were obtained from Eurostat, and patent data were taken from the RISIS Patent database (Table A1 in Appendix A).

Firm-level data. The VICO database includes 29,974 early-stage, venture-capital-backed firms in 27 EU countries, the United Kingdom and Israel for the period 1998–2014 (RISIS VICO dataset documentation, Tenca, n.d.). To identify companies, investors and investment deals, VICO relies on three main sources: Thomson One Private Equity, Zephyr and Crunchbase.

The Cheetah database covers 42,369 firms classified as medium-sized and fast-growing, defined by the OECD threshold of average annual employment or turnover growth above 20% over 3 years. A medium-sized firm is defined by combining the Eurostat small medium enterprises (SME) criteria with the French *entreprise de taille intermédiaire* definition. Fast-growing medium-sized firms are identified as those that experienced rapid growth in employment or turnover in at least one of the periods 2008–2011, 2009–2012 or 2010–2013 (RISIS Cheetah dataset documentation, Guerini, n.d.). The main data source for medium-sized firms is the Orbis database. Firms with missing turnover information were removed during the RISIS data construction stage. Missing employment values were imputed by RISIS developers using predicted values based on firm turnover, age, industry, country and year fixed effects.

Regional and patent-based data. Patent data were drawn from the RISIS Patent database, which

includes all priority patents applied for in all national or regional patent offices. Patents were geocoded according to the inventors' addresses, which represent the geographical locations where the invention was developed. These inventor-based locations are then assigned to NUTS2 regions, capturing the spatial distribution of local inventive activity underlying the regional knowledge base. Regional indicators were extracted from Eurostat to capture the economic and innovation environment of European NUTS2 regions.

Mapping and sample construction. All RISIS firm-level databases provide harmonised indicators such as firms' assets, employment and value added. Each firm was geographically assigned to a NUTS2 region according to its registered address, enabling the firm-level data to be linked to regional characteristics. This approach is more appropriate for start-ups and medium-sized firms than for large R&D firms, as activities of these firms are typically concentrated within a limited number of locations. Accordingly, the econometric analysis focuses on start-ups and medium-sized firms, while large firms are used only for contextual comparison.

Firms were also assigned to economic sectors according to their main NACE Rev. 2 activity codes. For start-ups (VICO) and fast-growing medium-sized firms (Cheetah), these codes were taken from the original firm-level sources and harmonised within the RISIS infrastructure. The RISIS facilities provided geocoded data with NUTS2-level regional identifiers and NACE Rev. 2 sector codes, which we applied directly in the analysis.

All procedures related to data construction and treatment of missing information are documented in the official RISIS database manuals; we used the harmonised data as provided. Because firms are observed for different numbers of years, the resulting panel is unbalanced. As the firm data originate from the RISIS Core Facility, the analysis is limited to the information available within these databases.

Tables D1–D5 in Appendix D provide descriptive statistics for the main variables and sample. As expected, the average number of employees increases with the firm size. Start-ups show the highest capital intensity, with the largest fixed assets per employee

across all economic sectors. Medium-sized firms are the least capital-intensive group, with lower fixed assets, except in the chemical sector, and lower productivity levels than start-ups and large firms. In comparison, large R&D corporate performers have the highest employment levels and productivity. At the regional level, indicators such as RD, R&D expenditure, population density and industrial diversity reveal substantial variation across NUTS2 regions, reflecting uneven development conditions in Europe.

Selection of economic sectors

Economic sectors were selected according to the policy interest and nature of patenting activities regarding new knowledge, based on the following criteria. Firstly, this study included economic sectors contributing to the development of ATs. The Advanced Technologies for Industry (ATI) project, launched by the EU to monitor technological trends and provide current data on ATs, identified 16 technologies (key enabling and digital) expected to substantially transform business and society (Izsak et al., 2020). Based on the ATI project's data on patenting activities across ATs that could be attributed to economic sectors via the applicant firms, we identified six economic sectors where new knowledge in ATs was patented: medical devices, pharmaceuticals, automotive, chemicals, machinery and electronic sectors. The share of AT patents in these sectors varied between approximately 1% and 17% (Table 1).

We then associated the general sector names from the ATI project with NACE Rev. 2 two-digit classifications. It was difficult to associate medical devices with a single NACE division, as related activities are distributed across several sectors (mainly C32.5 and C26.6). Therefore, we included five economic sectors (NACE two-digit) in the analysis. Activities related to medical devices were partly covered under C26 (Manufacture of computers, electronic and optical products).

The second step in evaluating the importance of these sectors involved assessing their role in regional smart specialisation priorities. Evidence from the S3 Platform indicates that these sectors have high

Table 1. Economic sectors and their role in AT and smart specialisation strategies.

Economic sector (NACE two digits)	AT share of patents with known sectoral attribution in overall patents in EU27, 2017	Number of EU regions with encoded S3 priorities
C20 Manufacture of chemicals and chemical products	5% (Chemicals)	58
C21 Manufacture of basic pharmaceutical products and pharmaceutical preparations	2% (Pharma)	68
C26 Manufacture of computer, electronic and optical products	17% (Electronics)	88
C28 Manufacture of machinery and equipment	12% (Machinery)	107
C29 Manufacture of motor vehicles, trailers and semi-trailers	4% (Automotive)	80

Source: ATI Data Dashboard by EU, S3 Platform Eye@RIS3, authors' calculation.

AT: advanced technology; ATI: Advanced Technologies for Industry; EU: European Union.

policy relevance across Europe, with 58–107 regions identifying them as priority areas (Table 1).

An approach to production efficiency modelling

Production efficiency modelling stems from the seminal work by Farrell (1957), whose initial idea was to examine how far a firm or sector could increase its outputs without increasing inputs by improving its efficiency. Over the past few decades, parametric and non-parametric approaches have been widely applied in efficiency studies, both having advantages and limitations (Sickles and Zelenyuk, 2019). Parametric models, such as the SFA, enable the estimation of production elasticities and identification of factors influencing inefficiency while controlling for statistical noise. External conditions and their influence on firm-level performance can also be incorporated into this type of analysis (Mandl et al., 2008). Conversely, non-parametric methods do not provide insights into the elasticities of explanatory variables and are limited in their ability to capture time dynamics and slacks (Sickles and Zelenyuk, 2019).

Given that our data combine firm-level and regional information over several years, we employ a panel SFA model to analyse production efficiency. This approach is well-suited to the research context, as it captures changes in efficiency over time and enables the estimation of production elasticities and inefficiency determinants within a single framework

(Battese and Coelli, 1995; Sickles and Zelenyuk, 2019). SFA was selected for its ability to capture firm-level inefficiency and shifts in the available frontier technologies (Mandl et al., 2008; Sickles and Zelenyuk, 2019). Moreover, stochastic frontier modelling is a common technique for measuring technical efficiency in regional studies (Glass and Kenjegaliev, 2024).

A key assumption of SFA is that all firms in a given industry share the same production technology. Because our sample involves different sectors, we estimated frontiers separately by sector rather than using one pooled model. In SFA, the estimated frontier represents the best-practice production technology under a given input–output structure. If sectors differ technologically, estimating a common frontier assumes a single production benchmark across industries, which may bias efficiency comparisons. For example, the pharmaceutical sector is characterised by long R&D cycles and strict regulation, whereas the electronics sector operates under short product life cycles, reflecting distinct production processes. Sector-specific estimation, therefore, improves precision by allowing each industry's technology and inefficiency parameters to vary (Coelli et al., 2005; Kumbhakar et al., 2012).

The model and variables

We apply a one-step estimation procedure rather than the two-step approach, as this enables inefficiency factors to be estimated jointly with the

production frontier, avoiding potential bias arising from the correlation between inefficiency effects and the error term (Battese and Coelli, 1995). We assume a translog production function with inefficiency terms following a half-normal distribution, consistent with standard applications in firm-level efficiency studies (Battese and Coelli, 1995; Mandl et al., 2008; Sickles and Zelenyuk, 2019). This approach provides consistent estimates of how firm-level characteristics and regional conditions affect production efficiency.

The SFA model follows the one-step estimation specification introduced by Battese and Coelli (1995). The model for firm i operating in sector s , region r and year t is specified as:

$$Y_{isrt} = \exp(X_{isrt}\beta + v_{isrt} - u_{isrt}), \quad (1)$$

where Y_{isrt} is the firm's production output, X_{isrt} is a vector of inputs and v_{isrt} is a random error term distributed as $N(0, \sigma_v^2)$. The inefficiency term is as follows:

$$u_{isrt} = Z_{isrt}\delta + w_{isrt}, w_{isrt} \sim N^+(0, \sigma_u^2) \quad (2)$$

where Z_{isrt} includes the determinant of inefficiency. A negative coefficient in the inefficiency equation indicates that the corresponding variable reduces inefficiency (i.e. increases efficiency).

The dependent variable was value added per employee, which reflects firm-level productivity. Per employee values are considered to standardise data and eliminate firm-size effects (Crépon et al., 1998; Kumbhakar et al., 2012). Firm-level inputs include fixed assets per employee (FA), approximating capital intensity and the number of employees (E), which captures scale effects in production (Kumbhakar et al., 2012). These variables were taken from the RISIS databases and expressed in logarithmic form.

Fixed assets consist of both tangible and intangible components. A tangible asset represents physical capital used in production, such as machinery, equipment and buildings (International Accounting Standard 16). An intangible asset, by contrast, is a non-monetary asset without physical substance, including items such as patented technologies,

trademarks, licences, software and trade secrets (International Accounting Standard 38). Intangible assets constitute an important component of firm performance because they represent the accumulated results of past knowledge investments and innovation efforts and therefore form a stock of knowledge capital used in current production.

However, not all intangible assets, understood as knowledge-based capital, are fully captured in official financial reports (Corrado et al., 2022). Due to inconsistencies in accounting practices, internally generated intangibles are often excluded from official financial reports. Productivity studies, such as those conducted under the EUKLEMS and INTANProd projects, have sought to address these limitations by providing a measurement of intangible assets (Bontadini et al., 2023). However, these data are not available at the micro level for our sample. Therefore, our analysis is constrained to the intangible assets reported in firms' official financial accounts.

Some studies measure innovation activity using R&D expenditures or patent counts. Such indicators, however, are not consistently available in the harmonised microdata used in this study. We therefore rely on accounting-based measures of intangible assets reported in firms' financial statements, which provide a proxy for the accumulated stock of knowledge capital resulting from past R&D and innovation investments and are treated here as productive inputs in the estimation of the production frontier. As these assets reflect accumulated firm-level investments rather than current innovation outputs, they represent knowledge capital used in production. An inefficiency term u_{isrt} is modelled as a function of regional characteristics that may impact firms' capacity to transform inputs into outputs. These regional-level indicators include the local knowledge base in a particular economic sector (RD) besides the other contextual conditions.

The local knowledge base (RD) was evaluated using a proximity approach originally developed by Hidalgo et al. (2007) to measure product space and later widely used in studies of technology space and industry diversification. Patent data were used as a proxy for knowledge creation and inventions (Boschma et al., 2015). First, the RD was calculated

for each technological domain at the NUTS2 regional level following Boschma et al. (2015) using the EconGeo package in R (Balland, 2017). An indicator relevant to economic sectors rather than specific technologies was obtained by aggregating these technology-level values to the industry level by weighting each technology's RD based on its share of patents within the corresponding sector. The weights were derived from patenting data for large R&D-performing firms, which enabled patents to be matched to firms and firms to their NACE sectors. A higher RD indicator in a region reflects a stronger local knowledge base and greater potential for knowledge recombination. Appendix C details the evaluation of RD .

Besides RD , we included regional variables that are commonly employed in regional and innovation studies to capture key contextual conditions influencing firms' productive efficiency (Amara, 2023; Broekel, 2015; Ketels and Protsiv, 2021). Population density (POP_DEN) proxies urbanisation economies and agglomeration effects (Broekel, 2015; Marshall, 1920). The gross domestic product per capita (GDP) reveals the size of the local market and regional demand conditions that shape firms' scale opportunities. The society's tertiary education (EDU), considering the quality of human capital and total intramural expenditure on R&D ($GERD$), capture the knowledge-generation capacity of regions, which is central to regional innovation systems (Asheim et al., 2019; Fritsch and Slavtchev, 2011).

Potential effects of a region's industrial agglomeration were accounted for by estimating the location coefficient of employees in a particular economic sector ($SPEC_LQ$). This indicator reflects sectoral specialisation and firms' exposure to Marshall–Arrow–Romer (MAR) localisation externalities and was calculated as:

$$SPEC_LQ_{rs} = \frac{(E_{rs} / E_r)}{(E_s / E)}, \quad (3)$$

where E_{rs} is employment in sector s in region r , E_r is total employment in region r , E_s is total employment in sector s and E is total employment in all regions. Additionally, to capture the advantages related to industrial diversification, we used an

inverted Hirschman–Herfindahl index ($DIVERSE_iHHI$), which represents Jacobs (1969) externalities arising from knowledge spillovers across diverse industries:

$$DIVERSE_iHHI_r = 1 / \sum_s \left(\frac{E_{rs}}{E_r} \right)^2. \quad (4)$$

Together, these regional variables reveal key structural features of local environments that may influence firms' productive efficiency.

Assuming a log-linear relationship, the empirical SFA model specification could then be summarised as:

$$\ln(Y)_{istr} = b_0 + b_1 \ln(FA)_{istr} + b_2 \ln(E)_{istr} + t + v_{istr} - u_{istr}, \quad (5)$$

where Y_{istr} = value added per employee at firm i in industry s at time period t in region r ($i = 1 \dots N$, $s = 1 \dots N$ and $t = 1 \dots T_i$), FA = fixed assets per employee,¹ E = number of employees; t = time trend for Hicks-neutral technical change; v_{istr} = random noise $\mathcal{N}(0, \sigma_v^2)$ and u_{istr} = time-varying technical inefficiency term of production, which is assumed to be independently distributed, such that u_{it} is obtained by truncation (at zero) of the normal distribution and assumed to be specified as:

$$u_{istr} = \delta_0 + \delta_1 RD_{str} + \delta_2 SPEC_LQ_{str} + \delta_3 DIVERSE_HHI_{str} + \delta_4 \ln(GDP)_{tr} + \delta_5 \ln(GERD)_{tr} + \delta_6 EDU_{tr} + \delta_7 POP_DEN_{tr} + \omega_{istr}, \quad (6)$$

where the random variable ω_{istr} is defined by the truncation of a normal distribution with zero mean and variance σ^2 . The panel data are unbalanced, and each firm may have a different number of time periods (T_i). The equations were estimated using maximum likelihood, which simultaneously estimates the parameters of the stochastic frontier and model for the technical inefficiency effects (Battese and Coelli, 1995).

Estimation results

Tables 2 and 3 present the estimation results for equations (3) and (4), whose parameters were estimated

Table 2. Model parameters and efficiency estimates for start-ups.

Model specification	Chemicals and chemical products		Pharmaceutical products and preparations		Computer, electronic and optical products		Machinery and equipment		Motor vehicles, trailers and semi-trailers	
	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value
Constant	9.080	0.000	4.963	0.000	8.085	0.000	8.654	0.000	6.022	0.000
ln(FA)	0.115	0.007	0.331	0.000	0.168	0.000	0.050	0.030	0.228	0.000
ln(E)	0.052	0.360	0.084	0.102	0.037	0.226	-0.078	0.001	0.139	0.109
Time	-0.089	0.002	0.058	0.218	-0.001	0.957	0.006	0.707	0.085	0.000
Determinants of inefficiency										
Constant	4.350	0.301	2.273	0.125	5.308	0.158	49.573	0.268	26.153	0.027
RD	-0.041	0.000	0.015	0.001	0.000	0.979	-0.182	0.100	-0.040	0.134
SPEC_LQ	-0.108	0.466	-0.111	0.002	-0.121	0.251	-1.494	0.067	0.649	0.059
DIVERSE_iHHI	0.288	0.000	-0.196	0.000	0.124	0.003	0.585	0.058	-0.283	0.028
ln(GDP)	-0.205	0.693	0.124	0.617	-0.290	0.553	-5.599	0.257	-1.527	0.373
ln(GERD)	-0.269	0.205	-0.324	0.107	-0.385	0.193	1.537	0.176	-1.225	0.291
EDU	-0.054	0.002	0.012	0.064	-0.010	0.282	-0.075	0.188	0.009	0.525
POP_DEN	0.0000	0.896	0.0002	0.358	0.0000	0.980	-0.0070	0.195	0.0026	0.041
Mean efficiency	0.46		0.79		0.41		0.88		0.69	
Log likelihood value	-112.91		-98.53		-450.31		-125.13		-1.37	
Firms	39		32		141		59		10	
Observations	131		78		392		189		38	

EDU: education; GDP: gross domestic product per capita; POP_DEN: population density; RD: relatedness density.

Table 3. Model parameters and efficiency estimates for fast-growing medium-sized firms.

Model specification	Chemicals and chemical products		Pharmaceutical products and preparations		Computer, electronic and optical products		Machinery and equipment		Motor vehicles, trailers and semi-trailers	
	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value
Constant	6.522	0.000	7.679	0.000	7.761	0.000	8.141	0.000	6.881	0.000
ln(FA)	0.351	0.000	0.221	0.000	0.261	0.000	0.181	0.000	0.273	0.000
ln(E)	-0.024	0.096	0.055	0.028	-0.043	0.007	-0.052	0.000	-0.060	0.000
Time	-0.013	0.072	-0.019	0.183	-0.009	0.284	-0.005	0.117	0.029	0.000
Determinants of inefficiency										
Constant	15.222	0.000	13.083	0.000	11.901	0.000	17.763	0.000	11.431	0.000
RD	-0.017	0.000	-0.023	0.000	0.007	0.002	-0.005	0.002	-0.003	0.043
SPEC_LQ	-0.318	0.000	-0.244	0.004	-0.005	0.742	-0.035	0.010	0.066	0.000
DIVERSE_iHHI	0.035	0.018	0.055	0.012	0.051	0.000	0.043	0.000	0.043	0.000
ln(GDP)	-1.383	0.000	-1.015	0.000	-1.056	0.000	-1.743	0.000	-1.137	0.000
ln(GERD)	-0.034	0.525	-0.362	0.000	-0.186	0.000	0.039	0.035	0.031	0.170
EDU	-0.003	0.370	0.025	0.000	0.009	0.001	0.005	0.001	0.001	0.594
POP_DEN	0.0002	0.000	0.0002	0.000	0.0000	0.886	0.0003	0.000	0.0001	0.127
Mean efficiency	0.50		0.47		0.35		0.46		0.45	
Log likelihood value	-2092.01		-1065.17		-2331.08		-6097.72		-3655.63	
Firms	485		223		568		1615		845	
Observations	2175		969		2258		7582		4104	

EDU: education; GDP: gross domestic product per capita; POP_DEN: population density; RD: relatedness density.

simultaneously. The analysis was performed separately for start-ups and fast-growing medium-sized firms across five economic sectors.

Across both firm categories, fixed assets per employee showed a positive and statistically significant effect on output efficiency, supporting that capital intensity shifts the production frontier outwards. In medium-sized firms (Table 3), the magnitude of the coefficients ranged from 0.18 to 0.35. Start-ups (Table 2) had lower elasticities for *FA* in most of the sectors analysed (0.05–0.33), highlighting that while capital investment is beneficial, the scale of operations and nature of start-ups may limit the impact relative to larger firms. Additionally, start-up firms may employ intangible assets not accounted for in official reports (Corrado et al., 2022). Overall, this pattern indicates increasing efficiency gains from physical capital accumulation as firms grow in scale and technological maturity.

The mean efficiency score, the average of firm-specific estimates of technical efficiency² within a specific firm size and economic sector, further reflects differences between firm types. The mean efficiency score was 0.41–0.88 for start-ups and 0.35–0.5 for medium-sized firms, indicating substantial efficiency gaps across firms.

The coefficient of the number of employees (*E*) showed mixed results. In medium-sized firms (Table 3), negative elasticities (–0.06 to –0.02) suggest decreasing returns to scale or underuse of labour. Conversely, for start-ups, the coefficients were small and mainly insignificant, implying that efficiency variations are driven more by technological and managerial factors than by labour scale (Brand et al., 2021).

The determinants of inefficiency show the importance of regional and structural factors. For fast-growing medium-sized firms, the coefficient of *RD* was negative and significant in many sectors (e.g. chemicals –0.017, pharmaceutical products –0.023 and machinery –0.005). This result indicates that firms located in regions with richer interconnections among technologies achieve lower inefficiency; that is, firms tend to be more efficient in these regions.

The coefficients for specialisation and diversity show that local industrial structures shape firm efficiency differently across sizes. For medium-sized

firms, *SPEC_LQ* had negative and significant coefficients (e.g. –0.32 for chemicals), implying that local technological specialisation enhances efficiency through *MAR* externalities. Conversely, start-ups benefited from a diversified environment in the pharmaceutical and automotive sectors, supporting the concept that new firms profit from cross-sectoral knowledge recombination (Jacobs' externalities).

This pattern is consistent with Duranton and Puga's (2001) nursery-city model. Their framework suggests that new products, and thus young firms, emerge in diversified, multi-industry cities where they could borrow trying processes from different activities. As firms grow, they tend to relocate to more specialised locations that offer concentrated sector-specific know-how and lower production costs.

The regional *GDP* significantly reduced inefficiency in fast-growing medium-sized firms in all sectors (–1.0 to –1.7), while *GERD* was significant only in the pharmaceutical and computer sectors. This result indicates that local market size and demand have a more positive influence on efficiency than *R&D* intensity. Education and population density mainly had weaker or insignificant effects.

Descriptive comparisons with large *R&D* firms indicate that these firms have higher efficiency scores and operate closer to the efficiency frontier (Table D4). This result indicates that in many cases, large firms are more capital-intensive than smaller firms and receive sufficient efficiency gains from asset accumulation (Kumbhakar et al., 2012; Moncada-Paternò-Castello, 2022). However, their weaker association with regional characteristics is consistent with their multi-establishment structures and reliance on global knowledge networks, as a substantial share of inventive activities of large firms is located outside the region in which the company is registered (Table B1 in Appendix B).

Overall, the empirical analysis supports that the local knowledge base enhances production efficiency, which is consistent with Hypothesis 1. The positive impact of *RD* on efficiency is particularly visible for fast-growing medium-sized firms and, to a lesser extent, start-ups, therefore also supporting Hypothesis 2. These findings highlight that regional

knowledge recombination acts as an enabling condition for firm-level production efficiency, with its effects mediated by the firm size, technological intensity and structure of regional specialisation.

Discussion and conclusions

This article analysed the production efficiency of start-ups and fast-growing medium-sized firms in European regions, with a particular focus on the role of regional knowledge structures. By combining a stochastic frontier framework with measures of technological relatedness, the study extended productive efficiency analysis by drawing on evolutionary and regional economics. It thus examined how internal firm resources and external regional knowledge conditions jointly shaped firms' performances in AT sectors.

Three main findings emerge from the analysis. First, firm-specific tangible and intangible fixed assets consistently shift the production frontier across all firm-size groups and sectors. Second, regional knowledge conditions, particularly the structure of the local knowledge base and market size, significantly reduce inefficiency for fast-growing medium-sized firms. Third, start-ups depend mainly on internal technological capabilities. The patterns observed for start-ups are consistent with existing empirical evidence. Conversely, the pronounced and consistent sensitivity of medium-sized firms to regional knowledge structures represents the most distinctive contribution of this study.

These results highlight the structural role of fast-growing medium-sized firms in European regional development. While many EU regions have selected ATs as smart specialisation priorities, only a limited number host large global R&D leaders. The estimates show that medium-sized firms in several regions operate close to the production efficiency frontier, indicating their potential to become future innovation leaders if supported by coherent regional knowledge environments. This supports the existence of a "missing middle" in regional innovation systems, suggesting that medium-sized firms constitute an important transmission channel between local knowledge bases and regional productivity growth.

From a European competitiveness perspective, the findings imply that strengthening regional

knowledge structures primarily benefits firms large enough to scale but still locally embedded. Start-ups may succeed in niche technological domains even in fragmented knowledge environments; however, their scaling into medium-sized firms increasingly depends on access to a broader and more coherent regional knowledge base. Conversely, large R&D corporate performers typically draw predominantly on global innovation networks. This asymmetry explains why regional policies often struggle to retain and scale successful start-ups into globally competitive firms.

Several policy implications follow. Regional and innovation policies could place greater emphasis on supporting fast-growing medium-sized firms, in addition to start-ups and large R&D players. The results suggest that policy effectiveness depends not only on expanding regional R&D intensity but also on strengthening the internal relatedness of regional technological portfolios to enhance knowledge recombination. Continued support for knowledge-based fixed asset investment remains important, given its robust frontier-shifting effect across all firm types in AT sectors. Together, these findings contribute to current debates on smart specialisation and cohesion policy in regions seeking to transcend the start-up phase.

The results also have managerial implications. Managers of start-ups should prioritise the accumulation of firm-specific technological and asset-based capabilities, as these dominate efficiency performance in the early stages. Managers of medium-sized firms may benefit substantially from active engagement with regional research organisations, universities and technologically related local firms to exploit local knowledge complementarities.

Finally, this study is subject to several limitations. The analysis relies on patent-based measures of knowledge that capture only codified innovation, while tacit and symbolic knowledge remains unobserved (Asheim, 2007; Asheim et al., 2017; Howells, 2012). Moreover, the empirical design focuses primarily on the regional knowledge base and therefore cannot directly observe the role of interregional knowledge linkages or firm-internal global networks that may complement local knowledge spillovers (Balland and Boschma, 2021; Bathelt et al., 2004). Additionally, the data cover the period 2009–2016,

with RISIS data, and therefore do not fully reflect post-pandemic adjustments and digital transformation dynamics. Future research could integrate differentiated knowledge bases using mixed-method approaches to further deepen understanding of regional–firm productivity links.

Acknowledgements

The data for this work was from the database of RISIS2 (Research Infrastructure for Innovation Policy Studies 2). RISIS2 was funded by the European Union's Horizon 2020 Research and Innovation Programme under the grant number 824091. The authors are thankful to the managers of the RISIS2 datasets for providing data.

ORCID iDs

Giedrė Dzemydaitė  <https://orcid.org/0000-0002-1806-7663>

Patricia Laurens  <https://orcid.org/0000-0002-3192-7736>

Laurynas Naruševičius  <https://orcid.org/0000-0002-2971-6737>

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research project has received funding from European Social Fund project No 09.3.3-LMT-K-712-19-0090 under a grant agreement with the Research Council of Lithuania (LMTLT).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. Current stock indicators were used as determinants because firm productivity is affected by the accumulated stock rather than lagged flows, and it also captures the “learning by doing effects” (Kumbhakar et al., 2012).
2. The technical efficiency for the i -th firm at the t -th observation is defined as: $TE_{it} = \exp(-u_{it})$.

References

Acs ZJ, Braunerhjelm P, Audretsch DB, et al. (2009) The knowledge spillover theory of entrepreneurship.

Small Business Economics 32(1): 15–30. <https://doi.org/10.1007/s11187-008-9157-3>

Amara M (2023) Agglomeration and firm performance in times of economic turmoil: Evidence from Tunisian firm-level data. *Growth and Change* 54(2): 446–481.

Asheim BT (2007) Differentiated knowledge bases and varieties of regional innovation systems, innovation. *Innovation: The European Journal of Social Science Research* 20(3): 223–241. <https://doi.org/10.1080/13511610701722846>

Asheim BT, Grillitsch M and Trippi M (2017) Introduction: Combinatorial knowledge bases, regional innovation, and development dynamics. *Economic Geography* 93(5): 429–435. <https://doi.org/10.1080/00130095.2017.1380775>

Asheim BT, Isaksen A and Trippi M (2019) *Advanced Introduction to Regional Innovation Systems*. Elgar Advanced Introductions.

Audretsch DB (1995) *Innovation and Industry Evolution*. MIT Press.

Audretsch DB and Lehmann EE (2005) Does the knowledge spillover theory of entrepreneurship hold for regions? *Research Policy* 34: 1191–1202. <https://doi.org/10.1016/j.respol.2005.03.012>

Balland PA (2017) *Economic Geography in R: Introduction to the EconGeo Package*. SSRN Electronic Journal. Social Science Research Network (SSRN), Elsevier. <https://doi.org/10.2139/ssrn.2962146>

Balland PA and Boschma R (2021) Complementary interregional linkages and Smart Specialisation: An empirical study on European regions. *Regional Studies* 55(6): 1059–1070. <https://doi.org/10.1080/0343404.2020.1861240>

Balland PA, Boschma R, Crespo J, et al. (2019) Smart specialization policy in the European Union: Relatedness, knowledge complexity and regional diversification. *Regional Studies* 53(9): 1252–1268. <https://doi.org/10.1080/00343404.2018.1437900>

Bathelt H, Malmberg A and Maskell P (2004) Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography* 28(1): 31–56. <https://doi.org/10.1191/0309132504ph4690a>

Battese GE and Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20(2): 325–332. <https://doi.org/10.1007/BF01205442>

Bishop P (2012) Knowledge, diversity and entrepreneurship: A spatial analysis of new firm formation in Great Britain. *Entrepreneurship and Regional Development* 24: 641–660. <https://doi.org/10.1080/0985626.2011.617786>

- Bonaccorsi A, Colombo MG, Guerini M, et al. (2014) The impact of local and external university knowledge on the creation of knowledge-intensive firms: Evidence from the Italian case. *Small Business Economics* 43(2): 261–287.
- Bontadini F, Corrado C, Haskel J, et al. (2023) *EUKLEMS & INTANProd: Industry Productivity Accounts with Intangibles*. Sources of growth and productivity trends: Methods and main measurement challenges, Luiss Lab of European Economics.
- Boschma R, Balland PA and 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 and Corporate Change* 24(1): 223–250. <https://doi.org/10.1093/icc/dtu012>
- Brand M, Tiberius V, Bican PM, et al. (2021) Agility as an innovation driver: Towards an agile front end of innovation framework. *Review of Managerial Science* 15(1): 157–187. <https://doi.org/10.1007/s11846-019-00373-0>
- Broekel T (2015) Do cooperative research and development (R&D) subsidies stimulate regional innovation efficiency? Evidence from Germany. *Regional Studies* 49(7): 1087–1110. <https://doi.org/10.1080/00343404.2013.812781>
- Burger MJ, Kounetas K, Napolitano O, et al. (2021) Do innovation and human capital actually narrow the technology gap? Champions and laggards of European regional productive performance. *Regional Studies* 56(10): 1655–1670. <https://doi.org/10.1080/00343404.2021.2000596>
- Camisón-Haba S, Clemente-Almendros JA and Gonzalez-Cruz T (2019) How technology-based firms become also highly innovative firms? The role of knowledge, technological and managerial capabilities, and entrepreneurs' background. *Journal of Innovation & Knowledge* 4(3): 162–170. <https://doi.org/10.1016/j.jik.2018.12.001>
- Capello R and Lenzi C (2015) The knowledge–innovation nexus. Its spatially differentiated returns to innovation. *Growth and Change* 46(3): 379–399. <https://doi.org/10.1111/grow.12098>
- Cieślak E, Golejewska A and Wójcik P (2025) Two faces of human capital and research and development activities: Effects of related technologies. *European Urban and Regional Studies* 32(4): 421–442. <https://doi.org/10.1177/09697764251331563>
- Coelli TJ, Prasada Rao DS, O'Donnell CJ, et al. (2005) *An Introduction to Efficiency and Productivity Analysis*. Springer US.
- Colombelli A (2016) The impact of local knowledge bases on the creation of innovative start-ups in Italy. *Small Business Economics* 47: 383–396.
- Corrado C, Haskel J, Jona-Lasinio C, et al. (2022) Intangible capital and modern economies. *Journal of Economic Perspectives* 36(3): 3–28. <https://doi.org/10.1257/jep.36.3.3>
- Crépon B, Duguet E and Mairessec J (1998) Research, innovation and productivity: An econometric analysis at the firm level. *Economics of Innovation and new Technology* 7(2): 115–158. <https://doi.org/10.1080/10438599800000031>
- Duranton G and Puga D (2001) Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review* 91(5): 1454–1477. <https://doi.org/10.1257/aer.91.5.1454>
- Dzemydaitė G and Naruševičius L (2023) Exploring efficiency growth of advanced technology-generating sectors in the European Union: a stochastic frontier analysis. *Journal of Business Economics and Management (JBEM)* 24(6): 976–995. <https://doi.org/10.3846/jbem.2023.20688>
- Dzemydaitė G, Naruševičius L and Rubčinskaitė R (2025) Pathways to economic diversification in the baltic States: An evolutionary approach. *Eastern European Economics* 1–30. <https://doi.org/10.1080/00128775.2025.2586603>
- European Commission (2009) Preparing for our future: Developing a common strategy for key enabling technologies in the EU. {SEC(2009) 1257} /* COM/2009/0512 final */
- European Commission (2015) A digital single market strategy for Europe, COM(2015) 192 final, Brussels, May 6, 2015.
- European Commission (2024) The future of European competitiveness – A competitiveness strategy for Europe. Prepared by Mario Draghi. Available at: https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4c-f152a8232961_en (accessed 19 May 2026).
- Farrel MJ (1957) The measurement of productive efficiency. *Journal of Royal Statistics Society A120*: 253–281.
- Fritsch M and Slavtchev V (2011) Determinants of the efficiency of regional innovation systems. *Regional Studies* 45(7): 905–918. <https://doi.org/10.1080/00343400802251494>
- Glass AJ and Kenjegaliev K (2024) The role of inefficiency in a productivity puzzle: Regional evidence for Great Britain. *Journal of Regional Science* 64(4): 1287–1322. <https://doi.org/10.1111/jors.12702>
- Guerini M (n.d.) Documentation of RISIS datasets: Cheetah. Available at: https://zenodo.org/record/3337963/files/Documentation_Cheetah_Final_edited.pdf?download=1 (accessed 5 May 2024).
- Hartmann D, Zagato L, Gala P, et al. (2021) Why did some countries catch-up, while others got stuck in the middle? Stages of productive sophistication and smart

- industrial policies. *Structural Change and Economic Dynamics* 58: 1–13. <https://doi.org/10.1016/j.strueco.2021.04.007>
- Hidalgo CA and Hausmann R (2009) The building blocks of economic complexity. *Proceedings of the National Academy of Sciences* 106(26): 10570–10575.
- Hidalgo CA, Klingler B, Barabási AL, et al. (2007) The product space conditions the development of nations. *Science* 317(5837): 482–487. <https://doi.org/10.1126/science.1144581>
- Holm JR, Timmermans B, Østergaard CR, et al. (2020) Labor mobility from R&D-intensive multinational companies: implications for knowledge and technology transfer. *The Journal of Technology Transfer* 45(5): 1562–1584. <https://doi.org/10.1007/s10961-020-09776-8>
- Howells J (2012) The geography of knowledge: Never so close but never so far apart. *Journal of Economic Geography* 12(5): 1003–1020. <https://doi.org/10.1093/jeg/lbs027>
- Howitt P and Aghion P (1998) Capital accumulation and innovation as complementary factors in long-run growth. *Journal of Economic Growth* 3(2): 111–130. <https://doi.org/10.1023/A:1009769717601>
- Hwang WS and Oh S (2023) The effectiveness of public R&D subsidy on SMEs' innovation capability and catch-up in the Korean manufacturing industry. *Technology Analysis & Strategic Management* 35(10): 1369–1384. <https://doi.org/10.1080/09537325.2021.2005781>
- Izsak K, Markianidou P, van de Velde E, et al. (2020) Advanced Technologies for Industry—Methodological report. European Commission. *European Union* 2020. <https://doi.org/10.2826/911991>
- Jacobs J (1969) *The Economy of Cities*. Random House.
- Jaffe AB, Trajtenberg M and Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* 108(3): 577–598. <https://doi.org/10.2307/2118401>
- Ketels C and Protsiv S (2021) Cluster presence and economic performance: A new look based on European data. *Regional Studies* 55(2): 208–220. <https://doi.org/10.1080/00343404.2020.1792435>
- Kumbhakar SC, Ortega-Argilés R, Potters L, et al. (2012) Corporate R&D and firm efficiency: Evidence from Europe's top R&D investors. *Journal of Productivity Analysis* 37: 125–140. <https://doi.org/10.1007/s11223-011-0223-5>
- Laffi M and Boschma R (2022) Does a local knowledge base in Industry 3.0 foster diversification in Industry 4.0 technologies? Evidence from European regions. *Papers in Regional Science* 101(1): 5–35. <https://doi.org/10.1111/pirs.12643>
- Lundvall BA (1992) *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*, vol. 242. Pinter.
- Luo X, Huang F, Tang X, et al. (2021) Government subsidies and firm performance: Evidence from high-tech start-ups in China. *Emerging Markets Review* 49: 100756. <https://doi.org/10.1016/j.ememar.2020.100756>
- Mandl U, Dierx A and Ilzkovitz F (2008) The effectiveness and efficiency of public spending (Economic Papers No 301). European Commission.
- Marshall A (1920) *Principles of Economics*. Palgrave Macmillan UK.
- Moncada-Paternò-Castello P (2022) Corporate R&D intensity, size, and performance in a changing world: A review of the evidence. *Eurasian Business Review* 12(2): 229–254. <https://doi.org/10.1007/s40821-022-00206-3>
- Nelson RR and Winter SG (1982) The Schumpeterian tradeoff revisited. *The American Economic Review* 72(1): 114–132.
- Romer PM (1986) Increasing returns and long-run growth. *Journal of Political Economy* 94(5): 1002–1037.
- Sanchez M (2021) What drives euro area labour productivity growth?: An international production-frontier examination. In: Kostis PC (ed.) *Bridging Microeconomics and Macroeconomics and the Effects on Economic Development and Growth*. IGI Global, pp.231–258. <https://doi.org/10.4018/978-1-7998-4933-9.ch012>
- Schoen A and Laurens P (n.d.) Documentation of RISIS datasets: RISIS Patent. Available at: https://zenodo.org/record/3338122/files/Documentation_CIB_Final_RISIS1_version.pdf?download=1 (accessed 5 May 2025).
- Schoen A, Villard L, Laurens P, et al. (2012) The network structure of technological developments; Technological distance as a walk on the technology map. In: *Proceedings of the 17th International Conference on Science and Technology Indicators (STI 2012)*, Montréal, Canada, 5–8 September 2012, pp.733-742.
- Sickles R and Zelenyuk V (2019) *Measurement of Productivity and Efficiency: Theory and Practice*. Cambridge University Press. <https://doi.org/10.1017/9781139565981>
- Solow RM (1956) A contribution to the theory of economic growth. *The Quarterly Journal of Economics* 70(1): 65–94. <https://doi.org/10.2307/1884513>
- Teirlinck P and Khoshnevis P (2022) SME efficiency in transforming regional business research and innovation investments into innovative sales output.

Regional Studies 56(12): 2147–2163. <https://doi.org/10.1080/00343404.2022.2046263>

Tenca F (n.d.) Documentation of RISIS datasets: VICO. Available at: https://zenodo.org/record/3337991/files/Documentation_Vico_Final_edited.pdf?download=1 (accessed 17 April 2025).

Toma P, Frittelli M, Schoen A, et al. (2019) Analysing technological specificities of industrial sectors using corporate patent profiles with a gravity center modelling. In: *Conference: 17th international conference on scientometrics and informetrics*, Rome, Italy.

Appendix A

Table A1. Statistical data sources.

Indicator/data	Firm-size	Code in database	Database
Value added (VA)	Start-ups	AddedvaluethEUR	RISIS VICO
	Medium-sized	added_valueYEAR	RISIS CHEETAH
	Large R&D corporate performers	Value added calculated from financial data: Profit for period + Depreciation + Taxation + Interests paid + Cost of employees	Orbis
Number of employees (E)	Start-ups	Numberofemployees	RISIS VICO
	Medium-sized	num_empYEAR	RISIS CHEETAH
	Large R&D corporate performers	Number_employees	RISIS CIB/CinnoB
Fixed assets (FA)	Start-ups	FixedassetsthEUR	RISIS VICO
	Medium-sized	fix_assetYEAR	RISIS CHEETAH
	Large R&D corporate performers	Total fixes assets	Orbis
Patent data for RD)	–	–	RISIS Patent
Working-age population with tertiary education (EDU)	NUTS2	Table [edat_lfse_04]	Eurostat
Population density, persons per square kilometre (POP_DEN)	NUTS2	DEMO_R_D3DENS	Eurostat
Gross domestic product, purchasing power standard (PPS, EU27 from 2020), per inhabitant (GDP)	NUTS2	NAMA_I0R_2GDP	Eurostat
Total intramural expenditure on R&D, purchasing power standard (PPS) per inhabitant at constant 2005 prices (GERD)	NUTS2	RD_E_GERDREG	Eurostat
Location coefficient of employees for sectors: C20, C21, C26, C28, C29 (SPEC_LQ)	NUTS2 two-digit NACE code	SBS_R_NUTS06_R2	Eurostat
Inverted Hirschman–Herfindahl index, on the basis of each industry's own employment and the employment of the other two-digit NACE code manufacturing industries in a region (DIVERS_iHHI)	NUTS2 two-digit NACE code	SBS_R_NUTS06_R2	Eurostat

EDU: education; GDP: gross domestic product per capita; RD: relatedness density; RISIS: Research Infrastructure for Science and Innovation Policy Studies.

Appendix B

Table B1. Location of inventories of large R&D corporate performers in European NUTS2 regions (EU, Norway and Switzerland) in 2000–2015.

Economic sector (NACE two digits)	Number of inventors	Number of inventors with geographical information	Number of inventors outside EU continent	% Inventors outside EU continent	Number of inventors in Europe	% Inventors inside Europe	Number of inventors in EU and same NUTS2 as a company's location	% Inventors in EU same NUTS2
C20	201,679	199,353	31,978	16.04	167,375	83.96	65,272	32.74
C21	259,441	253,159	97,298	38.43	155,861	61.57	42,173	16.66
C26	398,438	390,729	141,370	36.18	249,359	63.82	72,121	18.46
C28	424,438	418,182	50,088	11.98	368,094	88.02	115,692	27.67
C29	284,812	281,766	61,461	21.81	220,305	78.19	75,786	26.90

Source: Authors' calculations, RISIS Patstat and CIB/CinnoB databases.

EU: European Union; RISIS: Research Infrastructure for Science and Innovation Policy Studies.

Appendix C

Steps of evaluation of relatedness density indicator

The technology space was evaluated first using co-occurrence analysis of technology subfields in the same patent documents, which yielded a network of technology-relatedness. Each node i ($i=1, \dots, n$) in the $n \times n$ network defined technology subfields. Ties between nodes indicated their degree of relatedness, revealing how often two technologies come together in one patent document. We used the least aggregated data of 401 technological subfields according to IFRIS classification (Schoen et al., 2012). The relatedness $\phi_{i,j,t}$ between each pair of technology i and j was evaluated with the minimum of the pairwise conditional probabilities of regions patenting in one technological subfield i , given that they patent in another technological subfield j during the same period. To avoid noise from negligible patenting activity, only technological subfields listed on patent documents in regions that had a revealed comparative advantage (RCA) were considered:

$$\phi_{i,j,t} = \min \left\{ \begin{array}{l} P(RCA_{x_{i,t}} | RCA_{x_{j,t}}), \\ P(RCA_{x_{j,t}} | RCA_{x_{i,t}}) \end{array} \right\} \quad (C1)$$

where a region r has an RCA in technology i in time t if the share of technology i in the region's technological portfolio is higher than the share of technology i in the entire patent portfolio of European regions. For a more formal explanation, $RCA_{r,t}(i) = 1$ if:

$$\frac{\text{patents}_{r,t}(i) / \sum_i \text{patents}_{r,t}(i)}{\sum_r \text{patents}_{r,t}(i) / \sum_r \sum_i \text{patents}_{r,t}(i)} > 1 \quad (C2)$$

We computed the relatedness between each pair of technologies i and j for a rolling sum of 5 years using patent data from 2004 to 2016. The relatedness density indicator was calculated for each technology subfield (i) and NUTS2 region (r), yielding a region technology-level variable that reveals the density of knowledge production around a given i technological subfield in region r at time t :

$$\begin{aligned} \text{RELATEDNESS_DENSITY}_{i,r,t} \\ = \frac{\sum_{j \in r, j \neq i} \phi_{ij}}{\sum_{j \neq i} \phi_{ij}} \end{aligned} \quad (C3)$$

where technological relatedness $\phi_{i,j,t}$ of technological subfield i to all other technologies j in which the region has a RCA is divided by the sum of

technological relatedness of technology i to all the other technologies j in the reference region at time t .

From a region technology-level variable, we then calculated a region industry-level variable. In this study, we needed a local knowledge indicator that was not for a particular technology, but rather for an industry, where a variety of technologies could play a role; this became the regional industry-level variable. In previous studies, the relatedness density indicator showed local knowledge base on a particular technology while constructing a city technology-level variable (Balland and Boschma, 2021; Boschma et al., 2015). The main question this study addressed was which technologies (and knowledge

used in developing them) were significant for each economic sector involved in the analysis, as there are no technological fields that uniquely characterise a specific economic sector (Toma et al., 2019). To avoid bias, we used data on the patenting activities of large corporate R&D performers, as from that data, it was possible to relate the technology subclass identified in the patent document with an economic sector. We obtained information on which technologies were patented and how often in each sector. With this information, we calculated weights by share of patents of each technology within an economic sector and evaluated a weighted relatedness density indicator.

Appendix D

Table D1. Descriptive statistics of start-ups firm-level data.

Economic sector	Variable	n	Mean	SD	Min	Max
Chemicals and chemical products	Value added per employee (Y)	131	6900	4861	535	26,716
	Fixed assets per employee (FA)	131	31,584	67,573	173	445,984
	Number of employees (E)	131	71	102	2	692
Basic pharmaceutical products and pharmaceutical preparations	Value added per employee (Y)	78	10,923	15,512	277	93,358
	Fixed assets per employee (FA)	78	41,465	101,469	178	683,878
	Number of employees (E)	78	203	483	1	2,320
Computer, electronic and optical products	Value added per employee (Y)	392	7243	6040	218	33,621
	Fixed assets per employee (FA)	392	11,203	13,698	37	94,132
	Number of employees (E)	392	186	505	1	3725
Machinery and equipment	Value added per employee (Y)	189	6983	4603	268	35,640
	Fixed assets per employee (FA)	189	17,156	41,072	47	251,482
	Number of employees (E)	189	110	165	1	1026
Motor vehicles, trailers and semi-trailers	Value added per employee (Y)	38	7755	5087	576	20,265
	Fixed assets per employee (FA)	38	20,786	53,566	329	333,276
	Number of employees (E)	38	42	35	1	112

Table D2. Descriptive statistics of fast-growing medium-sized firms' firm-level data.

Economic sector	Variable	<i>n</i>	Mean	SD	Min	Max
Chemicals and chemical products	Value added per employee (Y)	2175	8087	7486	160	66,734
	Fixed assets per employee (FA)	2175	21,576	62,746	2	2,163,217
	Number of employees (E)	2175	252	415	2	4682
Basic pharmaceutical products and pharmaceutical preparations	Value added per employee (Y)	969	9972	9045	190	70,732
	Fixed assets per employee (FA)	969	23,451	105,502	36	1,821,428
	Number of employees (E)	969	334	410	3	2922
Computer, electronic and optical products	Value added per employee (Y)	2258	5841	5803	178	43,971
	Fixed assets per employee (FA)	2258	6441	15,542	0	452,809
	Number of employees (E)	2258	428	1166	2	21,557
Machinery and equipment	Value added per employee (Y)	7582	5523	3926	287	28,951
	Fixed assets per employee (FA)	7582	6048	11,428	1	377,606
	Number of employees (E)	7582	267	505	2	9131
Motor vehicles, trailers and semi-trailers	Value added per employee (Y)	4104	4264	3505	211	30,491
	Fixed assets per employee (FA)	4104	5927	13,278	1	416,458
	Number of employees (E)	4104	526	883	5	9365

Table D3. Descriptive statistics of large R&D corporate performers' firm-level data.

Economic sector	Variable	<i>n</i>	Mean	SD	Min	Max
Chemicals and chemical products	Value added per employee (Y)	68	10,234	5371	3371	42,465
	Fixed assets per employee (FA)	68	21,039	14,516	3221	65,424
	Number of employees (E)	68	12,418	23,764	521	112,435
Basic pharmaceutical products and pharmaceutical preparations	Value added per employee (Y)	93	12,993	5649	1649	28,755
	Fixed assets per employee (FA)	93	28,797	25,021	2337	148,400
	Number of employees (E)	93	12,182	24,732	23	116,800
Computer, electronic and optical products	Value added per employee (Y)	177	9503	5314	1423	37,422
	Fixed assets per employee (FA)	177	10,970	8472	937	45,931
	Number of employees (E)	177	7565	16,149	75	116,281
Machinery and equipment n.e.c.	Value added per employee (Y)	196	8569	3600	3629	26,929
	Fixed assets per employee (FA)	196	12,156	10,621	1003	52,569
	Number of employees (E)	196	15,989	53,433	53	374,778
Motor vehicles, trailers and semi-trailers	Value added per employee (Y)	53	7677	4106	800	20,841
	Fixed assets per employee (FA)	53	15,521	18,247	2232	98,604
	Number of employees (E)	53	40,696	69,730	346	284,015

Table D4. Descriptive statistics of NUTS2 regional data.

Variable	<i>n</i>	Mean	SD	Min	Max
RD	11,130	27.0	12.5	0.0	75.2
SPEC_LQ	9625	1.0	1.3	0.0	36.5
DIVERSE_iHHI	9625	9.9	2.7	1.2	16.2
GDP	2005	25,284	10,116	5900	78,500
GERD	1851	383	408	6	3047
EDU	2369	28.2	9.6	8.3	74.7
POP_DEN	2336	460	1199	3	11,201

EDU: education; GDP: gross domestic product per capita; POP_DEN: population density; RD: relatedness density.

Table D5. Model parameters and efficiency estimates for large R&D corporate performers.

Model specification	Chemicals and chemical products		Pharmaceutical products and preparations		Computer, electronic and optical products		Machinery and equipment		Motor vehicles, trailers and semi-trailers	
	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value	Coefficient	p-Value
Constant	5.945	0.000	6.973	0.000	6.617	0.000	7.812	0.000	6.131	0.000
ln(FA)	0.311	0.000	0.154	0.001	0.376	0.000	0.192	0.000	0.517	0.000
ln(E)	0.024	0.338	0.066	0.010	-0.089	0.000	-0.064	0.000	-0.155	0.000
Time	0.013	0.587	0.070	0.006	0.009	0.669	-0.002	0.892	-0.035	0.168
Determinants of inefficiency										
Constant	5.853	0.038	-16.385	0.638	-23.754	0.482	-5.785	0.000	110.259	0.259
RD	0.018	0.000	0.047	0.273	-0.025	0.617	-0.014	0.000	-0.166	0.127
SPEC_LQ	-0.022	0.422	-0.148	0.247	0.166	0.695	-0.107	0.000	1.991	0.036
DIVERSE_iHI	0.039	0.308	-0.392	0.319	0.067	0.646	0.054	0.000	0.871	0.172
ln(GDP)	-0.781	0.016	2.325	0.563	2.450	0.478	0.857	0.000	-10.557	0.291
ln(GERD)	0.076	0.648	-0.523	0.606	-0.148	0.822	-0.516	0.000	-0.894	0.733
EDU	0.026	0.001	-0.080	0.320	-0.101	0.377	0.004	0.008	-0.122	0.491
POP_DEN	0.00002	0.487	-0.0002	0.506	-0.0003	0.752	0.00001	0.788	-0.0039	0.559
Mean efficiency	0.89		0.75		0.70		0.95		0.80	
Log likelihood value	-1.49		-36.12		-105.89		-27.08		-4.28	
Firms	24		33		73		69		19	
Observations	68		93		177		196		53	

The table reports stochastic frontier estimations for large R&D corporate performers. Given the multi-establishment structures of these firms, regional assignment based on headquarters location may not fully capture the regional conditions affecting firm efficiency. Therefore, the regional determinants of inefficiency for this group are not interpreted in the main analysis. The results are provided for descriptive comparison. Data for the large R&D corporate performers are taken. The CIB/CinnoB dataset comprised 3992 large corporations, representing about 80% of global private R&D expenditure (CIB/CinnoB dataset documentation, Schoen and Laurens, n.d.). Of this number, 1000 companies were based in the EU. Firms are classified as large R&D performers according to annual R&D investments. Financial data are based on a headquarters firm and its subsidiaries and derived from the Orbis database.

EDU: education; EU: European Union; GDP: gross domestic product per capita; POP_DEN: population density; RD: relatedness density.