CARMA 2025

7th Int. Conf. on Advanced Research Methods and Analytics
2-4 July 2025, Rome, Italy
DOI: https://doi.org/10.4995/CARMA2025.2025.20579

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

Franseo del Puente
University Parthenope, Italy.

How to cite: Del Puente, F. 2025. Asessing the Impact of a European Union's Policy on Agricultural Innovation in Italy. In: 7th International Conference on Advanced Research Methods and Analytics (CARMA 2025). Rome, 2-4 July 2025. https://doi.org/10.4995/CARMA2025.2025.20579

Abstract

This study assesses innovation outputs of Operational Groups (OGs), a key mechanism for promoting innovation in European agriculture. Using a Large Language Model, we construct an Italian OG dataset to quantify innovation outputs and contributing factors. Count data analyses reveal significant disparities in innovation output across regions, thematic areas, commodities, partner composition, and leadership engagement. Higher innovation output is associated with thematic areas focused on market competitiveness, supply chain management, and resource management; commodities like forestry, industrial crops, and vegetables and fruits; collaborations involving farmers, research and education institutes, and training organizations; and OG leaders involved in multiple OGs.

Keywords: *Innovation, Operational Groups, Italian Agriculture, Large Language Model.*

1. Introduction

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 20142020 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_net1, with a budget of € 784,244 and a consortium of six farmers, two freelance advisors, and a university, aims to create a new mulberrybased agri-food supply chain in the Piedmont region. This initiative not only expands mulberry sales but also increases their economic value by converting pruning waste into biomass. The 20212027 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

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

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. 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.

2. Data and Methodology

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.

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:

$$E[Ii|OGi,Rit0,Ti,\tau i] = \mu i\tau i = e\delta OGi+\beta Rit0+Ti+ i \quad (1)$$

where I_i represents the number of innovative projects, OG_i includes OG and project characteristics, and R_{it0} includes regional characteristics at base time t_0. Equation (1) also controls for fixed effects for OG establishment year (T_i). The error term $_i$ 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 $\tau_i = e^{\ i}$ is independent of regressors conditional on OG_i, R_{it0}, 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 $e^{\delta OGi+\beta Rit0+Ti}$ 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.

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.

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.

3. 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.

On average, each OG develops approximately three innovations. The kernel density distributions of the number of innovations, shows that 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).

There is a 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%),[2] 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), while the total budget was highest for OGs focusing on vegetables and fruits and second lowest for those focused on forests. Using Trienekens (2011)'s framework, we classify OGs into key thematic areas: productivity improvement (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. The average budget per OG is higher for those focusing on market competitiveness and supply chain management than for those emphasizing productivity.

2 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.

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%.[3] 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.

4. 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. 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.

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. Using the IRRs, we calculate the percentage change in innovation output as follows: (IRR-1) *100%. 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

3 "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.

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.

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. 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.

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, 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.

5. Conclusion

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, highquality 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.

References

Arzeni, A., Giarè, F., & Vagnozzi, A. (2023). The role of Operational Groups in fostering agricultural innovation: Evidence from Italy. *Agricultural and Food Economics*, 11(1), 1–20.

Barca, F. (2009). *An agenda for a reformed cohesion policy. A place-based approach to meeting European Union challenges and expectations*. Bruxelles: DG REGIO, European Commission.

Castelnovo, P., Morone, P., & Sisto, R. (2020). Territorial capital and regional disparities in Italy: The role of innovation and productivity. *Regional Studies*, 54(7), 905–917.

Chen, L., Giray, L., & Wang, X. (2024). Prompt engineering and structured system instructions for reliable large language model outputs. *Journal of Artificial Intelligence Research*, 79, 1123–1150.

Cristiano, S., & Proietti, P. (2018). Farmers as local innovators: The role of Operational Groups in the Italian AKIS. *Bio-based and Applied Economics*, 7(3), 263–281.

Devetak, M., & Mandel, A. (2023). Using large language models to construct economic datasets: Opportunities and challenges. *Journal of Economic Methodology*, 30(4), 345–362.

del Puente, M., Proietti, P., & Cristiano, S. (2024). Innovation goals and traditional food valorisation in Italian Operational Groups: An Oslo Manual perspective. *Food Policy*, 124, 102613.

European Commission. (2021). *Regulation (EU) 2021/2115 of the European Parliament and of the Council of 2 December 2021 establishing rules on support for strategic plans under the Common Agricultural Policy (CAP)*. Official Journal of the European Union.

European Commission. (2023). *The Common Agricultural Policy 2023–2027: Supporting a sustainable future for EU agriculture*. Brussels: European Commission.

European Union. (2024). *The future of agriculture in the European Union: A vision for sustainable and resilient food systems*. Brussels: Publications Office of the European Union.

FAO. (2024). *The State of Food and Agriculture 2024*. Rome: Food and Agriculture Organization of the United Nations.

Fuglie, K., Gautam, M., Goyal, A., & Maloney, W. (2024). Harvesting Prosperity: Technology and Productivity Growth in Agriculture. Washington, DC: World Bank.

Giarè, F., & Vagnozzi, A. (2021). Financial procedures and innovation processes in Italian EIP-AGRI Operational Groups. *Agricultural and Food Economics*, 9(1), 1–18.

Giray, L. (2023). Prompt engineering techniques for improving large language model reliability. *AI & Society*, 38(2), 521–534.

Guerrero-Ocampo, J., Klerkx, L., & Dessein, J. (2022). Agricultural innovation partnerships and multi-actor governance: Lessons from EIP-AGRI. *Journal of Rural Studies*, 89, 143–154.

Maziliauskas, A., Klerkx, L., & Liepins, I. (2018). The role of Operational Groups in agricultural knowledge and innovation systems. *Agricultural Systems*, 165, 65–75.

Molina, A., Proietti, P., & Cristiano, S. (2021). Drivers of participation in agricultural innovation partnerships: Evidence from Italian Operational Groups. *Sustainability*, 13(14), 7894.

SCAR AKIS. (2019). *Preparing for future AKIS in Europe*. Brussels: Standing Committee on Agricultural Research.

Takahashi, K., Muraoka, R., & Otsuka, K. (2020). Technology adoption, impact, and extension in developing countries: A review. *Annual Review of Resource Economics*, 12, 293–315.

Trienekens, J. H. (2011). Agricultural value chains in developing countries: A framework for analysis. *International Food and Agribusiness Management Review*, 14(2), 51–82.

Wang, Y., Li, H., & Zhang, T. (2024). Evaluating factual consistency and verifiability in large language model outputs. *Transactions of the Association for Computational Linguistics*, 12, 245–260.