Assessing the Impact of a European Union’s Policy on Agricultural Innovation in Italy

del Puente Francesco, University of Naples- Parthenope, francesco.delpuente@gmail.com

Sapio Alessandro University of Naples- Parthenope, alessandro.sapio@uniparthenope.it

Jin Yanhong, Rutger the State University of New Jersey (USA) Department of Agriculture Food and Resource Economics, yanhong.Jin@rutgers.edu

Pray Carl, Rutger the State University of New Jersey (USA) Department of Agriculture Food and Resource Economics, pray@sebs.rutgers.edu

Lightning Paper session prepared for presentation at the 2025 AAEA & WAEA Joint Annual Meeting in Denver, CO; July 27-29, 2025

Assessing the Impact of a European Union's Policy on Agricultural Innovation in Italy

Agricultural productivity in Italy grew rapidly until 2010 but has since declined (Fuglie et al., 2024). Recognizing the need for innovation to enhance agricultural productivity and promote sustainable agriculture, both European and Italian policymakers have emphasized fostering innovation as a cross-cutting objective (European Regulation 2021/2115 art.77; art.114, art.127). Innovation also plays a crucial role in addressing challenges such as climate change, biodiversity loss, soil degradation, geopolitical instability, energy challenges, and inflation (FAO, 2024). The European Commission's agricultural vision prioritizes innovation through multi-level stakeholder cooperation within an integrated governance framework (European Union; 2024).

The *European Innovation Partnership for Agricultural Productivity and Sustainability* (EIP-AGRI) is a key initiative to improve agricultural innovation through collaboration. Launched during the 2014-2020 Common Agricultural Policy (CAP) programming period, EIP-AGRI supports innovation and knowledge-sharing through Operational Groups (OGs). Each OG forms a partnership with farmers, food processors, producer cooperatives, consultants, universities, firms, and government entities to develop an innovative project (SCAR AKIS, 2019; Guerrero-Ocampo et al., 2022; Maziliauskas et al., 2018). Within EIP-AGRI, innovation is broadly defined to include both novel practices/products and the application of traditional practices/products to new geographical or environmental contexts (Article 126 of Regulation 2115/2021). del Puente et al., (2024) find that OGs in Italy integrate tradition in their innovation strategies, creating demand for traditional foods and expanding market opportunities. For example, *GELSO_net*[1], with a budget of € 784,244 and a consortium of six farmers, two freelance advisors, and a university, aims to create a new mulberry-based agri-food supply chain in the Piedmont region. This initiative not only expands mulberry sales but also increases their

[1] Gelso-Net project: https://www.innovarurale.it/it/pei-agri/gruppi-operativi/bancadati-go-pei/valutazione-di-innovative-strategie-di-adattamento. Last access on January 29, 2025.

economic value by converting pruning waste into biomass. The 2021-2027 CAP reaffirms EIP-AGRI as the EU's primary strategy for promoting agricultural innovation and knowledge transfer (European Commission, 2023).

Among EU-25 countries, Italy had the highest CAP expenditure during the 2014-2020 CAP period (€ 27 billion - EAFRD funds). As part of this investment, over €267 million was allocated to establish 787 OGs across 21 regions and autonomous provinces, with 183 in the North, 372 in the Centre, and 232 in the South. Despite its policy significance and over a decade of implementation, research on the impacts of EIP-AGRI in Italy remains limited, primarily due to data constraints (Proietti & Cristiano, 2023) and reliance on general information rather than on OG-specific data (Giarè & Vagnozzi, 2021). No quantitative analysis has evaluated OG's effectiveness in fostering innovation. Exiting assessments, mainly from Member State reports, highlight best practices and measure impacts based on the number of funded projects, in line with the EU Regulation 2021/2115 (Annex I p.134). Studies on Italian OGs primarily focus on operational aspects rather than innovation output. For example, Giarè & Vagnozzi (2021) examine financial procedures as a key factor in initiating innovation projects. Arzeni et al. (2023), use surveys and interviews to evaluate participant characteristics, partnership interactions, and satisfaction levels, highlighting strong engagement from farmers and research institutes. Molina et al. (2021) develop a qualitative framework identifying key factors driving participation, such as motivation, commitment, communication, trust, network learning, and knowledge-sharing and co-creation. del Puente et al. (2024) classify OGs' innovation goals using the Oslo Manual and the Community Innovation Survey (CIS), showing a strong focus on green innovations.

This study aims to address these gaps by a) constructing a dataset on OGs and quantifying innovation outputs using a Large Language Model (LLM) based on OG final reports and b) investigating factors contributing to OGs' innovation output using count data analysis.

Based on a newly constructed dataset of 646 OGs completed as of May 2024, developed using web-scrapping and LLM techniques, we discover significant disparities in innovation output across regions, thematic areas, commodities, partner composition, and leadership engagement. Specifically, OGs emphasizing market competitiveness, supply chain management, and resource management are more innovative than those focused on productivity improvement, despite receiving more intensive funding. Regarding commodity focus, OGs centred on forestry products are the most innovative, followed by those focused on industrial crops, vegetables and fruits, and other products (aquaculture, beekeeping, and floriculture). In contrast, OGs focused on main crops and viticulture exhibit the lowest levels of innovation, even though OGs in main crops receive the highest funding. OGs in the centre and northwest regions have lower innovation output than those in the northern region. Regional business diversity related to agriculture is positively associated with innovation output, but regional costs of production and labour are not statistically associated with innovation output. Leadership types and their prior funding experience are not found to affect innovation, whereas leaders actively collaborating with other OGs is significantly associated with increased innovation. Partner composition affects innovation output, with OGs involving farmers, research and education institutes, and training organizations being more innovative than those involving advisories, farmers' associations, and other private sector entities. These findings offer valuable policy insights to enhance the effectiveness of the EIP-AGRI initiative, particularly by refining funding allocation and fostering collaboration among key stakeholders.

This study makes the following significant contributions. First, it is the first to construct a dataset on OGs, including their characteristics and innovation output, using LLM, enabling a comprehensive economic analysis of policy impact on agricultural innovation in Italy. Second, our findings provide valuable insights for policymakers, assessing the effectiveness of the EIP-AGRI in fostering agricultural innovation and highlighting the role of partnership in shaping policy output.

The rest of this paper is organized as follows. The Background section presents a brief literature review on R&D collaboration and its effect on innovation, along with an overview of EIP-AGRI. The Methodology section explains the LLM model used for the extraction of the innovation output and focuses on the estimation strategy used for count-data analyses. The Data section describes the data sources. The Results section presents the findings, followed by an in-depth analysis in the Discussion section. Finally, the Conclusion section provides final remarks.

BACKGROUND

The role of R&D alliance in fostering innovation has been extensively explored in the literature (Becker, 2015; Pereira et al., 2023). R&D cooperation has become a cornerstone of innovation and a strategic priority for firms and farmers (Martínez-Noya & Narula, 2018). These collaborations facilitate knowledge exchange, technology transfer, increased efficiency and quality, expanded networks and markets, access to complementary resources, and accelerate time-to-market (Boiko, 2022).

Farmers, however, face significant challenges, including climate change, rural abandonment, the digital divide, resource scarcity, food quality issues, and growing global food demand (FAO, 2024). The literature on the agricultural sector explores R&D collaboration from a strategic perspective, emphasizing the importance of knowledge-sharing and the motivations behind farmers' participation (Klerkx et al., 2012). Aligned with this perspective, the EIP-AGRI initiative follows the broader Agricultural Knowledge and Innovation System (AKIS) approach, in which innovation emerges from interactions among farmers and other stakeholders (Touzard et al., 2015). Innovation is widely recognized as an interactive rather than a linear process, driven by dynamic exchanges among partners, with farmers playing a central role in advancing new solutions (Kok & Klerkx, 2023). These interactions facilitate developing and disseminating innovative practices (Knierim et al., 2015).

However, innovation is inherently uncertain, and collaboration entails significant costs, including externalities that make participation difficult for smaller, credit-constrained farms. Fieldsend et al.

(2021) highlight that the EIP-AGRI struggles to engage harder-to-reach groups, such as those with low literacy or digital skills who live in marginalized rural areas limiting its full potential. Policy support is, therefore, essential to facilitate stakeholder interactions (Hermans et al., 2023). Cuervo-Cazurra et al. (2018) emphasize that external financial support, such as governmental funding, can amplify the impact of R&D initiatives.

According to the European Commission, OGs must develop innovative projects addressing various agricultural challenges through diverse partnerships. For example, the VINSACLIMA project in Emilia Romagna (total budget: € 347,870) involved five research partners, two farmers, and three farmer associations. It conducted three on-field trials to improve viticulture techniques and three trials on processed wines to enhance Oenological techniques.[2] Similarly, the INPOSA project in Sicily (total budget: € 493,964) brought together five farmers, a university, and a farmer association to develop and commercialize a patented industrial invention for yellow tomatoes.[3]

Despite these efforts, farmers often remain locked into routine practices and traditional production methods due to limited openness to innovation (Bopp et al., 2019). They tend to collaborate with trusted, familiar partners, which may constrain opportunities for new pathways (Li et al., 2008). Moreover, they hesitate to adopt new practices when changes lack clear incentives or revenue gains (Tensi et al., 2022). As a result, when relying solely on internal knowledge, farmers are less likely to invest in R&D alliances (Cuervo-Cazurra & Annique Un, 2010).

The partner diversity encouraged by the EIP-AGRI strengthens R&D collaborations and offers broad advantages. Since the demonstrated success of their peers highly influences farmers (Wang et al., 2023), collaborating with other farmers and farmer associations could be a key driver of innovation.

[2] VINSACLIMA project: https://www.innovarurale.it/it/pei-agri/gruppi-operativi/bancadati-go-pei/filiera-agroalimentare-del-gelso-frutto-foraggio. Last access on January 1, 2025.
[3] INPOSA: https://www.innovarurale.it/it/pei-agri/gruppi-operativi/bancadati-go-pei/innovazione-nel-pomodoro-e-sostenibilita-agricoltura Last access on January 23, 2025.

Additionally, engaging in coopetition--collaborating with competitors--can be beneficial. Mariani & Belitski (2023) find that coopetition positively impacts firms that are not traditionally innovative, mainly rely on imitation, and have little experience in creating new-to-market innovations. Partnering with university researchers in OGs offers mutual benefits (D'Este & Patel, 2007). For farmers, collaboration facilitates the application of university-generated research (Siegel et al., 2003). For researchers, it provides exposure to real-world agricultural challenges and creates opportunities for applied research (Perkmann et al., 2021). Furthermore, collaboration with universities improves human capital (Audretsch et al., 2022), which fosters future innovation and partnerships.

Geographical proximity influences innovation collaboration (Audretsch & Belitski, 2024; Audretsch et al., 2023). Close physical proximity among partners fosters strong local networks, which help build trust, traditions, and routines (Balland et al., 2015). It also enables farmers to observe innovation output firsthand through on-farm demonstrations or structured meetings with partners (Ingram et al., 2018). Recognizing these benefits, the European Commission actively promotes both intra- and inter-regional R&D partnerships within the EIP-AGRI initiative.

The European Commission (2023) emphasizes the importance of national support in disseminating innovative solutions developed by OGs. Each member state publicly shares OG projects through an online database, ensuring visibility. In Italy, the *Innovarurale* database has been established, requiring OGs to submit standardized project reports (Article 126 of Regulation 2115/2021). However, a key limitation of this system is the lack of quantitative innovation metrics, as project reports are recorded solely as textual descriptions, preventing researchers from conducting rigorous quantitative assessments of OGs' innovation output.

METHODOLOGY

We sequentially employ web-scrapping and LLM techniques to construct a database of Italian OGs based on the publicly available *Innovarurale* dataset. Web-scraping enables us to collect detailed information on each OG, including project title, objectives, duration, budget, partner characteristics, involved commodities, milestone activities, innovation descriptions, and main results. A thorough data-cleaning process supplements missing information in *Innovarurale* using data from OGs' websites.

Building on recent advancements in using LLMs to build economic datasets (Devetak & Mandel, 2023), we employ the LlaMa_3 model developed by Meta to extract quantitative information on innovations and OGs. To ensure structured and contextually relevant outputs, we develop a structured system prompt following methodologies outlined by Chen et al. (2024) and Giray (2023). Thus, incorporating insights from agricultural innovation experts, we formulate the prompt to define innovation within the framework of Italian OGs, aligning with the Oslo Manual (IV edition) and the member state reports on OGs' best practices. Specifically, we distinguish between product and process innovations as in the Oslo Manual (IV edition): a new or improved good/service or business process, that differs significantly from the firm's previous and has been introduced in the market or (for process innovations) into the firm. Moreover, due to the context of the Italian agriculture and OGs, we define innovations relating traditional foods and processes. We classify a new approach to bringing traditional foods to markets as process innovation, while revaluing traditional foods falls under product innovation classification. Appendix B details the procedures, prompts, and parameters used to generate the data.

We implement a multi-stage annotation scheme to assess verifiability and detect factual inconsistencies in LLM-generated textual outputs based on established verification methods, as recommended by Wang et al. (2024). Following authors' approach, we assess quantitative outputs through multiple steps. First, we compare a sample of extracted results with findings from a preliminary study on OGs in Italy (del Puente et al., 2024). Second, we conduct randomized tests by submitting both incorrect and correct queries about specific projects and requesting justifications for responses. Finally, we run 13 rounds with the LLM, averaging the extracted values across iterations to construct the final dataset. As a robustness check, we also use the medium values across.

Since our outcome variable--the number of innovations delivered by each OG—is a count variable, we employ count data models, specifically Poisson and negative binomial (NB) regression, to examine the factors influencing OGs' innovation output. Formally, the number of innovations delivered by OG *i* is expressed as

$$E\left[I_i | OG_i, R_{it_0}, T_i, \tau_i\right] = \mu_i \tau_i = e^{\delta OG_i + \beta R_{it_0} + T_i + \varepsilon_i} \quad (1)$$

where I_i represents the number of innovative projects, OG_i includes OG and project characteristics, and R_{it_0} includes regional characteristics at base time t_0. Equation (1) also controls for fixed effects for OG establishment year (T_i). The error term ε_i is assumed to be clustered by region, allowing for correlation within but not between regions. The parameters δ's and β's are to be estimated.

If the unobserved heterogeneity term $\tau_i = e^{\varepsilon_i}$ is independent of regressors conditional on OG_i, R_{it_0}, and T_i, it follows a Poisson distribution with conditional mean and conditional variance $\mu_i \tau_i$. However, the Poisson assumption of equidispersion (equal mean and variance) is often violated due to unobserved heterogeneity or an excessive number of zero values in the dependent variable. In our case, overdispersion is present (approximately 20%), making Poisson estimates inefficient. However, excess zeros are not a major concern, as only 0.15% of OGs report no innovation outputs. To account for unobserved heterogeneity, we employ the NB model, which introduces τ_i as a gamma-distributed random term with mean one and variance α. The NB model assumes a conditional mean of $e^{\delta OG_i + \beta R_{it_0} + T_i}$ and a conditional variance of $\mu_i(1 + \alpha \tau_i)$, where α captures the degree of overdispersion.

To determine the appropriate model, we conduct Pearson's chi-squared tests following Poisson estimations. A non-significant result from Pearson's chi-squared test suggests that the Poisson model fits the data reasonably well, while a significant result indicates overdispersion, implying that NB model may be more suitable. After the NB estimations, we evaluate the statistical significance of the dispersion parameter α and conduct likelihood ratio tests. These tests compare the nested models (Poisson vs. NB), testing the null hypothesis that $\alpha = 0$. If the null hypothesis is rejected, the NB model is preferred, confirming the presence of overdispersion. Additionally, we evaluate in-sample prediction accuracy as an additional measure of model fit.

DATA

This study uses two data sources: a) innovation and OG characteristics, constructed using web scrapping and LLM techniques, and b) regional characteristics obtained from Eurostat. Our analysis focuses on 646 OGs established between 2016 and 2024 and completed their projects as of May 2024, representing approximately 84% of all Italian OGs.

OG characteristics include total budget, thematic areas (agricultural productivity, environmental and resources management, supply chain management, and market competitiveness), project duration, year of establishment, number of partners, and involved commodities. For OG members, we capture leader type and partner composition (farmers, research and education institutes, training entities, advisory entities, farmers' associations, other private entities, other public entities, and others). Additionally, we extract information on whether OG leaders currently collaborate with other OGs and have prior experience in securing public funding, Table A1 in Appendix A present the variables used in the estimation.

To address potential selection bias, we control for pre-OG establishment regional characteristics from 2016, including per capita income, geographical location, agricultural output, percentage of farmers engaged in agriculture-related business activities, production costs, and labour costs. Table AI in Appendix A presents the definitions of the key variables and data sources.

RESULTS

Characteristics of OGs and Innovation Delivered by OGs

We construct the OG data using web scraping and LLM techniques, which help us better understand Italian OGs. Table A2 in Appendix A presents summary statistics for key variables, including characteristics and innovation output of 646 OGs in Italy. We highlight some key features of OGs.

On average, each OG develops approximately three innovations. Figure 1 illustrates the kernel density distributions of the number of innovations, comparing two estimations: one based on the average and the other based on the median across 13 iterations using LLM. The highest density is observed around one to two innovations, indicating that most OGs produce only a few innovations, with fewer OGs achieving high counts of up to 10. Since the kernel distribution based on the mean and median are similar, the OG innovation output process may be less prone to extreme values.

More than half of OGs (55%) were established between 2019 and 2020, while 21% were established in 2016-2018, and 24% in 2021-2024. In our sample, geographically, OGs are primarily concentrated in the Northeastern regions (N = 320), followed by the South (231), Central regions (124), and North (95). Emilia-Romagna leads in the number of OGs, highlighting the role of territorial capital in affecting regional disparities between northern and southern Italy in both innovation and productivity (Castelnovo et al., 2020).

Figure (2) shows significant variations in OG funding across commodity types and thematic areas. In terms of sectoral focus, approximately 22.6% of OGs focus on vegetables and fruits, followed by muti-sector projects (20.1%),[4] livestock (19.4%), and viticulture (13.3%). The average budget per OG is highest for those focused on main crops, such as cereals and forage (€408,610), while the lowest average budget is for forestry (€235,805) (Figure 2(b)), while the total budget was highest for OGs focusing on vegetables and fruits and second lowest for those focused on forests (Figure 2(a)). Using Trienekens (2011)'s framework, we classify OGs into key thematic areas: productivity improvement

[4] According to Eurostat, multi-sector crops include a diverse range of commodities, including oilseeds, fiber crops, tobacco, hemp, hops, aromatic and culinary plants, medicinal plants, seeds for herbaceous oilseed plants, linseed seeds (including fiber flax), energy crops, and crops cultivated for renewable energy production.

(36.2%), supply chain management (18.6%), environmental sustainability (18.4%), market competitiveness (17.7%), and resource management (8.9%). The total budget per OG is highest for those focused on the productivity thematic area (€96 million), nearly double the budget for supply chain management, market competitiveness, and environmental sustainability, respectively (Figure 2(c)). The average budget per OG is higher for those focusing on market competitiveness and supply chain management than for those emphasizing productivity (Figure 2(d)).

Italian OGs involve diverse partners, with an average of eight types of participants per OG. Figure 3 shows that farmers are the most prevalent, participating in 91% of OGs, followed by research and educational institutes (88%), farmers' associations (38%), advisory consultants (34%), and public and private training institutes (26%). Oher private and public entities each account for 14%.[5] Figure 3 also shows that 35% of OGs are led by farmers, demonstrating a strong commitment from agricultural enterprises, while 29% are led by research and education institutes, reflecting a solid alignment with experimental and basic research. Additionally, 26% of OG leaders have prior experience securing funding from EIP-AGRI, and 22% are currently involved in other OGs as a partner.

Factors Affecting Innovation Outputs Delivered by OGs

We employ both Poisson and NB regression models to investigate the relationship between OG's innovation output and contributing factors. The outcome variable is the number of innovations completed by each OG. The total OG budget is used as the exposure variable accounting for factors that increase innovation output. Each model follows a stepwise approach, introducing additional controls at each specification. As shown in Table A3, Models (1) and (4) include only OG establishment year fixed effects and region fixed effects. Models (2) and (5) additionally control for regional characteristics in the base year, including per capita income, production costs per unit of

[5] "Other private entities" include environmental and consumer groups as well as other private organizations. "Other Public entities" refer to agencies and functional entities (e.g., environmental protection agencies), development agencies (e.g., agricultural districts), territorial authorities, and garden and park entities.

agricultural output, labour costs per unit of agricultural output, and the percentage of farmers engaged in agriculture-related business activities alongside their main agricultural operations. Models (3) and (6) incorporate additional controls for OG-specific characteristics, such as thematic areas, project duration, commodities involved, partner types, leader types, whether leaders have prior funding experience, and whether leaders participate as partners in other OGs. The estimation results for these six models are presented in Table A3 of Appendix A, with the corresponding marginal effects and IRRs (incidence-rate ratios) shown in Table A4. Marginal effects indicate the direction and magnitude of each variable's impact on OG innovation output, while IRRs show the relative changes in innovation output associated with each factor.

As shown in Table 1, incorporating regional and OG characteristics significantly improves in-sample prediction, increasing from 37.78% to 40% for the Poisson models and from 37.01% to 40.13% for NB models. Furthermore, the Pearson Chi^2 tests, the likelihood ratio tests, and the statistical significance of the dispersion parameter confirm the superior fit of the NB models, indicating that the Poisson model's assumption of equal mean and variance is violated due to overdispersion caused by heterogeneity. Additionally, the NB model demonstrates better predictive performance in the full-control specification than the Poisson model (40.13% vs. 40%).

The following section discusses the factors statistically associated with OG's innovation output based on the full-control NB model. Figure (4) illustrates the marginal effects and IRRs associated with thematic areas, partner types, and main commodities. Using the IRRs, we calculate the percentage change in innovation output as follows: (IRR-1) *100%. Figure 4(a) shows that OGs engaged in thematic areas such as environmental sustainability, market competitiveness, supply chains, and resource management have higher innovation counts than those focused on productivity improvement. Specifically, compared to productivity-focused OGs, those focusing on environmental sustainability experience a 38.6% increase in innovations, corresponding to one additional innovation at the mean. OGs emphasizing market competitiveness have a 36% increase (0.95 additional innovations), while those in supply chain management show a 39% increase (0.8 additional innovations). Similarly, OGs focused on resource management experience a 25% increase (0.65 additional innovations). These findings align with the CAP goals set by the European Commission (2024), which prioritizes environmental and market-related challenges in agriculture.

As shown in Table A4 of Appendix A, leader types and their prior funding experience are not statistically associated with innovation output, suggesting a strong and impartial funding process that operates independently of established partners. Additionally, leaders' current participation in other OGs as partners is statistically and positively associated with innovation output, highlighting the importance of collaboration between OGs. This contrasts with existing qualitative literature, which often focuses on the role of leaders in driving innovation within the same organization (Arzeni et al., 2023). Our findings, however, highlight the value of broader collaborative networks that extend beyond individual OG in fostering innovation.

The types of partners involved in OGs influence innovation output. As shown in Figure 4(b), farmer involvement is associated with the highest increase of 0.73 additional innovation counts (a 26% increase), aligning with the goal of OGs to develop innovative solutions for farmers. Participation from research and educational institutes, as well as other public sectors, is linked to higher innovation output, with increases of 16% and 19%, respectively. In contrast, involvement from advisories, farmers' associations, and other private sectors is associated with a 12-14% decrease in innovation output, though these effects are not statistically significant. These results suggest that public sector engagement is more likely to improve innovation of OGs. Additionally, partnerships with training organizations, whether private or public, are associated with a 17% increase in innovation output, corresponding to 0.48 additional innovation counts, which may reflect these organizations' strong connection to certification processes and their role in ensuring compliance with industry standards.

Figure 4(c) shows that compared to the base category (livestock), OGs focused on forestry have the highest increase in innovation output (aside of Others that involves several commodities), with a 67% increase, corresponding to additional 1.79 innovation counts at the mean. This is followed by OGs focused on industrial crops (a 42% increase), vegetable and fruits (a 24% increase), and others (e.g., aquaculture, beekeeping, and floriculture) with a twofold increase in innovation output. However, the lack of significant innovation output in OGs focusing on main crops and viticulture warrants further investigation. Understanding why these well-supported sectors do not exhibit higher innovation output could provide valuable insights into the effectiveness of funding allocation and the structural challenges within these industries. The prominence of forestry aligns with the European regulation (European Commission, 2021) that aims to foster the sector, especially in the northern regions of Italy, with a positive impact on sustainability.

At the regional level, as shown in Table A4 of Appendix A, regional production and labour costs are not statistically associated with innovation output, suggesting it is less likely that OG innovations focus on cost-saving. Furthermore, we find the percentage of farmers engaged in non-agricultural businesses related to agricultural operations, such as agritourism, restaurants, and recreation gardens, is associated with a 2% increase in innovation output. This suggests that regional business diversity related to agriculture is positively associated with agricultural innovation. Additionally, regional fixed effects show that, compared to those in the northern region, OGs in the centre and northwest regions have fewer innovation counts, with 44% and 34% decreases, respectively. The highest innovation output for OGs in the Northeast region highlights a potential limitation for policymakers who may rely solely on absolute project counts as indicators for policy impact within each region without considering the actual innovation output of OGs.

As a robustness check, we re-estimate the models using the median innovation output across 13 iterations with LLM and present the marginal effects and IRRs in Table A5 of Appendix A. The results are consistent with those estimated using the mean values for the main analysis.

DISCUSSION

One key finding is that OGs focusing on non-productivity thematical areas, especially environmental sustainability, tend to achieve higher innovation output. This aligns with the policy objectives of the European Commission (2024) and existing research on the role of R&D alliances for environmental innovations (De Marchi, 2012). Additionally, our findings reveal a divergence in regional funding priorities. Specifically, OGs with productivity-focused objectives receive the highest level of funding, despite being less innovation-intensive. This suggests that while regional policies align with national CAP objectives, funding allocation disproportionately favours the less innovative "Productivity" theme at the expense of more innovation-driven thematic areas. Future EIP-AGRI programs should provide more support to themes that yield higher innovation output.

From a policymaker's perspective, the relatively low significance of prior funding experience of OG leaders suggests that the funding process is robust and impartial, without systematically favouring or disfavouring established partners. At the same time, collaboration between OGs through shared partners boosts innovation and knowledge sharing, likely due to the strategic selection of collaboration partners based on access to specialized knowledge. These findings align with research on the meso-level influence of firms in fostering innovation. Giuliani (2007) argues that knowledge diffusion occurs within firm-selected clusters--facilitated by local proximity--to improve innovation. Specifically, through network analyses of the wine sector in Chile and Italy, Giuliani (2007) shows that innovation arises from a deliberate and highly selective search process rather than random or widespread knowledge diffusion, leading to the formation of localized knowledge communities.

Un et al. (2010) highlight the importance of accessing external knowledge as a driver of agricultural innovation. Our findings support this by demonstrating that innovation increases in regions where OGs leverage multiple sources of expertise—such as agrotourism, hospitality, and food processing--or partnering with diverse OGs, farmers can better identify sector-specific needs, ultimately enhancing innovation in agriculture.

Partnership diversity plays a crucial role in driving innovation (Yang et al., 2022), as they facilitate learning and accelerate innovation, particularly for firms that rely on imitation and lack prior experience in developing new-to-market products and processes (Audretsch & Belitski; 2024). Moreover, collaboration among farmers is further facilitated by cognitive proximity, which enhances the updating of business routines (Jiang et al., 2024). These dynamics create opportunities for farmers to expand their market reach and adopt technologies from competitors, contributing to overall growth (Mariani & Belitski, 2023).

The involvement of research and education institutes in OGs has statistically significant and substantial positive impact on innovation, highlighting the importance of strong networks that bridge research and practical applications. This supports the hypothesis of Ozdemir et al. (2023), who argue that collaborations with primary stakeholders, such as universities and research centres, improve innovativeness more than partnerships with other entities.

Conversely, our findings indicate that advisory entities (primarily individual consultants) and farmers' associations (such as commodity consortia and farmers' unions) negatively impact OGs' innovation output. This could be due to a 'lock-in' effect, where these partners reinforce existing practices rather than fostering innovation. Advisors in Italy primarily focus on implementing standards and regulations, which can limit OGs' capacity for innovation. This aligns with Allen & Sriram (2000), who highlight the negative relation between strict standards and innovation for low-tech sectors like agriculture. Similarly, Gorman (2019) argues that individual advisors, who often develop technical qualifications through an informal, hands-on approach, may not be as effective as at fostering innovation. These findings highlight the need for more structured advisory entities rather than individual consultants to better facilitate coordination and support farmers' innovation activities

(Lybaert et al., 2022). On the other hand, farmers' associations play a key role in negotiating labor standards and contract agreements within Italy's agricultural supply chain. Their involvement in OGs may constrain cost-saving labor, potentially hindering innovation output. While regional agricultural production and labor costs are associated with an increased innovation output, the results are not statistically significant. Nevertheless, as Hamilton et al. (2022) note, improving agricultural labour productivity requires the adoption of technologies that enhance workforce efficiency and maintain competitiveness, ultimately fostering greater innovation in agriculture.

CONCLUSIONS

This study is the first to construct a unique dataset on Italian OGs by using web-scraping and LLM techniques, based on final textual reports of 644 Italian OGs. This newly constructed database provides better understanding of OGs, investigate factors which facilitate innovation, and address the effectiveness of the EIP-AGRI initiative. On one hand, this study showcases the potential of LLMs as complementary tools for addressing data scarcity and bridging information gaps in agriculture. On the other hand, to further the exploration of policy impact, comprehensive, high-quality quantitative data collected from or documented by OGs are still needed.

OGs serve as valuable platforms for innovation and knowledge-sharing among diverse partners, encouraging farmers as act as local innovators, as highlighted by Cristiano & Proietti (2018). Our findings reaffirm the important role of public entities, especially research and educational institutions, in fostering innovation through collaboration, as they face fewer regulatory constraints compared to other entities. Additionally, we find that training entities play a significant role in driving innovation, aligning with recent literature (Takahashi et al., 2020). To maximize their impact, the EIP-AGRI initiative should encourage and incentivize the participation of these entities.

Our results highlight the limitations of measures in effectively monitoring innovation. Policymakers may face challenges when relying solely on absolute project counts as indicators of policy impact, as this approach does not necessarily capture the actual contribution of OGs to innovation. For example, consistent with Wesseler (2022) who find a negative relationship between sustainability challenges, the CAP, and agricultural productivity, we find that OGs focused on profitability improvement are heavily funded but show lower innovation compared to others. Similarly, while OGs focused on main crops attract the most funding, they exhibit lower innovation levels. In contrast, forest-focused OGs, despite receiving less funding, are more innovative than those focusing on other crops. The EIP-AGRI initiative should consider balancing funding allocation with innovation output across thematic areas and commodities to maximize the policy impact. Meanwhile, greater attention is needed on funding allocation across regions, as the southern and central regions still showing significantly lower levels of innovativeness compared to the Northeast. This suggests that the policy fails to reduce the Italian south-north divide. In this context, intra-regional collaborations could play a key role in improving effective knowledge-sharing among sector actors.

This study has several limitations. First, it does not rely on Retrieval-Augmented Generation (RAG), a technique that integrates information retrieval with text generation to improve AI responses using relevant external data sources. RAG requires constructing an internal knowledge library to enhance the model's performance. However, the literature on OGs is still quite limited and primarily focuses on best practices and partner commitments rather than innovation, making it challenging to create a sufficiently comprehensive library for defining innovation in the OG context. Future research could address this limitation by developing a more extensive data on OGs as well as other R&D networks, enabling the effective implementation of RAG to enhance AI-generated responses and comparing various LLM to assess their output. Second, while the number of innovations is an important indicator of project effectiveness, we are unable to estimate broader economic impact of these innovations or conduct a cost-benefit analysis due to data constraint. Future research will be needed once such data become available.

References

Allen, R. H., & Sriram, R. D. 2000. "The Role of Standards in Innovation" *Technological Forecasting and Social Change*, *64*(2), 171–181. https://doi.org/10.1016/S0040-1625(99)00104-3

Arzeni, A., Giarè, F., Lai, M., Lasorella, M. V., Ugati, R., & Vagnozzi, A. 2023. "Interactive Approach for Innovation: The Experience of the Italian EIP AGRI Operational Groups." *Sustainability*, *15*(19), 14271. https://doi.org/10.3390/su151914271

Audretsch, D. B., & Belitski, M. 2024. "Geography of knowledge collaboration and innovation in Schumpeterian firms." *Regional Studies*, *58*(4), 821–840. https://doi.org/10.1080/00343404.2023.2222137

Audretsch, D. B., Belitski, M., Caiazza, Rosa, & Phan, P. 2023. "Collaboration strategies and SME innovation performance." *Journal of Business Research*, *164*, 114018. https://doi.org/10.1016/j.jbusres.2023.114018

Audretsch, D. B., Belitski, M., Guerrero, M., & Siegel, D. S. 2022. "Assessing the Impact of the UK's Research Excellence Framework on the Relationship between University Scholarly Output and Education and Regional Economic Growth." *Academy of Management Learning & Education*, *21*(3), 394–421. https://doi.org/10.5465/amle.2021.0240

Balland, P.-A., Boschma, R., & Frenken, K. 2015. "Proximity and innovation: From statics to dynamics." *Regional studies*, *49*(6), 907–920.

Becker, B. 2015. "Public R&d Policies and Private R&d Investment: A Survey of the Empirical Evidence." *Journal of Economic Surveys*, *29*(5), 917–942. https://doi.org/10.1111/joes.12074

Boiko, K. 2022. "R&D activity and firm performance: Mapping the field." *Management Review Quarterly*, *72*(4), 1051–1087. https://doi.org/10.1007/s11301-021-00220-1

Bopp, C., Engler, A., Poortvliet, P. M., & Jara-Rojas, R. 2019 . "The role of farmers' intrinsic motivation in the effectiveness of policy incentives to promote sustainable agricultural practices." *Journal of Environmental Management*, *244*, 320–327. https://doi.org/10.1016/j.jenvman.2019.04.107

Castelnovo, P., Morretta, V., & Vecchi, M. 2020. "Regional disparities and industrial structure: Territorial capital and productivity in Italian firms." *Regional Studies*, *54*(12), 1709–1723. https://doi.org/10.1080/00343404.2020.1763941

Chen, B., Zhang, Z., Langrené, N., & Zhu, S. 2024. "Unleashing the potential of prompt engineering in Large Language Models: A comprehensive review" (No. arXiv:2310.14735). arXiv. https://doi.org/10.48550/arXiv.2310.14735

Cristiano, S., & Proietti, P. 2018. "Do EIP interactive innovation approaches interact each other?" *International Journal of Agricultural Extension*, *6*(3), Articolo 3.

Cuervo-Cazurra, A., & Annique Un, C. 2010. "Why some firms never invest in formal R&D. "*Strategic Management Journal*, *31*(7), 759–779. https://doi.org/10.1002/smj.836

Cuervo-Cazurra, A., Nieto, M. J., & Rodríguez, A. 2018. "The impact of R&D sources on new product development: Sources of funds and the diversity versus control of knowledge debate." *Long Range Planning*, *51*(5), 649–665. https://doi.org/10.1016/j.lrp.2017.06.004

De Marchi, V. 2012. "Environmental innovation and R&D cooperation: Empirical evidence from Spanish manufacturing firms." *Research Policy*, *41*(3), 614–623. https://doi.org/10.1016/j.respol.2011.10.002

del Puente, F., Buonomo, S., & Sapio, A. 2024. "Promoting agricultural innovation through EIP-AGRI Operational Groups: Insights from the Southern Italian projects." *Economia Marche - Journal of Applied Economics*, *43*(2). https://doi.org/10.57638/3034-8234EIPAGRI

D'Este, P., & Patel, P. 2007. "University–industry linkages in the UK: What are the factors underlying the variety of interactions with industry?" *Research Policy*, *36*(9), 1295–1313. https://doi.org/10.1016/j.respol.2007.05.002

Devetak, M., & Mandel, A. 2023. "Using Large Language Models to generate socio-economic datasets." *In Review*. https://doi.org/10.21203/rs.3.rs-3686756/v1

EU SCAR AKIS. 2019. "Preparing for future AKIS in Europe". *Brussels, European Commission*, *2*(0), 1.

European Commission. 2021. "Stakeholders' consultation and evidence base. New EU Forest strategy 2030." *Document 52021SC0652. Brussel* 16.07.2021

European Commission. 2023. "Evaluating the AKIS Strategic Approach in CAP Strategic Plans Guidelines*".*

European Commission. 2024. *"Study on outcomes achieved by EIP-AGRI Operational Group projects under the CAP."*

European Commission. 2024. "EU agricultural outlook, 2024-2035." *European Commission, DG Agriculture and Rural Development, Brussels.*

FAO. 2024. "World Food and Agriculture – Statistical Yearbook 2024." FAO. https://openknowledge.fao.org/handle/20.500.14283/cd2971en

Feiock, R. C., Lee, I. W., Park, H. J., & Lee, K.-H. 2010. "Collaboration Networks Among Local Elected Officials: Information, Commitment, and Risk Aversion." *Urban Affairs Review*, *46*(2), 241–262. https://doi.org/10.1177/1078087409360509

Fieldsend, A. F., Cronin, E., Varga, E., Biró, S., & Rogge, E. 2021. "'Sharing the space' in the agricultural knowledge and innovation system: Multi-actor innovation partnerships with farmers and foresters in Europe." *The Journal of Agricultural Education and Extension*, *27*(4), 423–442. https://doi.org/10.1080/1389224X.2021.1873156

Fuglie, K. O., Morgan, S., Jelliffe, J., & United States. Department of Agriculture. Economic Research Service,. 2024. "World agricultural production, resource use, and productivity, 1961-2020." *Economic Reserach Service,*. https://doi.org/10.32747/2024.8327789.ers

Giarè, F., & Vagnozzi, A. 2021. "Governance's effects on innovation processes: The experience of EIP AGRI's Operational Groups (OGs) in Italy." *Italian Review of Agricultural Economics (REA)*, *76*(3), Articolo 3. https://doi.org/10.36253/rea-13206

Giray, L. 2023. "Prompt Engineering with ChatGPT: A Guide for Academic Writers." *Annals of Biomedical Engineering*, *51*(12), 2629–2633. https://doi.org/10.1007/s10439-023-03272-4

Giuliani, E. 2007. "The selective nature of knowledge networks in clusters: Evidence from the wine industry." *Journal of Economic Geography*, *7*(2), 139–168. https://doi.org/10.1093/jeg/lbl014

Gorman, M. 2019. "Becoming an agricultural advisor – the rationale, the plan and the implementation of a model of reflective practice in extension higher education." *The Journal of Agricultural Education and Extension*, *25*(2), 179–191. https://doi.org/10.1080/1389224X.2018.1559742

Hamilton, S. F., Richards, T. J., Shafran, A. P., & Vasilaky, K. N. 2022. "Farm labor productivity and the impact of mechanization." *American Journal of Agricultural Economics*, *104*(4), 1435–1459. https://doi.org/10.1111/ajae.12273

Hermans, T. D. G., Smith, H. E., Whitfield, S., Sallu, S. M., Recha, J., Dougill, A. J., Thierfelder, C., Gama, M., Bunderson, W. T., Museka, R., Doggart, N., & Meshack, C. 2023. "Role of the interaction space in shaping innovation for sustainable agriculture: Empirical insights from African case studies." *Journal of Rural Studies*, *100*, 103012. https://doi.org/10.1016/j.jrurstud.2023.103012

Ingram, J., Chiswella, H., Mills, J., Debruyne, L., Cooreman, H., Koutsouris, A., Pappa, E., & Marchand, F. 2018. "Enabling learning in demonstration farms: A literature review." *International Journal of Agricultural Extension*, *6*(3), Articolo 3.

Jiang, H., Gai, J., Chen, C., & Yang, J. 2024. "Influence of standards alliance coopetition relationship on enterprises' technology innovation performance—A dual-path perspective of knowledge acquisition and routine updating." *Technology in Society*, *78*, 102683. https://doi.org/10.1016/j.techsoc.2024.102683

Knierim, A., Boenning, K., Caggiano, M., Cristóvão, A., Dirimanova, V., Koehnen, T., Labarthe, P., & Prager, K. 2015. "The AKIS concept and its relevance in selected EU member states." *Outlook on AGRICULTURE*, *44*(1), 29–36.

Kok, K. P. W., & Klerkx, L. 2023. "Addressing the politics of mission-oriented agricultural innovation systems." *Agricultural Systems*, *211*, 103747. https://doi.org/10.1016/j.agsy.2023.103747

Li, D., Eden, L., Hitt, M. A., & Ireland, R. D. 2008. "Friends, Acquaintances, or Strangers? Partner Selection in R&D Alliances." *Academy of Management Journal*, *51*(2), 315–334. https://doi.org/10.5465/amj.2008.31767271

Lybaert, C., Debruyne, L., Kyndt, E., & Marchand, F. 2022. "Competencies for Agricultural Advisors in Innovation Support." *Sustainability*, *14*(1), Articolo 1. https://doi.org/10.3390/su14010182

Mariani, M. M., & Belitski, M. 2023. "The effect of coopetition intensity on first mover advantage and imitation in innovation related coopetition: Empirical evidence from UK firms." *European Management Journal*, *41*(5), 779–791. https://doi.org/10.1016/j.emj.2022.05.001

Martínez-Noya, A., & Narula, R. 2018. "What More Can We Learn from R&D Alliances? A Review and Research Agenda." *Business Research Quarterly*, *21*(3), 195–212. https://doi.org/10.1016/j.brq.2018.04.001

Molina, N., Brunori, G., Favilli, E., Grando, S., & Proietti, P. 2021. "Farmers' Participation in Operational Groups to Foster Innovation in the Agricultural Sector: An Italian Case Study." *Sustainability*, *13*(10), Articolo 10. https://doi.org/10.3390/su13105605

Ozdemir, S., Carlos Fernandez de Arroyabe, J., Sena, V., & Gupta, S. 2023. "Stakeholder diversity and collaborative innovation: Integrating the resource-based view with stakeholder theory." *Journal of Business Research*, *164*, 113955. https://doi.org/10.1016/j.jbusres.2023.113955

Pereira, V., Bamel, U., Temouri, Y., Budhwar, P., & Del Giudice, M. 2023. "Mapping the evolution, current state of affairs and future research direction of managing cross-border knowledge for innovation." *International Business Review*, *32*(2), 101834. https://doi.org/10.1016/j.ibusrev.2021.101834

Perkmann, M., Salandra, R., Tartari, V., McKelvey, M., & Hughes, A. 2021. "Academic engagement: A review of the literature 2011-2019." *Research Policy*, *50*(1), 104114. https://doi.org/10.1016/j.respol.2020.104114

Proietti, P., & Cristiano, S. 2023. "Innovation support services: An evidence-based exploration of their strategic roles in the Italian AKIS." *The Journal of Agricultural Education and Extension*, *29*(3), 351–371. https://doi.org/10.1080/1389224X.2022.2069828

Robinson, G. M. 2024. "Global sustainable agriculture and land management systems." *Geography and Sustainability*, *5*(4), 637–646. https://doi.org/10.1016/j.geosus.2024.09.001

Siegel, D. S., Waldman, D. A., Atwater, L. E., & Link, A. N. 2003. "Commercial knowledge transfers from universities to firms: Improving the effectiveness of university–industry collaboration." *The Journal of High Technology Management Research*, *14*(1), 111–133. https://doi.org/10.1016/S1047-8310(03)00007-5

Sutherland, L.-A., Adamsone-Fiskovica, A., Elzen, B., Koutsouris, A., Laurent, C., Stræte, E. P., & Labarthe, P. 2023. "Advancing AKIS with assemblage thinking." *Journal of Rural Studies*, *97*, 57–69. https://doi.org/10.1016/j.jrurstud.2022.11.005

Takahashi, K., Muraoka, R., & Otsuka, K. 2020. "Technology adoption, impact, and extension in developing countries' agriculture: A review of the recent literature." *Agricultural Economics*, *51*(1), 31–45. https://doi.org/10.1111/agec.12539

Tensi, A. F., Ang, F., & van der Fels-Klerx, H. J. 2022. "Behavioural drivers and barriers for adopting microbial applications in arable farms: Evidence from the Netherlands and Germany." *Technological Forecasting and Social Change*, *182*, 121825. https://doi.org/10.1016/j.techfore.2022.121825

Touzard, J. M., Temple, L., Faure, G., & Triomphe, B. 2015. "Innovation systems and knowledge communities in the agriculture and agrifood sector: a literature review." *Journal of Innovation Economics & Management*, *17*(2), 117-142.

Trienekens, J. H. 2011. "Agricultural value chains in developing countries a framework for analysis." *International food and agribusiness management review*, *14*(2), 51-82. https://doi.org/10.22004/ag.econ.103987

Un, C. A., Cuervo-Cazurra, A., & Asakawa, K. 2010. "R&D Collaborations and Product Innovation." *Journal of Product Innovation Management*, *27*(5), 673–689. https://doi.org/10.1111/j.1540-5885.2010.00744.x

Wang, Y., Möhring, N., & Finger, R. 2023. "When my neighbors matter: Spillover effects in the adoption of large-scale pesticide-free wheat production." *Agricultural Economics*, *54*(2), 256–273. https://doi.org/10.1111/agec.12766

Wang, Y., Reddy, R. G., Mujahid, Z. M., Arora, A., Rubashevskii, A., Geng, J., Afzal, O. M., Pan, L., Borenstein, N., Pillai, A., Augenstein, I., Gurevych, I., & Nakov, P. 2024. "Factcheck-Bench: Fine-Grained Evaluation Benchmark for Automatic Fact-checkers" (No. arXiv:2311.09000). arXiv. https://doi.org/10.48550/arXiv.2311.09000

Wesseler, J. 2022. "The EU's farm-to-fork strategy: An assessment from the perspective of agricultural economics." *Applied Economic Perspectives and Policy*, *44*(4), 1826–1843. https://doi.org/10.1002/aepp.13239

Yang, N., WangDu, F., Zhang, Y., Du, J., & Sharif, S. M. F. 2022. "Impact of partner heterogeneity on innovation performance through interaction behaviour in Chinese complex product R&D industries." *International Journal of Technology Management*, *89*(3–4), 232–252. https://doi.org/10.1504/IJTM.2022.123740

Figure 1. Density Distributions of Innovation Output for Operation Groups based on the Mean or Median of Extracted Values Across 13 Iterations using the LLM Technique

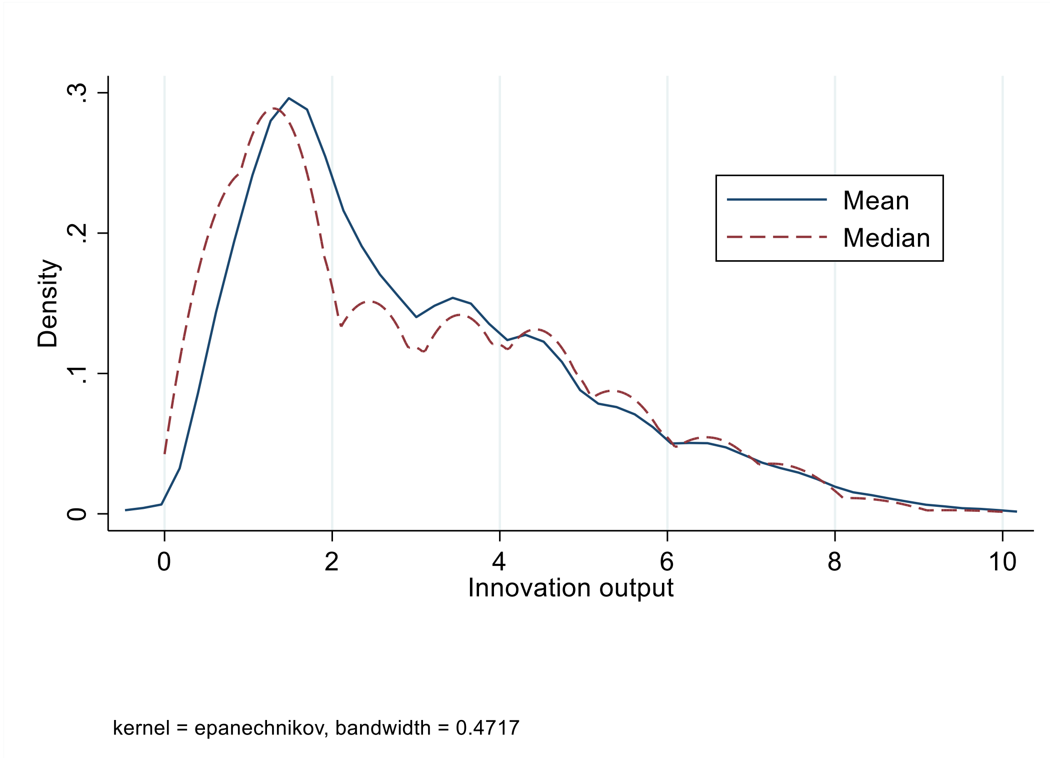
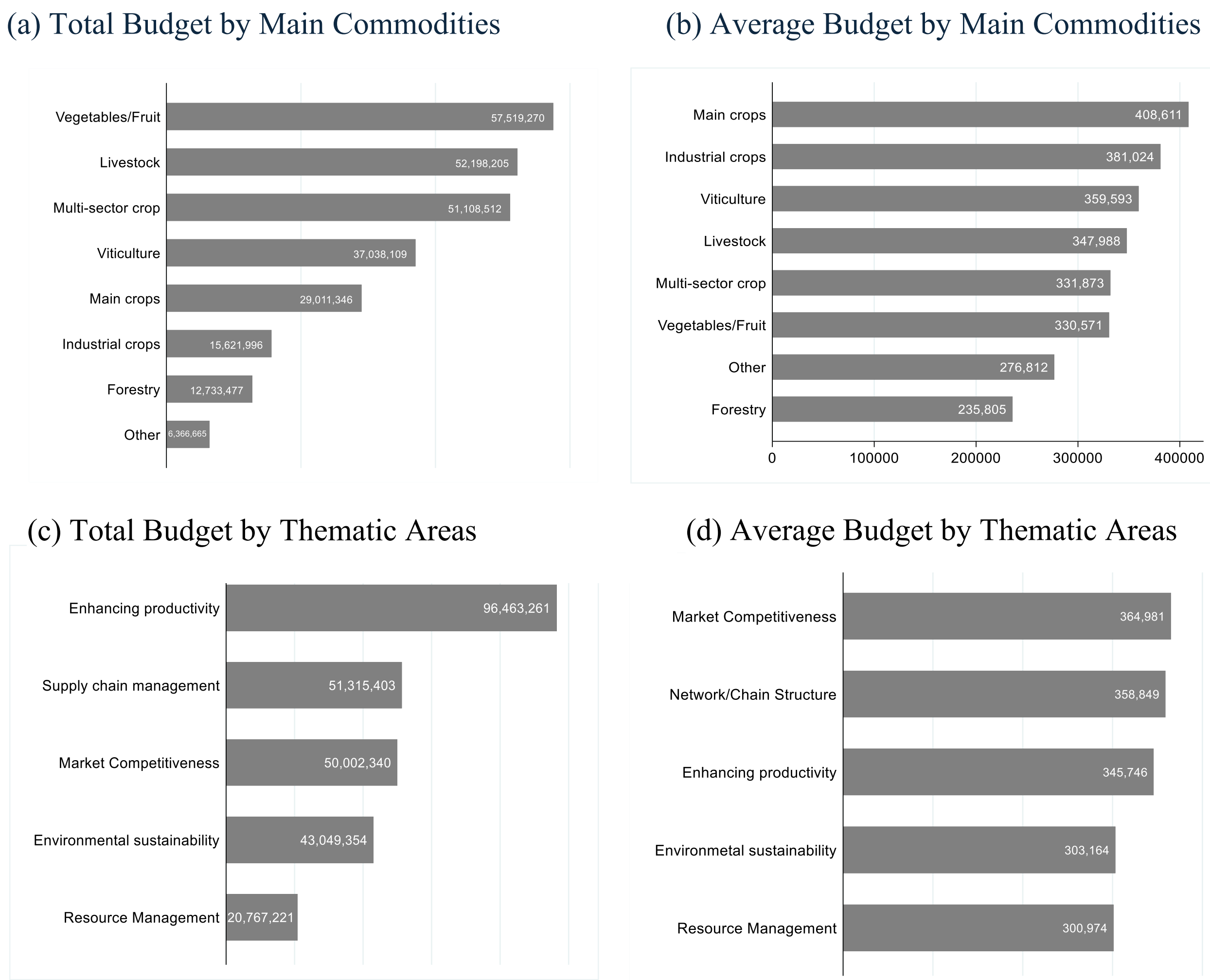
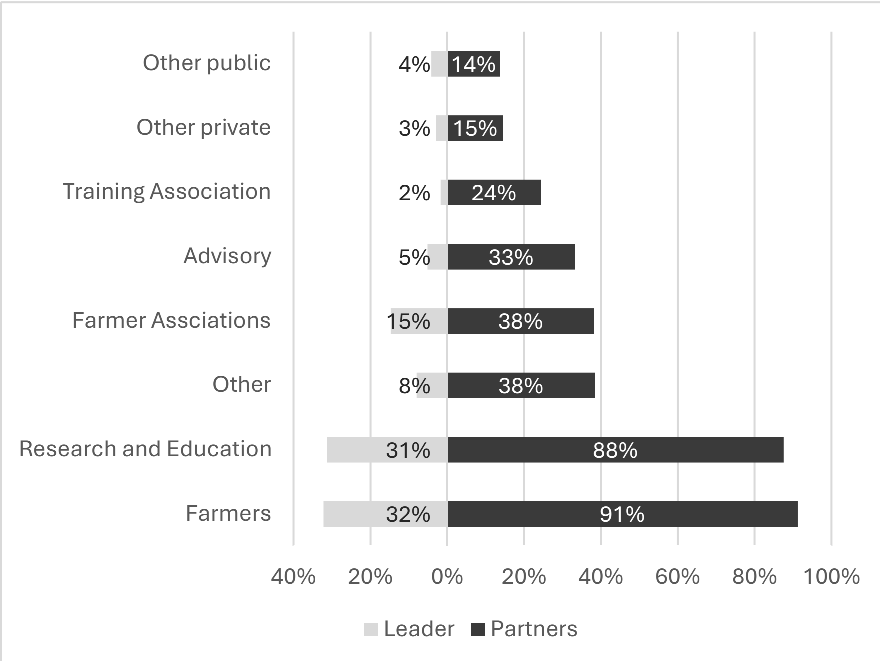


Figure 2. (a) Average Budget and (b) Total Budget per OG by Main Commodities Involved and per OG by Thematic Areas (in €)



Note: The “Other” category for main commodities includes beekeeping, floricultural, aquacultural products.

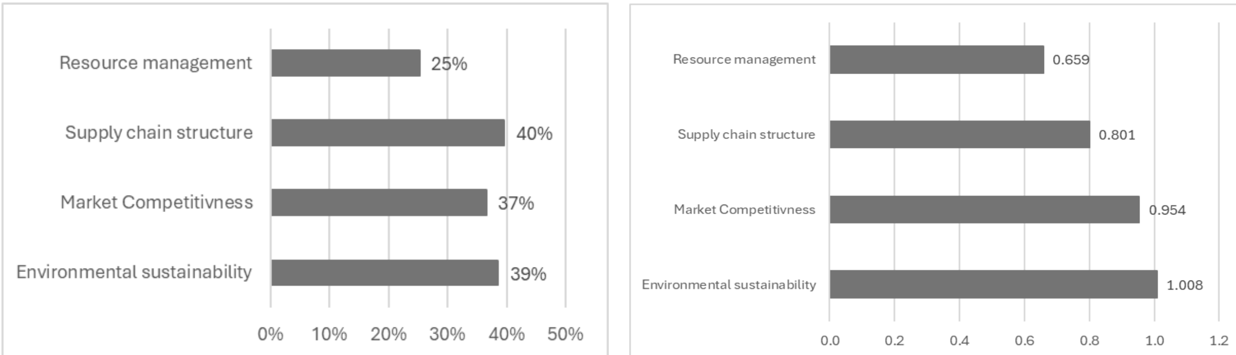
Figure 3. Distribution of Leader and Partner Types among Operation Groups



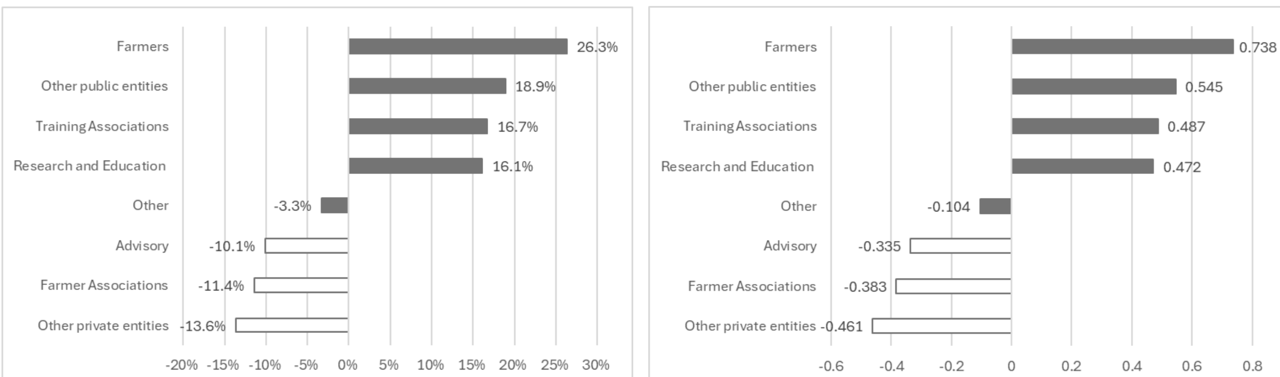
Note: The "Other" category involves OGs with no reference to Partner types

Figure 4. Marginal Effects at Mean and Percentage Change derived from IRR Estimates, Associated with Thematic Areas, Partner Types, and Main Commodities

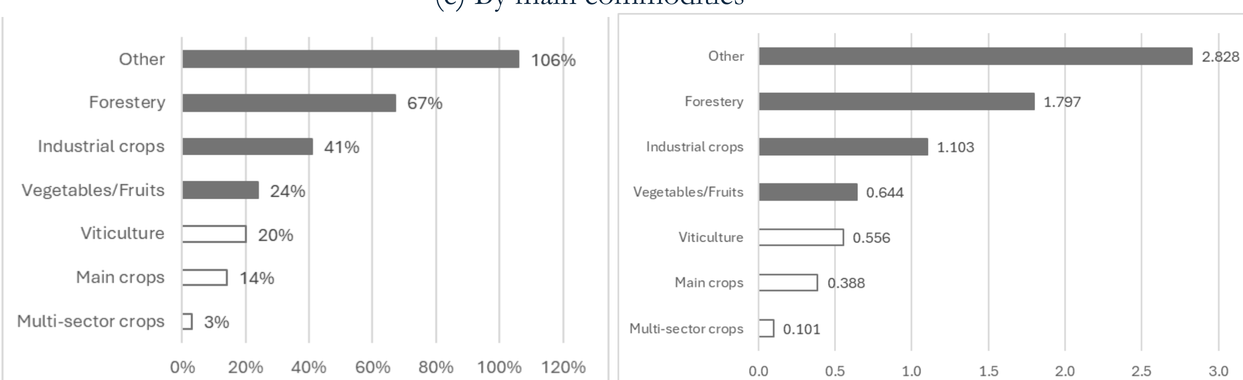
(a) By thematic areas



(b) By partner types



(c) By main commodities



Notes: The solid bars indicate statistical significance at least 10% significance level, while hollow bars indicate statistical insignificance. Percentage changes are calculated based on the following formular: (1-IRR)*100. The base category is “Enhancing productivity” for thematic areas and “livestock” for main commodities. The “Other” category for main commodities includes beekeeping, floricultural, aquacultural products.

Table 1. Model Fit across Different Models (Poisson vs. Negative Binomial) and Specifications

Model	Poisson			Negative Binomial		
	(1)	(2)	(3)	(4)	(5)	(6)
Establishment year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional characteristics		Yes	Yes		Yes	Yes
OGs' characteristics			Yes			Yes
Pearson's Chi^2 test	1172.8***	1071***	849.63***			
Ln (Alpha)				0.168***	0.145***	0.066***
				(0.054)	(0.029)	(0.018)
Likelihood ratio test				69.11***	52.58***	13.11***
% correct predictions	37.78%	37.79%	40%	37.01%	37.66%	40.13%

Note: We incorporate the following regional characteristics: per capita income, production costs and labour cost per unit of agricultural output, and the percentage of farmers engaged in agricultural-related business activities alongside their agricultural operations in the base year 2016 before OGs were established. OGs' characteristics are thematic areas, project duration, commodities involved, partner types, leader types, whether leaders have prior funding experience and participate as partners in other OGs.
*** 1%, ** 5%, and * 10%

Appendices

For

Assessing the Impact of a European Union's Policy on Agricultural Innovation in Italy

Appendix A: Additional Figures and Tables

Table A1 Variables used in the regression model, definitions and data sources

Variable	Description	Source
(y) Innovations	Number of actual innovation	Extracted with LLM
Total Budget	In Millions €	web-scraped 'Innovarurale'
Thematic	Orientation of the main OG objective (Environmental; Market Competitiveness; Network/Chain structure; Resources; Productivity)	web-scraped 'Innovarurale'
Duration	Number of months	web-scraped 'Innovarurale'
Partners	Number of partners involved	web-scraped 'Innovarurale'
Collaboration	Yes/No head collaboration with other OGs	web-scraped 'Innovarurale'
Past experiences	Yes/No head past experiences in achieving public funding	web-scraped 'Innovarurale'
Leader type	Head Partner typology (e.g., Farmer Research&Education, Training; Advisory; Association of farmers; Other private; other public; other entities)	web-scraped 'Innovarurale'
Actors involves	Typologies of actor involved in the project (dummy for each)	web-scraped 'Innovarurale'
Established Year	number of the first year per OG	web-scraped 'Innovarurale'
Income per capita	in Millions € per region	Eurostat
Geographical location	categorical (Northeast; Northwest; Centre; South)	Eurostat
Commodities	Typologies of commodities (Livestock; Main crops; industrial crops; Veg&Fruits; Viticulture; Multi-sector; Forestry; Others)	web-scraped 'Innovarurale'
Labour costs	Ratio number of labour in agriculture on agricultural costs per Region (in Millions €)	Eurostat
Percentage farmer Other activities	percentage of number of non-agricultural activities Managed by farmers on total number of farmers Per region in 2016	Eurostat
Production costs	ratio of total costs per agr output per region	Eurostat

Table A2 Summary Statistics of Key Variables

Variable	Mean	Std. Dev.	Min	Max
Continuous variables				
Number actual Innovations	3.03	1.91	0	9.7
Established year	2018.9	1.37	2016	2022
Per capita income by region	30144	7267	17500	45400
Production costs per unit of agricultural output value	.17	.04	.042	.288
Labour costs per unit of agricultural output value	.56	.38	.179	1.421
Percentage of farmers engaged in agricultural-related business activities	7.87	8.36	1.251	41.894
Duration (month)	31.18	7.69	11	59
Number of partners	8.30	4.36	1	48
Dummy variables				
OG leader collaborating in other OGs	.24	.43	0	1
OG leader past OG experience	.28	.45	0	1
Partner Type:				
Farmers	.91	.28	0	1
Training entities	.26	.44	0	1
Advisory	.34	.47	0	1
Research and educational institutes	.87	.32	0	1
Farmers' associations	.38	.48	0	1
Other public entities	.14	.35	0	1
Other private entities	.14	.35	0	1
Others	.40	.49	0	1
Categorical variables				
Region:				
Mezzogiorno	.254	.436	0	1
Centre	.167	.373	0	1
North-East	.463	.499	0	1
North-West	.116	.321	0	1
Thematic Areas:				
Environmental	.184	.388	0	1
Market Competitiveness	.181	.385	0	1
Network/Chain Structure	.189	.392	0	1
Productivity	.35	.477	0	1
Resources	.096	.295	0	1
Commodities:				
Livestock	.192	.394	0	1
Vegetables/Fruit	.212	.409	0	1
Main crops	.101	.301	0	1
Industrial crops	.054	.227	0	1
Viticulture	.135	.342	0	1
Multi-sector crop	.201	.401	0	1
Forestry	.077	.267	0	1
Other	.028	.165	0	1
Leader Typology:				

Farms	.322	.468	0	1
Training entities	.017	.129	0	1
Advisory entities	.051	.22	0	1
Research and education	.313	.464	0	1
Associations of farmers	.147	.354	0	1
Other private entities	.029	.169	0	1
Other public entities	.04	.197	0	1
Other	.08	.272	0	1

Table A3 Estimation analysis Poisson and Negative Binomial, fixed effects and control variables

	Poisson Models			Negative Binomial Models		
	(1)	(2)	(3)	(4)	(5)	(6)
Established: Base 2016						
2017	1.516**	1.402	1.421*	1.644***	1.511**	1.447**
	(0.310)	(0.323)	(0.276)	(0.256)	(0.315)	(0.266)
2018	0.791	0.756	0.804	0.882	0.815	0.818
	(0.237)	(0.163)	(0.146)	(0.255)	(0.187)	(0.147)
2019	0.994	1.039	0.939	0.993	1.060	0.941
	(0.219)	(0.169)	(0.182)	(0.226)	(0.198)	(0.188)
2020	1.107	1.196**	0.986	1.136	1.209**	0.982
	(0.146)	(0.107)	(0.112)	(0.147)	(0.110)	(0.116)
2021	1.579***	1.499*	1.139	1.614***	1.557**	1.150
	(0.145)	(0.319)	(0.155)	(0.182)	(0.313)	(0.165)
2022	3.747***	4.026***	2.796***	3.787***	4.017***	2.760***
	(0.607)	(0.845)	(0.775)	(0.616)	(0.870)	(0.762)
Region: Base Northeast						
South	0.752	1.180	0.876	0.730*	1.003	0.850
	(0.140)	(0.596)	(0.258)	(0.133)	(0.584)	(0.272)
Centre	0.721	0.750	0.673**	0.708	0.710	0.663**
	(0.174)	(0.212)	(0.127)	(0.172)	(0.215)	(0.127)
North-West	0.674	0.425***	0.547***	0.763	0.438***	0.561***
	(0.215)	(0.106)	(0.108)	(0.314)	(0.122)	(0.117)
Regional income per capita		1.000	1.000		1.000	1.000
		(0.000)	(0.000)		(0.000)	(0.000)
Production costs		2.951	1.954		2.175	1.896
		(5.152)	(2.401)		(4.018)	(2.371)
Labour costs		0.676	1.097		0.682	1.102
		(0.236)	(0.206)		(0.257)	(0.215)
Ratio of farmer engaged in non-agricultural activities		1.032*	1.022**		1.033**	1.022**
		(0.018)	(0.010)		(0.015)	(0.010)
Thematic: Base Productivity						
Environmental			1.384***			1.386***
			(0.109)			(0.111)
Market Competitiveness			1.358***			1.365***
			(0.055)			(0.057)
Network/Chain Structure			1.301***			1.307***
			(0.105)			(0.106)
Resources			1.270***			1.252***
			(0.090)			(0.091)
Duration (in months)			0.986**			0.987**
			(0.006)			(0.006)
Leader collaboration in other OGs			1.166***			1.171***
			(0.061)			(0.058)
Leader Past Experience in OG projects			1.005			1.000
			(0.107)			(0.104)
Base Leader: Research & Education						

LEADER: Farms			0.996			1.006
			(0.065)			(0.065)
LEADER: Training entities			0.918			0.922
			(0.243)			(0.253)
LEADER: Advisory entities			1.082			1.073
			(0.095)			(0.096)
LEADER: Associations of farmers			0.892			0.895
			(0.069)			(0.071)
LEADER: Other private entities			1.154			1.163
			(0.251)			(0.256)
LEADER: Other public entities			1.085			1.086
			(0.137)			(0.137)
LEADER: Other firms			0.976			0.988
			(0.107)			(0.107)
Number of Partners			1.006			1.005
			(0.009)			(0.008)
Typology partner involved:						
Farms			1.255***			1.263***
			(0.089)			(0.092)
Training			1.165			1.167*
			(0.110)			(0.109)
Advisory			0.900			0.899
			(0.060)			(0.059)
Research&Education			1.159**			1.161**
			(0.076)			(0.075)
Farmers'Associations			0.883			0.886
			(0.071)			(0.071)
Other public			1.182**			1.189**
			(0.089)			(0.094)
Other private			0.866**			0.864**
			(0.058)			(0.059)
Other firms			0.973			0.967
Base Commodity: Livestock			(0.062)			(0.065)
Veg&Fruit			1.239**			1.242**
			(0.117)			(0.118)
Main crops			1.146			1.146
			(0.111)			(0.111)
Industrial crops			1.411**			1.415**
			(0.224)			(0.217)
Viticulture			1.212			1.209
			(0.162)			(0.162)
Multi-sector			1.033			1.038
			(0.085)			(0.082)
Forestry			1.662***			1.676***
			(0.248)			(0.249)
Other			1.998***			2.064***
			(0.501)			(0.553)
Constant	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
lnalpha				0.168***	0.145***	0.066***
				(0.054)	(0.029)	(0.018)

Observations	646	646	646	646	646	646
% Y predicted count	37.78%	37.79%	40%	37.01%	37.66%	40.13%

Robust SE in parentheses *** p<0.01, ** p<0.05, * p<0.1 Exposure Total Budget per OGs

Table A4 Estimation analysis Poisson and Negative Binomial, IRR and marginal effects

	Poisson Models		Negative Binomial Models	
	IRR	Marginal Effects	IRR	Marginal Effects
	(3)	(3)	(6)	(6)
Average Number of actual innovations per OG				
Base established 2016				
2017	1.421*	1.280*	1.447**	1.376**
	(0.276)	(0.735)	(0.266)	(.701)
2018	0.804	-0.595	0.818	-0.559
	(0.146)	(0.489)	(0.147)	(0.495)
2019	0.939	-0.185	0.941	0.180
	(0.182)	(0.571)	(0.188)	(0.595)
2020	0.986	-0.043	0.982	0.054
	(0.112)	(0.345)	(0.116)	(0.362)
2021	1.139	0.423	1.150	0.461
	(0.155)	(0.447)	(0.165)	(0.479)
2022	2.796***	5.459***	2.760***	5.415***
	(0.775)	(1.878)	(0.762)	(1.858)
Base NorthEast				
South	0.876	-0.453	0.850	-0.560
	(0.258)	(0.993)	(0.272)	(1.083)
Centre	0.673**	-1.192**	0.663**	-1.257**
	(0.127)	(0.607)	(0.127)	(0.642)
North-West	0.547***	-1.649***	0.561***	-1.637***
	(0.108)	(0.583)	(0.117)	(0.639)
Income per region	1.000	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)
Costs on output	1.954	2.036	1.896	1.973
	(2.401)	(3.740)	(2.371)	(3.860)
Labour on output	1.097	0.280	1.102	0.299
	(0.206)	(0.568)	(0.215)	(0.599)
Ratio farm other activities	1.022**	0.067**	1.022**	0.066**
	(0.010)	(0.031)	(0.010)	(0.030)
Base Productivity				
Environmental	1.384***	0.961***	1.386***	0.979***
	(0.109)	(0.288)	(0.111)	(0.296)
Market Competitiveness	1.358***	0.895***	1.365***	0.927***
	(0.055)	(0.124)	(0.057)	(0.132)
Network/Chain Structure	1.301***	0.754***	1.307***	0.779***
	(0.105)	(0.232)	(0.106)	(0.239)
Resources	1.270***	0.675***	1.252***	0.640***
	(0.090)	(0.216)	(0.091)	(0.223)
Duration (in months)	0.986**	-0.041**	0.987**	-0.041***
	(0.006)	(0.016)	(0.006)	(0.017)
Leader collaboration in other OGs	1.166***	0.466***	1.171***	0.487***
	(0.061)	(0.158)	(0.058)	(0.153)
Leader Past Experience in OG projects	1.005	0.016	1.000	-0.009
	(0.107)	(0.322)	(0.104)	(0.319)
Base Leader: Research&Education				
LEADER: Farms	0.996	-0.012	1.006	0.017
	(0.065)	(0.198)	(0.065)	(0.201)
LEADER: Training entities	0.918	-0.250	0.922	-0.241

	(0.243)	(0.751)	(0.253)	(0.788)
LEADER: Advisory entities	1.082	0.252	1.073	0.227
	(0.095)	(0.284)	(0.096)	(0.292)
LEADER: Associations of farmers	0.892	-0.331	0.895	-0.325
	(0.069)	(0.210)	(0.071)	(0.218)
LEADER: Other private entities	1.154	0.473	1.163	0.504
	(0.251)	(0.768)	(0.256)	(0.792)
LEADER: Other public entities	1.085	0.261	1.086	0.266
	(0.137)	(0.426)	(0.137)	(0.428)
LEADER: Other firms	0.976	-0.073	0.988	-0.038
	(0.107)	(0.331)	(0.107)	(0.332)
Number Partners	1.006	0.018	1.005	0.015
	(0.009)	(0.026)	(0.008)	(0.026)
Typology of partners:				
Farms	1.255***	0.690***	1.263***	0.721***
	(0.089)	(0.220)	(0.092)	(0.229)
Training	1.165	0.464*	1.167*	0.475*
	(0.110)	(0.270)	(0.109)	(0.271)
Advisory	0.900	-0.321	0.899	-0.327
	(0.060)	(0.204)	(0.059)	(0.151)
Research&Education	1.159**	0.448**	1.161**	0.461**
	(0.076)	(0.197)	(0.075)	(0.195)
Farmers'Associations	0.883	-0.377	0.886	-0.374
	(0.071)	(0.253)	(0.071)	(0.255)
Other public	1.182**	0.507**	1.189**	0.533**
	(0.089)	(0.225)	(0.094)	(0.240)
Other private	0.866**	-0.436**	0.864**	-0.451**
	(0.058)	(0.215)	(0.059)	(0.222)
Other firms	0.973	-0.083	0.967	-0.102
	(0.062)	(0.194)	(0.065)	(0.207)
Base commodity: Livestock				
Veg&Fruit	1.239**	0.616**	1.242**	0.632**
	(0.117)	(0.271)	(0.118)	(0.276)
Main crops	1.146	0.376	1.146	0.381
	(0.111)	(0.270)	(0.111)	(0.273)
Industrial crops	1.411**	1.060*	1.415**	1.082**
	(0.224)	(0.559)	(0.217)	(0.547)
Viticulture	1.212	0.547	1.209	0.546
	(0.162)	(0.400)	(0.162)	(0.407)
Multi-sector	1.033	0.086	1.038	0.098
	(0.085)	(0.214)	(0.082)	(0.209)
Forestry	1.662***	1.705***	1.676***	1.763***
	(0.248)	(0.598)	(0.249)	(0.610)
Other	1.998***	2.571**	2.064***	2.776*
	(0.501)	(1.272)	(0.553)	(1.422)
Constant	0.000***		0.000***	
	(0.000)		(0.000)	
lnalpha			0.066***	
			(0.018)	
Observations	646	646	646	646

*Robust SE in parentheses *** p<0.01, ** p<0.05, * p<0.1 Exposure Total Budget per OGs*

Table A5 Estimation analysis for median distribution of OG actual innovations, Poisson and Negative Binomial, IRR and marginal effects

	Poisson Models		Negative Binomial Models	
	IRR	Marginal Effects	IRR	Marginal effects
	(7)	(7)	(8)	(8)
Median Number of actual Innovations per OG				
Base established: 2016				
2017	1.370	2.801	1.397*	3.077*
	(0.284)	(1.922)	(0.268)	(1.972)
2018	0.873	0.702	0.911	0.777
	(0.178)	(0.375)	(0.181)	(0.419)
2019	0.991	0.975	0.995	0.987
	(0.213)	(0.580)	(0.223)	(0.623)
2020	1.105	1.339	1.104	1.343
	(0.155)	(0.536)	(0.159)	(0.562)
2021	1.268*	2.110*	1.289*	2.268
	(0.175)	(0.936)	(0.199)	(1.143)
2022	3.164***	414.750***	3.141***	428.344***
	(1.038)	(932.757)	(1.014)	(947.701)
Base NorthEast				
South	0.968	0.896	0.945	0.821
	(0.314)	(0.989)	(0.343)	(1.027)
Centre	0.691*	0.342*	0.682*	0.322*
	(0.144)	(0.219)	(0.145)	(0.221)
North-West	0.497***	0.175***	0.519***	0.180***
	(0.097)	(0.097)	(0.113)	(0.115)
Income per region	1.000	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)
Costs on output	2.026	8.218	1.910	7.221
	(2.770)	(33.582)	(2.648)	(30.614)
Labour on output	1.129	1.435	1.149	1.531
	(0.212)	(0.796)	(0.229)	(0.922)
Ratio farm other activities	1.024**	1.074**	1.024**	1.074**
	(0.011)	(0.034)	(0.011)	(0.035)
Base Productivity				
Environmental	1.452***	2.946***	1.456***	3.052***
	(0.130)	(0.937)	(0.132)	(1.019)
Market Competitiveness	1.428***	2.780***	1.439***	2.926***
	(0.070)	(0.369)	(0.071)	(0.408)
Network/Chain Structure	1.350***	2.307***	1.349***	2.350***
	(0.133)	(0.624)	(0.133)	(0.653)
Resources	1.397***	2.585***	1.386***	2.569***
	(0.111)	(0.634)	(0.113)	(0.662)
Duration(month)	0.987*	0.963**	0.988*	0.962**
	(0.006)	(0.018)	(0.007)	(0.019)
Leader collaboration in other OGs	1.116	1.387	1.128*	1.447*
	(0.074)	(0.282)	(0.072)	(0.291)
Leader past Experience in OG projects	1.063	1.198	1.043	1.136
	(0.127)	(0.421)	(0.123)	(0.406)
Base Leader: Research&Education				
Leader: Farms	0.980	0.940	0.986	0.958
	(0.065)	(0.189)	(0.068)	(0.203)

Leader: Training entities	0.929	0.806	0.935	0.817
	(0.257)	(0.637)	(0.278)	(0.714)
Leader: Advisory entities	1.046	1.151	1.036	1.119
	(0.119)	(0.414)	(0.122)	(0.424)
Leader: Associations of farmers	0.894	0.724	0.895	0.722
	(0.094)	(0.209)	(0.097)	(0.220)
Leader: Other private entities	1.064	1.217	1.058	1.199
	(0.244)	(0.903)	(0.249)	(0.930)
Leader: Other public entities	1.101	1.360	1.094	1.340
	(0.196)	(0.824)	(0.191)	(0.806)
Leader: Other firms	0.923	0.792	0.926	0.793
	(0.122)	(0.301)	(0.121)	(0.304)
Number of Partners	1.006	1.019	1.004	1.013
	(0.009)	(0.027)	(0.009)	(0.027)
Typology of Partners:				
Farmer	1.332***	2.353***	1.369***	2.609***
	(0.101)	(0.558)	(0.104)	(0.640)
Training	1.086	1.279	1.078	1.260
	(0.103)	(0.349)	(0.103)	(0.358)
Advisory	0.893	0.713	0.889	0.699
	(0.074)	(0.178)	(0.075)	(0.181)
Research&Education	1.039	1.122	1.038	1.122
	(0.080)	(0.256)	(0.083)	(0.273)
FarmsAssociation	0.869*	0.657	0.873*	0.659
	(0.073)	(0.173)	(0.071)	(0.173)
Other public	1.138	1.472	1.148	1.523
	(0.102)	(0.374)	(0.108)	(0.418)
Other private	0.943	0.840	0.943	0.837
	(0.048)	(0.130)	(0.055)	(0.149)
Other firms	1.037	1.116	1.039	1.124
	(0.066)	(0.212)	(0.069)	(0.225)
Base commodity: Livestock				
Veg&Fruit	1.228**	1.769**	1.240**	1.842**
	(0.120)	(0.473)	(0.123)	(0.514)
Main crops	1.136	1.406	1.141	1.430
	(0.117)	(0.384)	(0.117)	(0.398)
Industrial crops	1.418**	2.847**	1.438***	3.042**
	(0.204)	(1.386)	(0.199)	(1.458)
Viticulture	1.218	1.726	1.212	1.714
	(0.181)	(0.743)	(0.183)	(0.762)
Multi-sector	1.056	1.150	1.067	1.185
	(0.094)	(0.262)	(0.089)	(0.259)
Forestry	1.726***	6.142***	1.761***	6.929***
	(0.260)	(3.694)	(0.263)	(4.324)
Other	2.135***	17.118**	2.256***	24.394*
	(0.584)	(24.380)	(0.656)	(39.777)
Constant	0.000***		0.000***	
	(0.000)		(0.000)	
Observations	646	646	646	646
lnalpha			0.120***	
			(0.021)	

*Robust SE in parentheses *** p<0.01, ** p<0.05, * p<0.1 Exposure Total Budget per OGs*

APPENDIX B: Definitions and codes for LLM

The selection of the Large Language Model (LLM) was based on the objectives of this study. Specifically, we employed LLaMA 3, an open-source model developed by Meta, designed for text comprehension and generation. Similar to ChatGPT, LLaMA 3 is a general-purpose model, making it well-suited for extracting a specific data variable from unstructured text. However, a limitation of this study is the need to integrate multiple models to assess the accuracy and effectiveness of the generated responses. Future research can build upon our approach by utilizing the code, which is freely available on GitHub, to facilitate comparative analyses.

The present appendix presents the parameters used in the LLM and the definition of innovation. Accordingly, the main parameter underlined is the *temperature* of the model. Results are compared and collected with temperature settled at level 0.3 and level 0.1, while the number of *CPU* used are settled at 32.

To assist the AI in extracting the number of actual innovations within each Operational Group (OG), a clear definition of innovation is provided. The prompt was formulated based on the Oslo Manual (4th edition) and relevant literature on OGs best practices.

The following code snippet outlines the definition of innovation as applied in this study.

```
response_format={ "type": "json_object" },
    messages=[
        {"role": "system", "content": "You are an expert about the
definitions contained in the Oslo Manual 4th edition. You are an expert in
agricultural innovations and agricultural economics. According to the Oslo
Manual IV edition the term “innovation” can be used in different contexts to
refer to either a process or an outcome (product). \
“A product innovation is a new or improved good or service that differs
significantly from the firm’s previous goods or services and that has been
introduced on the market.” Product innovations must provide significant
improvements to one or more characteristics or performance specifications.
Relevant functional characteristics include quality, technical
specifications, reliability, durability, economic efficiency during use,
affordability, convenience, usability, and user friendliness. Product
innovations do not need to improve all functions or performance
specifications and must be made available to potential users.\
A process innovation is the implementation of a new or significantly improved
production or delivery method. This includes significant changes in
techniques, equipment and/or software.\
“A business process innovation is a new or improved business process for one
or more business functions that differs significantly from the firm’s
previous business processes and that has been brought into use in the firm.”
A business process innovation can involve improvements to one or more aspects
of a single business function or to combinations of different business
functions. According to Eurostat (2013) the functions can be divided into: \
```

• distribution and logistics: transportation activities, warehousing and order processing.\
• marketing, sales and after-sales services: market research, advertising, direct marketing services (telemarketing), exhibitions, fairs and other marketing or sales services; also included are call-centre services and after-sales services such as help-desks and other customer support services.\
• information and communication technology (ICT) services: information technology (IT) services and telecommunication (IT services including hardware and software consultancy, customised software data processing and database services, maintenance and repair, web-hosting, as well as other computer-related and information services, but excluding packaged software and hardware).\
• administrative and management functions: legal services, accounting, book-keeping and auditing, business management and consultancy, human resources (HR) management (e.g. training and education, staff recruitment, provision of temporary personnel, payroll management as well as health and medical services), corporate financial and insurance services; also included are procurement functions.\
• engineering and related technical services: engineering and related technical consultancy, technical testing, analysis and certification; also included are design services.\
According to Nelson and Winter (1982), technical innovation often emerges as firms experiment and modify their routines in response to competitive pressures and market demands. Thus, in the Operational Group context, innovations often refer to technical innovations that arise from experimental and scientific research and can often be the answer to farmers' needs, bringing the new technology to the local/national market. Technical innovations are not understood as radical innovations, but also as incremental innovations, which improve certain technologies to adapt them to the needs of agriculture. In addition, Italy's relationship with tradition and agricultural culture has led to the view that a new way of bringing into the market traditional foods is a process innovation, while a revaluation of traditional foods is a product innovation.\
The innovation minimum requirement is the novelty. In Italian Operational Groups novelty of innovation can be measured as new for the patterns of the project, or for the entire commodity sector. For Italian operational Groups is not an innovation: routine changes or updates; simple capital replacement or extension; product introductions that only involve minor aesthetic changes; the outputs of creative and professional service firms, such as reports for clients are not by default an innovation for the firms that develop them. "+