

The Hidden Innovators: Uncovering Innovation Capabilities and Industrial Dynamics in European Rural Areas

Bernardo Caldarola, Tommaso Ciarli, Emanuele Pugliese, Simone Sasso, Šimon Trlifaj, Alessio Bumbea

Regional disparities, and particularly the innovation gap between urban and rural regions, remain a central challenge for both competitiveness and cohesion in Europe. While policy frameworks such as Smart Specialisation Strategies aim to foster place-based innovation, they often rely on conventional indicators, such as patents and R&D expenditure, that tend to capture innovation in urban contexts while undercounting forms of innovation typical of rural economies (Balland et al., 2020; Eder, 2019). These metrics privilege codified, formalised innovation outputs and risk overlooking organisational change, incremental innovation, and skill-based capability accumulation that is frequently present also in rural economic structures. This measurement bias may result in a systematic misrepresentation of the innovation capabilities present outside metropolitan areas and may lead to policy designs that inadvertently reinforce existing spatial inequalities. Moreover, the reliance on traditional innovation indicators implicitly assumes that innovation is primarily the result of formal R&D investments, thereby neglecting the role of production capabilities, occupational structures, and tacit knowledge embedded in local labour markets. In many rural regions, innovation is closely linked to specialised manufacturing niches, advanced agro-industrial activities, or the upgrading of traditional sectors through digitalisation and process innovation. These transformations may not immediately translate into patent applications but nonetheless reflect meaningful shifts in technological capabilities. A capability-based perspective is therefore essential to capture the underlying drivers of structural transformation across diverse territorial contexts.

In this paper, we address this gap by investigating the skill portfolios and capabilities that characterise European regions, and how these co-evolve with firm dynamics to shape regional development trajectories. Our conceptual framework builds on the idea that regional innovation is the outcome of coevolutionary dynamics between skills, entrepreneurship, firms' capabilities and broader regional characteristics (Pugliese et al., 2019; Coad et al., 2021). Rather than focusing exclusively on traditional innovation outputs, we analyse innovation potential through the structure of regional skill demand, entrepreneurial activity, and industrial dynamics. This approach enables us to identify “hidden innovators” and to better understand how different regions combine capabilities over time to support structural transformation and growth. In this framework, skills are not treated merely as labour inputs, but as carriers of productive knowledge that condition the set of feasible industrial activities in a region. Firms draw upon local skill endowments when diversifying into new technologies, while at the same time their hiring decisions

reshape the regional skill base. This bidirectional relationship may generate path-dependent development trajectories, where the relatedness and complementarity among skills influence the likelihood of successful diversification. By focusing on the structure and coherence of skill portfolios, rather than on their sheer volume, we seek to uncover the qualitative dimensions of regional capabilities that underpin innovation potential.

Our empirical strategy is grounded in the construction of a granular panel dataset for European regions at NUTS3 level, complemented—where possible—by localisation at Local Administrative Unit (LAU) level. The analysis integrates multiple data sources capturing firms, skills, technologies and regional enabling conditions. To measure skills and occupations, we leverage more than 200 million Online Job Vacancies (OJV) covering 27 European countries between 2019 and 2023, complemented by longitudinal data for a subset of 11 countries starting in 2014. These data provide detailed information on occupations and skills, allowing us to observe firms' skill demand with unprecedented spatial and sectoral granularity. Using name-entity recognition techniques, we match vacancies to firms in Orbis-IP and Dealroom, linking firm characteristics, startup activity and patenting behaviour to local skill demand.

For the regional context, we integrate indicators on demographics, infrastructure, and finance from Eurostat, ARDECO, and the JRC Rural Observatory. This multidimensional dataset allows us to characterise regional innovation capabilities beyond single metrics and to capture the interplay between skills, firms and local conditions.

To characterise regional innovation capabilities, we develop skill-based indicators inspired by the economic complexity literature (Neffke and Henning, 2013; Hidalgo and Hausmann, 2009; Tacchella et al., 2012; Aufiero et al., 2024). Using OJV data, we construct bipartite networks linking region-industries to skills and apply statistical validation methods to identify significant connections. From these validated networks, we compute measures of diversity, coherence and sophistication that capture different dimensions of local capabilities. Diversity reflects the breadth of skills demanded within a region-industry; coherence measures the relatedness and internal consistency of the skill portfolio; and sophistication captures the extent to which regions specialise in complex and less ubiquitous skills. These indicators allow us to move beyond counts of high-tech occupations and to assess the structural quality of regional skill endowments.

In addition, we identify occupations and skills associated with innovative activity using complementary approaches. First, at NUTS3 level, we regress regional patent counts on occupation shares derived from OJV data (Dotzel and Wojan, 2022). Second, at the firm level, leveraging our OJV–Orbis-IP matching, we analyse skill co-occurrence within patenting firms and measure the internal coherence and diversity of their skill demand profiles. Third, we use transformer-based natural language processing models (e.g. MPNet) to match emerging digital technologies to job-ad text, thereby detecting skill-

occupation profiles directly associated with frontier technological domains. These approaches allow us to map innovative occupations and innovation potential even in regions with limited patenting activity, providing an alternative lens on regional capability structures.

Building on this framework, we examine how these capability measures relate to firm entry, survival and broader industrial dynamics. We estimate econometric models linking regional skill coherence, diversity and sophistication to startup creation rates, and we employ hazard models to analyse firm survival and age dynamics. To address endogeneity concerns, we exploit variation in regional exposure to emerging digital technologies developed outside the European Union, constructing shift-share instruments that capture exogenous technological opportunities interacting with local skill endowments. We further explore heterogeneous effects across rural, intermediate and urban regions, as well as across sectors, to understand how capability structures interact with territorial and industrial composition.

Initial results on regional skill portfolios point to a central role for geographic proximity and local spillovers. Neighbouring regions display more similar skill structures, suggesting that knowledge diffusion and labour mobility contribute to spatial clustering of capabilities. At the same time, we observe marked spatial heterogeneity in the structure and quality of skill portfolios across Europe. Importantly, coherent skill portfolios are positively associated with startup creation differently across different territorial typologies. In contrast, fragmented skill structures are negatively associated with entrepreneurship across all territorial types. These findings highlight that not only the presence of advanced skills, but their configuration and internal complementarity, are crucial for regional industrial dynamics.

Taken together, the analysis provides a new geography of innovation potential in Europe. By leveraging web-based skills data and integrating them with firm-level and regional information, we uncover rural regions with substantial latent capabilities that remain invisible in traditional metrics. Beyond its empirical contributions, the paper advances the measurement of regional innovation by operationalising a capability-based framework at fine spatial scales. By integrating economic complexity methods, network analysis and natural language processing, the paper seeks to contribute to next generations of skills intelligence tools capable of capturing evolving technological and occupational structures. This approach provides a comparable analysis of regional capabilities across Europe and offers a richer understanding of how skills, firms and regional characteristics co-evolve.

From a policy perspective, our findings underscore the importance of moving beyond traditional innovation metrics. Identifying and strengthening coherent local capability structures may be more effective than focusing exclusively on patent counts. By revealing hidden innovation potential in rural areas, the paper contributes to the design of more

place-sensitive innovation and regional development policies, aligned with broader European objectives of cohesion, competitiveness and inclusive growth.

References

- Aufiero, S., De Marzo, G., Sbardella, A., & Zaccaria, A. (2024). Mapping job fitness and skill coherence into wages: an economic complexity analysis. *Scientific Reports*, 14(1), 11752. <https://doi.org/10.1038/s41598-024-61448-x>
- Balland, P.-A., Boschma, R., Crespo, J., & Rigby, D. L. (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>
- Coad, A., Segarra, A., & Teruel, M. (2016). Innovation and firm growth: Does firm age play a role? *Research Policy*, 45(2), 387–400. <https://doi.org/10.1016/j.respol.2015.10.015>
- Dotzel, K. R., & Wojan, T. R. (2022). *An occupational approach to analyzing regional invention* (NCSES Working Paper No. 22-202). National Center for Science and Engineering Statistics.
- Eder, J. (2019). Innovation in the periphery: A critical survey and research agenda. *International Regional Science Review*, 42(2), 119–146. <https://doi.org/10.1177/0160017618764279>
- Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26), 10570–10575. <https://doi.org/10.1073/pnas.0900943106>
- Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3), 297–316. <https://doi.org/10.1002/smj.2014>
- Pugliese, E., Cimini, G., Patelli, A., Zaccaria, A., Pietronero, L., & Gabrielli, A. (2019). Unfolding the innovation system for the development of countries: coevolution of science, technology and production. *Scientific Reports*, 9(1), 16440. <https://doi.org/10.1038/s41598-019-52767-5>
- Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., & Pietronero, L. (2012). A new metrics for countries' fitness and products' complexity. *Scientific Reports*, 2, 723. <https://doi.org/10.1038/srep00723>