Andrea BONFIGLIO*

Do EIP-AGRI operational groups improve farmers' performance? An analysis of treatment effects in intensive farming systems

The Operational Groups (OGs) of the European Innovation Partnership for Agricultural Productivity and Sustainability (EIP-AGRI) were introduced by the 2014-2020 Common Agricultural Policy to foster competitive and sustainable farming and forestry. The objective of this paper is to assess the economic and environmental impacts of participating in the EIP-AGRI OGs located in the Italian region of Emilia-Romagna. Performance of participants in OGs is compared with that of non-participants, who are selected by applying propensity score matching techniques to an Italian farm accountancy data network sample of 3204 farmers observed in the period 2017-2020. Logistic regressions are used to measure both propensity scores and the average treatment effect on the treated, while one-to-many optimal matching without replacement is adopted to form the control group. The resulting sample is composed of 270 observations, of which 45 are treated subjects. Results indicate that the OGs analysed might have contributed to improving fertiliser management and profitability levels in participating farms, but they failed to preserve biodiversity and reduce the consumption of pesticides and other inputs such as water, energy, and fuels. To increase the effectiveness of OGs, policy makers are advised to condition projects on the actual experimentation and implementation of agricultural innovations and apply a performance-based system of indicators for the assessment of the ex-ante and ex-post impacts of farm management.

Keywords: European Innovation Partnership for Agricultural Productivity and Sustainability; Operational Groups; farm performance; treatment effects; propensity score matching

JEL classification: Q12

* Research Centre for Agricultural Policies and Bioeconomy, CREA – Council for Agricultural Research and Economics, Via Thomas Edison 2, Osimo (AN), Italy, 60027. E-mail: andrea.bonfiglio@crea.gov.it.

Received: 14 March 2023; Revised: 3 July 2023; Accepted: 10 July 2023.

Introduction

The Operational Groups (OGs) of the European Innovation Partnership for Agricultural Productivity and Sustainability (EIP-AGRI) represent an important policy instrument that was introduced by the 2014-2020 Common Agricultural Policy (CAP) to foster competitive and sustainable farming and forestry by using an interactive approach to innovation (European Commission, 2019; Van Oost and Vagnozzi, 2021). OGs take the form of groups of actors having diverse practical and scientific backgrounds, such as farmers, agribusinesses, researchers, and advisors, who come together for a practical reason, i.e., to respond to real problems through the implementation of innovative solutions. To be financed, OGs are asked to draw up a plan containing a description of the innovative project to be developed, tested, adapted, or implemented as well as a description of the expected results and the contribution to the EIP objective of enhancing productivity and sustainable resource management. Moreover, they are required to disseminate the results of projects through the EIP network in order to favour the adoption and the diffusion of innovation amongst farmers (Art. 55-57 of Regulation EU No 1305/2013).

The assessment of OGs involves several aspects related to the implementation of the funding programme, the selection of projects, and their results and effects (Gehrlein and von Kutzleben, 2016). The existing studies about OGs mainly focus on the first two aspects as well as on key factors for successful projects, governance, and consistency with policy objectives, while little or no emphasis is placed on the assessment of impacts (Cristiano and Proietti, 2018; Eckerberg *et al.*, 2023; Giarè and Vagnozzi, 2021; Harrahill *et al.*, 2022; Knotter *et al.*, 2019; Maziliauskas *et al.*, 2018; McCarthy *et al.*, 2021; Molina *et al.*, 2021; Parzonko *et al.*, 2022; Piñeiro *et al.*, 2021; Schreuder *et al.*, 2022). Factors related to availability and methods of collection of project data as well as the stage at which these studies were carried out may explain this shortcoming. The evaluation of the results represents a step that is fundamental to understanding the real effectiveness of OGs. Through a comparison of impacts with initial objectives, it makes it possible to verify whether the OG setting should be adjusted to remove its limits and improve its potential in the 2023-2027 programming period.

The aim of this paper is to assess the economic and environmental impacts of participating in OGs on farmers. In other words, the objective is to verify if the participation in OGs, through the application and the experimentation of agricultural innovations, helped to improve performance of farmers. To the authors' knowledge, this research represents one of the first attempts in this direction and can therefore be considered as a novel contribution.

For the purposes of this study, propensity score matching is adopted (Caliendo and Kopeinig, 2008; Guo *et al.*, 2020; Rosenbaum and Rubin, 1983). This is a statistical technique that matches treated subjects with one or more untreated cases based on their propensity scores. This helps to reduce selection bias in quasi-experimental and observational studies. In this study, the treatment is represented by the participation in concluded projects of OGs while the potential outcomes are assessed by comparing the variations of a set of monetary and quantitative indicators measured in the period 2017-2020 for both treated and untreated subjects. Monetary indicators include output, variable costs, fertiliser and pesticide expenditure, expenditure on water, energy and fuels, and net farm income. Quantitative indicators comprise the used quantity of phosphorus and nitrogen contained in fertilisers, the used quantity of fertilisers and pesticides, and the number of cultivated crops.

Logistic regressions are used to measure both propensity scores and the Average Treatment effect on the Treated (ATT), i.e., the average effect of treatment on those subjects who ultimately received the treatment (Imbens, 2004). As covariates, a set of socio-economic variables that are supposed to affect both the treatment and the outcomes are analysed.

This research is carried out by using the Farm Accountancy Data Network (FADN) sample of Italian farmers. FADN is employed to retrieve information about the variables investigated concerning the farmers participating in OGs whose projects were concluded. The focus of this study is on the Italian region of Emilia-Romagna. This region is a particularly suitable case for this analysis. According to Italian National Rural Network (NRN) statistics, in Italy, in September 2021, there were overall 656 OGs, of which 213 (over 30%) concentrated in Emilia-Romagna. The second region for number of OGs is Sicily with 61 projects financed. Moreover, according to the national database of OGs (containing detailed information about 633 OGs), 92 out of 144 projects for which it is possible to know if they were completed are in Emilia-Romagna. Another reason why this region represents an interesting case concerns its main morphological characteristics: about 70% of Utilised Agricultural Area (UAA) is situated on flat land compared with a national average of 33%. This has made possible a wide diffusion of mechanisation and intensive agriculture with significant negative impacts on the environment (Menta *et al.*, 2017). OGs are therefore called upon to favour the diffusion of more environmentally friendly techniques and reduce the pressure of agriculture on the environment.

The rest of this paper is organised as follows. Section 2 offers a brief overview of existing studies on OGs. In addition, it examines the main issues in measuring the effects of OGs related to the availability and collection of data and the type of impacts to be assessed consistently with the objectives of EIP-AGRI. Section 3 illustrates the methodology, the variables and the data used. Sections 4 and 5 present and discuss the results of this analysis, respectively. Finally, Section 6 provides some concluding remarks.

Literature review

The main question when evaluating EIP-AGRI is to which extent innovation, cooperation, and building the knowledge base in rural areas are supported by Rural Development Policy (RDP) interventions (European Commission, 2014). To answer this question, aspects such as the implementation of the funding programme, the selection of projects, and their results and effects need to be examined in more detail as part of the evaluation (Gehrlein and von Kutzleben, 2016). In terms of implementation, the programmed funding objects, funding conditions, and procedures are relevant. The central issue is whether regulations are capable of fostering innovations. As regards the selection procedure, the criteria that guide the decision on financing projects are also of great importance because the identification of innovative projects that respond to real problems of farmers strongly depend on them. However, if the objective is to evaluate the real effectiveness of the funded projects, the knowledge of their results and impacts becomes essential.

In literature, existing studies about OGs have focused on topics such as progress in implementation of OGs (Knotter *et al.*, 2019; Schreuder *et al.*, 2022), key factors for successful projects (Harrahill *et al.*, 2022; Maziliauskas *et al.*, 2018; McCarthy *et al.*, 2021; Molina *et al.*, 2021; Parzonko *et al.*, 2022), performed functions (Piñeiro *et al.*, 2021), governance processes (Giarè and Vagnozzi, 2021), and consistency with the objectives of European strategies (Cristiano and Proietti, 2018; Eckerberg *et al.*, 2023; Giarè and Vagnozzi, 2021).

More specifically, Knotter *et al.* (2019) assessed the state-of-play of the setting-up and implementation of OGs until 2018. By combining several methods of investigation (cluster analysis, survey, and case studies), they concluded that OGs are effective in tackling farmers' needs in a practical and collaborative way on topics related to both competitiveness and environmental sustainability. Schreuder *et al.* (2022) reviewed the OGs focused on topics related to grassland using the EIP-AGRI database and an online survey. They observed that the themes addressed by OGs are less focused on environmental issues than the recommendations coming from specific EIP-AGRI focus groups.

Maziliauskas *et al.* (2018) identified the external and internal factors that influence the effectiveness of OGs by a force field analysis. They found that the biggest negative impact comes from the lack of cooperation between partners and that internal factors such as partner involvement and constant monitoring of achievements based on a list of indicators play an important role in a project's success. McCarthy *et al.* (2021) explored the motivations of a small group of actors who established an OG in Ireland using an assemblage-based approach. Their main conclusion is that the motivations of different subjects involved influence each other and take into consideration future scenarios and new possibilities. The outcomes of the EIP-AGRI initiative are therefore affected by this process of reciprocal influence. Molina *et al.* (2021), through the analysis of a case study of an Italian OG, highlighted the factors that could influence and foster the interactive innovation process. They concluded that farmers are active players in the design and implementation phases and that motivation, commitment, trust, and an open communication among different actors are key factors for the success of a project. Parzonko *et al.* (2022) analysed the role of innovation brokers in the setting up of OGs in Poland by a survey addressed to a selected group of people who participated or showed interest in a web initiative realised by an advisory centre to support the creation of OGs during the COVID-19 pandemic. They demonstrated that the innovation broker played a key role in identifying subjects willing to cooperate, obtaining funds and preparing project proposals and documents related to the functioning of the OG. Harrahill *et al.* (2022) examined the degree of involvement of farmers in an Irish OG aimed at producing and transforming biomass into energy. By using social network analysis combined with interviews conducted with farmers and non-farmer participants in the OG, they found that, despite farmers were highly involved as input suppliers, the level of influence they exerted in several other areas, such as the logistical and managerial ones, was relatively limited and this can hinder the success of future projects having similar objectives.

Piñeiro *et al.* (2021) conducted an online survey addressed to members of Spanish OGs in order to identify the intermediary functions carried out by OGs. They found that OGs can manage the entire innovation process by encouraging collaboration, sharing information, and developing joint projects. OGs also make innovation demand emerge by identifying opportunities, developing studies, and seeking solutions that meet the needs of OGs and their members. Finally, they search for economic and institutional support and encourage external collaboration to find resources and disseminate knowledge and solutions.

Cristiano and Proietti (2018) investigated the relationship between Italian OGs and research programs, specifically Horizon 2020, by collecting data from direct interviews, semi-structured questionnaires, focus groups, and workshops. They highlighted that there is no interaction between research and innovation projects, and this slows down innovation processes and contrasts with the objectives of EIP-AGRI of creating synergies and value added by integrating different policy tools. Giarè and Vagnozzi (2021) compared the rules and implementation criteria adopted by some Italian managing authorities to finance OGs in order to analyse the impact of different governance choices on the functioning of OGs. They concluded that rules and criteria are inadequate in some cases, mainly regarding the definition of innovation needs, the involvement of all actors, the construction of a common strategy, and the connection with the measures addressed to finance investments, and this can negatively affect the effectiveness and consistency of projects with the objectives of RDP. More recently, Eckerberg *et al.* (2023) analysed the state's steering capacity of spreading "green innovation" in the agricultural sector of Sweden through the implementation of EIP-AGRI. By examining the information from the national database of OGs financed in Sweden and from interviews with key individuals engaged in the program administration, they found that, in contrast with policy objectives both at the general policy level and in the EIP-AGRI regulation, "green innovation" was only marginally supported by prioritising aspects related to competitiveness and placing less emphasis on those related to the environment and climate change.

Although these studies offer interesting indications for the aims of evaluating the EIP-AGRI initiative, no conclusion is provided about the real impact of participation in OGs on farm performance. One reason is related to the fact that several studies were conducted when few or no projects had yet been completed. Another reason that makes impact assessment difficult concerns data availability. The main instrument used for dissemination of innovative projects aimed at rural development and agriculture is represented by the publication of project data on online databases (Ibáñez-Jiménez *et al.*, 2022). The official database of European OGs can be freely consulted on the EIP-AGRI website. The available data (last access in December 2022) provide clear information about the objectives pursued, the activities to be carried out and the main innovations planned. However, little or no information is provided with reference to the results obtained. This mainly depends on the system of data collection that was implemented to retrieve information about OGs. In fact, the data requested adhere to an official template, which only asks for some qualitative information (European Commission, 2016). In addition, only a part of this information is categorised and is thus in a format suitable for processing. Moreover, much desirable information, such as the detailed characteristics of the participating farmers as well as the changes in economic aggregates (output, costs, inputs, income, etc.) following the execution of the project, is not present, impeding the environmental and economic impact assessment of OGs.

At the time this research was conducted, several projects were completed, and the relevant impacts could therefore be assessed. For the investigation of results and impacts, different methodologies can be adopted such as document analyses (interim and final reports), ad-hoc surveys, and self-assessment of performance (Gehrlein and von Kutzleben, 2016). However, these methods not only are costly and time-consuming, but the relevant results could be affected by the widespread absence of internal accounting systems, especially in countries such as Italy, which prevents farmers from knowing exactly if and how the variables of interest have changed over time. Further issues, which can negatively affect the goodness of the results, are interpretation difficulties and farmers' reluctance to provide truthful answers in consideration of the public subsidies received. An alternative approach involves the use of an already existing and official accounting system, i.e., the FADN data, by matching the information about the partnership of OGs with that contained in FADN. This system offers a great quantity of socio-economic and environmental information and can therefore be used to effectively assess the performance of farmers participating in OGs (treated), comparing it with that of farmers who did not participate (untreated).

Another important issue in evaluating OGs concerns what kind of impacts should be measured. A central question should be if and to what extent the productivity of farms has increased. Another crucial question is whether progress toward sustainability has been achieved. This is because improvements in productivity and sustainability represent the main objectives of the EIP-AGRI initiative. Therefore, the assessment of the impacts on these two main aspects is of great importance.

Productivity is commonly defined as the relationship between outputs and inputs. There are several ways to measure productivity, which depend on the purpose of measurement and the availability of data (OECD, 2001). In this study, output is measured as market value and is expressed per hectare. Therefore, land productivity is considered. The notion of sustainable agriculture is particularly complex, and this makes its use and implementation quite hard (Velten *et al.*, 2015). According to Pretty (2008), the key principles for sustainability are: to integrate biological and ecological processes into food production processes, to minimise the use of those non-renewable inputs that are harmful to the environment or to the health of farmers and consumers, and to make productive use of the knowledge and skills of farmers as well as of people's collective capacities to work together to solve common agricultural and natural resource problems. Agricultural sustainability is thus a very broad concept involving three "pillars": environmental, economic, and social (Purvis *et al.*, 2019). This study only concentrates on some environmental and economic aspects. As regards the environmental dimension, the focus is on the capability of reducing environmental impact by diminishing the used quantity of inputs and the level of specialisation, i.e., the tendency towards monoculture, which can undermine biodiversity (Altieri, 1999), soil fertility (Liu *et al.*, 2006), and the capability of facing climate change (Lin, 2011). Besides the rationalisation in the use of inputs, specialisation is another issue that can be faced by OGs through projects aimed at introducing new or rediscovered crop varieties. With reference to economic sustainability, this study focuses on farmers' ability to reduce their costs and improve their profitability, i.e. generate income. Both productivity and sustainability are tightly connected with profitability. An increase in the output-input ratio, as well as the adoption of environmentally friendly techniques that serve to reduce the quantity of used inputs, can increase profitability. The latter is one of the motivations, or, in some cases, may be the only motivation, which might induce farmers to decide to participate in OGs. Understanding the impact of EIP-AGRI on profitability is thus extremely important for policy makers since the degree of participation in OGs and, by extension, the success of this policy instrument, which has also been proposed again for the 2023-2027 programming period, strongly depends on it.

Materials and methods

The model

Propensity score matching allows the building of matched sets of treated and untreated subjects who share similar propensity scores (Caliendo and Kopeinig, 2008; Guo *et al.*, 2020; Rosenbaum and Rubin, 1983). A propensity score is defined as the conditional probability of being selected into the treatment group given a set of covariates or observed characteristics for group members, i.e.:

$$p(\mathbf{X}) = Pr\{Tr = 1|\mathbf{X}\} = E\{Tr|\mathbf{X}\} \tag{1}$$

where Tr={0,1} is an indicator variable for treatment group selection and $\mathbf{X}$ is a multidimensional vector of covariates. Propensity scores therefore describe the likelihood that a population member would be selected into the treatment group based on a set of model covariates. Propensity score estimates are used to construct a comparison group. The Average Treatment Effect (ATE), based on an outcome measure (Y), is then estimated as:

$$ATE = E\{Y_1 \mid Tr = 1\} - E\{Y_0 \mid Tr = 0\} \tag{2}$$

where Y_1 and Y_0 are the outcome measures for treated and untreated subjects, respectively. The ATE refers to the entire population. The ATT, used in this study, is a related measure of treatment effect and measures the ATE only on those subjects who received the treatment (Imbens, 2004).

In contrast to randomised designs, propensity scoring techniques use a set of covariates to model the treatment group selection process. Moreover, these methods cannot adjust for unobserved covariates. The main assumption is therefore that observations with the same propensity score have the same distributions for observable and unobservable characteristics. This connects propensity scoring with the assumption of ignorable treatment group assignment and the conclusion that the ATE estimate is unbiased (Stone and Tang, 2013).

Following a commonly used approach (Austin, 2011; D'Agostino, 1998; Rosenbaum and Rubin, 1983), propensity scores are estimated by logistic regression where the dichotomous outcome is treatment group assignment (1 and 0 for treated and untreated subjects, respectively) and predictors are a set of measured covariates. Once propensity scores are computed, the following step consists in creating balanced intervention and comparison groups. There are several approaches for creating these groups, some of which include exact matching, nearest neighbour matching, and optimal matching (Rosenbaum, 1989; Rubin, 1973). Further decisions concern the number of nonparticipants to be matched to each participant (one-to-one or one-to many matching) and whether replacement (i.e., matching nonparticipants multiple times to participants) is allowed. The choice can be made on the basis of different considerations (Stuart, 2010). Several studies have empirically demonstrated the potential benefits of one-to-many matching and proposed the optimal matching ratio for decreasing bias but increasing power (Austin, 2010; Cenzer *et al.*, 2020; Rassen *et al.*, 2012). In particular, Cenzer *et al.* (2020) focused on situations where the number of treated subjects is very small. Through a Monte Carlo simulation, they showed that, when the number of treated subjects available is between 25 and 50, the use of optimal matching without replacement and with one-to-five matching ratio proves to be the best option. Compared to greedy matching (such as nearest neighbour matching), optimal matching is a more complex approach whose goal is to find the matched samples with the smallest average absolute propensity score distance across all the matched pairs. In consideration of the limited size of the sample available (see Section 3.2), this method is therefore adopted in this study.

Once the matches are created, the quality of the matches is assessed in order to ensure that the comparison group has a distribution of propensity scores similar to the intervention group. Matches are assessed by comparing the balance both numerically and visually (Stuart, 2010). Visual diagnosis of balance is conducted here by inspecting distribution of propensity scores before and after matching. Numerical diagnosis of balance is instead carried out by evaluating the covariate balance. This is made by comparing the standardised difference of group propensity score means (SMD). For continuous and dichotomous variables, SMD for covariate X takes the following form, respectively:

$$SMD_X = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{(Var_1 - Var_0)/2}} \tag{3}$$

$$SMD_X = \frac{\hat{p}_1 - \hat{p}_0}{\sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_0(1-\hat{p}_0)/2}} \tag{4}$$

where $\bar{X}_1$ and $\bar{X}_0$ are sample means, Var_1 and Var_0 are sample variances, and, finally, $\hat{p}_1$ and $\hat{p}_0$ are the prevalence of dichotomous variables in the treated and untreated units, respectively. X is considered as balanced if the absolute SMD value is lower than 0.25 (Imbens and Wooldridge, 2009).

The ATT is then estimated by running a logistic regression over matched subjects with cluster-robust standard errors (Abadie and Spiess, 2022), where the dichotomous variable is the outcome analysed while the only predictor is represented by the treatment group selection. The regression gives an estimate of the logarithm of odds ratio, i.e., the ratio of the probability that a given outcome occurs in treated subjects to the probability that the same outcome occurs in untreated units.

Analyses were conducted using packages MatchIt 4.4.0, for propensity score matching; stats, for logistic regressions; lmtest 0.9-40 and sandwich 3.0-2, for estimating cluster-robust standard errors, in statistical software R 4.2.1.

The variables and the dataset used

The outcomes analysed in this study concern both economic and environmental aspects and are measured as monetary and quantitative indicators. For the choice of indicators, the approach followed is that of Cisilino *et al.* (2019), who carried out a conceptually similar analysis consisting in evaluating the environmental and economic effects of organic farming subsidies using propensity score matching techniques applied to a sample of FADN farms. More specifically, the monetary indicators used to assess performance of farmers are output (i.e., total revenues), variable costs, fertiliser and pesticide expenditure, and net farm income. Since the rationalisation in the use of water and energy represents another important objective of EIP-AGRI, expenditure on water, energy, and fuels is also considered. All variables are expressed per hectare. The quantitative indicators used to measure farm performance are instead the used quantity of phosphorus and nitrogen contained in fertilisers, the used quantity of pesticides, and the number of cultivated crops as an indicator of biodiversity. The overall quantity of fertilisers is also considered to integrate the analysis of the pesticides used. The quantity of used water, energy and fuels could not be analysed because of data unavailability. Quantities of fertilisers, phosphorus, nitrogen, and pesticides are expressed as quintals and per hectare. Outcomes are assessed as binary variables, which take one if the average variation of the indicators is positive and zero if null or negative.

As for the covariates to be included in the propensity score model, the choice is not straightforward since there are several possible variables that can be selected (Austin, 2011). They can be all baseline covariates, all baseline covariates that are associated with treatment, all covariates that influence the outcome (i.e., the potential confounders), and all covariates that affect both treatment and the outcome (i.e., the true confounders).

Since the propensity score is defined to be the probability of treatment assignment, there are theoretical reasons in favour of the inclusion of only those variables that affect treatment assignment (Austin, 2011). However, Austin *et al.* (2007) showed that including potential or true confounders does not introduce additional bias and results in estimates of treatment effect with greater precision. Similarly, Brookhart *et al.* (2006) suggested that potential confounders should be preferred to variables only affecting treatment since the inclusion of the latter increase the variance of the estimated treatment effect without a concomitant reduction in bias. In practice, it is quite hard to distinguish between different types of variables. Moreover, most baseline covariates likely affect both treatment assignment and the outcome. Therefore, it is better to include all measured baseline characteristics in the propensity score model. However, an important condition is that variables are measured at baseline and are not post-baseline covariates, since the latter may be influenced or modified by the treatment (Austin, 2011).

The data used in this study come from the Italian FADN. This database offers a very large set of variables. To contain the number of features, a subset of all variables available was selected. Data selection was focused on variables that can affect both the participation in OGs and outcomes. Moreover, the selection process was led by the need to consider both farmer and farm characteristics as well as various economic, environmental, social, and formal aspects in such a way to focus the analysis on a homogeneous sample. Subjective factors related to personal attitudes and motivations, which could also influence participation (Molina *et al.*, 2021), were neglected for data unavailability. As regards farmer characteristics, gender, age, education, and access to measures of RDP are considered while, with reference to farm features, altitude, productive specialisation, organic farming, on-farm diversification, legal form, land, livestock, labour, family work, and machinery are investigated. Most variables are categorical except for on-farm diversification, land, livestock, labour, family work, and machinery, which are measured as continuous. Gender takes value of one for females and zero for males. Age is modelled by a dichotomous variable taking unitary values if farmers are young according to the threshold set by the CAP for accessing specific measures in favour of farmers with no more than 40 years of age. Education is also a binary variable taking one in the case of a high-medium level of education. The variable relating to access to measures of RDP takes the value of one (zero) if farmers applied (did not apply) for measures of RDP other than those relating to OGs (i.e., measure 16.1). This variable is introduced since both participation in OGs and outcomes can also be affected by the knowledge of RDPs and the activation of other RDP measures. Altitude is represented by two binary variables that take unitary value if farms are localized in flat areas and in hills, respectively, while they are zero if farms are situated in the mountains. Productive specialisation is measured by four dummies related to arable crops, horticulture, livestock, and permanent crops, respectively. Zero values indicate mixed specialisation. The organic farming variable takes value of one if farms are certified as organic, there is at least one organic product, or there is one process that is carried out with organic methods. On-farm diversification is measured as a share of revenues produced by on-farm diversification activities. Legal form is represented by two dummies indicating if a farm is registered as either an individual holding or a company, which take value of zero in the case of other legal forms. Land is measured as number of hectares of UAA, livestock as number of units, labour as number of Annual Worked Hours (AWH), family work as a share of Annual Work Units (AWU), and, finally, machinery is measured as machine power in terms of number of kilowatts (kW).

Information about the participation of farms in completed OG projects is not available in the FADN data and was retrieved from the national database of OGs that is managed by the Italian NRN. This database is publicly available on the Innovarurale website. It was built on the basis of the European one, in order to share the same information and reduce the workload for those who have to introduce the data, but, unlike the European database, it contains more information such as the details of the partners involved. At the time of this research, the FADN data were available until 2020. Therefore, the projects concluded within 2020 are considered. The relevant typology, objectives, duration, and expected results are reported in Table 1.

The observations available in FADN are represented by different farms observed in few or more years. Since the farms that are present within FADN are subject to be changed over years, the analysis is conducted on pooled data. Outcomes are derived by calculating an average of annual variations of the indicators described above from 2017 to 2020. The period analysed mostly overlaps the one of realisation of the concluded projects, which have a duration of up to 36 months. Including periods prior to 2017 (i.e., 2016, corresponding to the start of some projects) was not possible for issues related to the correct application of the chosen matching ratio, which, in turn, depend on the characteristics of FADN. To remove a possible bias deriving from different periods in which farms are observed, the applied propensity score matching technique is time constrained. More specifi-

Table 1: Typology, objectives, duration, and expected results of the concluded OG projects related to the farms observed in the FADN sample, Emilia-Romagna, Italy.

Project	Typology*	Objectives	Start year	Duration (months)	Expected results**
1	Practice	Application of innovative protection strategies to fruit crops	2016	36	Pesticides (–)
2	Mixed	Application of sustainable techniques and methodologies for protection, irrigation, and nutrition in viticulture	2016	36	Water (–) Pesticides (–)
3	Practice	Improvement of forage systems to support the production of quality cheeses	2016	36	Output (+)
4	Research	Improving the management of soils for the maintenance of organic matter and carbon sequestration	2016	36	Output (+)
5	Practice	Introducing ancient cereals and hemp as a trap crop for the reduction of inputs	2016	36	Crops (+) Fertilisers (–) Pesticides (–)
6	Practice	Introducing innovative products to increase the resistance of plant production to adversities	2016	36	Pesticides (–)
7	Practice	Reducing the consumption of antibiotics in milk production	2016	36	Variable costs (–)
8	Practice	Enhancing by-products of the wine industry to produce energy products, nutraceuticals, and fertilisers	2017	36	Output (+) Fertilisers (–) Energy (–)
9	Practice	Enhancing by-products of vegetable supply chains for food, agronomic and energy purposes	2017	24	Output (+) Fertilisers (–) Energy (–)
10	Practice	Implementing conservation agriculture techniques and bioenergetic buffer strips	2017	24	Water (–) Fertilisers (–) Pesticides (–) Energy (–)
11	Research	Monitoring of the carbon footprint of the fruit sector	2017	36	Output (+) Fertilisers (–)
12	Practice	Reducing ammonia emissions from pig shelters with sewage recovery for soil fertilisation	2017	36	Fertilisers (–)

* "Research" identifies projects that are mainly addressed to monitoring activities and production of methodological guidelines, "practice" refers to projects that involve the application and the experimentation of agricultural innovations in the participating farms, while "mixed" identifies projects that combine research with practical activities.
** A common expected result is an increase in farm income, which may come from an increase in output and/or a decrease in costs.
Source: Authors' elaborations on the national database of OGs

cally, farms of a given year and observed for a given period are only matched with similar observations of the same year and having the same period of observation.

Table 2 shows some descriptive statistics about the sample used. The total number of observations available over the period 2017-2020 amounts to 3204, of which 45 related to farms that participated in OGs. Compared to the average, there are several differences in treated subjects, some of which are particularly evident. It turns out that participants obtain lower revenues, incur lower variable costs, pay less expenditure for the consumption of water, energy, and fuels, and make use of lower quantities of nitrogen, phosphorus, and pesticides per hectare. Moreover, much more than the average, they are younger, have higher levels of education, are more familiar with RDP measures, use organic methods, are formally established as companies, and are specialised in livestock (44% against an average of 18%). Finally, they have on average a far larger number of livestock units (246 against 53), consistently with the prevalent productive specialisation, and there are not treated subjects who are specialised in horticulture.

Table 2: Descriptive statistics about the sample used, Emilia-Romagna, Italy, 2017-2020.

	Treated (n=45)				**All (n=3204)**			
	Mean	**Std. dev.**	**Min**	**Max**	**Mean**	**Std. dev.**	**Min**	**Max**
Monetary outcome indicators								
Output (euro/ha)	7,972	6,404	1,384	41,290	15,152	167,674	360	6,697,958
Variable costs (euro/ha)	3,523	4,262	516	29,147	9,910	142,397	4	5,964,421
Fertilisers (euro/ha)	377	434	0	1,943	357	889	0	36,737
Pesticides (euro/ha)	302	336	0	1,241	356	2,528	0	141,922
Water, energy, and fuels (euro/ha)	169	145	4	723	305	3,005	0	135,196
Farm income (euro/ha)	3,683	2,676	281	11,035	3,946	26,731	–12,493	846,879
Quantitative outcome indicators								
Fertilisers (q/ha)	0.58	1.08	0	4.87	1.17	5.74	0	181.78
Nitrogen (q/ha)	0.009	0.021	0	0.123	0.015	0.060	0	1.457
Phosphorus (q/ha)	0.004	0.009	0	0.049	0.011	0.041	0	0.834
Pesticides (q/ha)	0.01	0.01	0	0.06	0.02	0.21	0	11.08
Crops per farm (no.)	4.20	2.46	1	10	4.28	2.44	0	19
Farm characteristics								
Female	0.09	0.29	0	1	0.11	0.31	0	1
Young (40 years)	0.13	0.34	0	1	0.06	0.24	0	1
With high-medium level education	0.64	0.48	0	1	0.45	0.50	0	1
Accessing to RDP	0.73	0.45	0	1	0.47	0.50	0	1
Located in flat land	0.64	0.48	0	1	0.70	0.46	0	1
Located in hills	0.22	0.42	0	1	0.23	0.42	0	1
With organic production	0.29	0.46	0	1	0.12	0.33	0	1
Individual holding	0.40	0.50	0	1	0.73	0.45	0	1
Company	0.60	0.50	0	1	0.27	0.45	0	1
Specialised in arable	0.11	0.32	0	1	0.34	0.47	0	1
Specialised in horticulture	0.00	0.00	0	0	0.07	0.26	0	1
Specialised in permanent crops	0.36	0.48	0	1	0.30	0.46	0	1
Specialised in livestock	0.44	0.50	0	1	0.18	0.39	0	1
Diversified (share of revenues)	0.03	0.11	0	0.51	0.02	0.11	0	1
Land (ha of UAA)	68.64	75.63	3.12	275.25	37.56	73.02	0.22	1,754.00
Livestock (units)	246.57	669.15	0.00	4,226.00	53.22	275.27	0	8,184.20
Labour (AWH)	8,176.98	10,982.61	1,800.00	72,320.00	4,540.11	7,500.64	900	201,960
Family work (share of AWU)	0.75	0.32	0.11	1.00	0.84	0.26	0	1.00
Machinery (kW)	381.33	626.08	31.00	2,284.00	278.67	315.87	0	4,816

Source: Authors' elaborations on Italian FADN data

Results

Figure 1 shows the distribution of propensity scores calculated for treated and untreated subjects. As can be noted, raw distributions are largely different, and this justifies the use of matching techniques to remove potential sources of bias. After matching, distributions are mostly identical, so demonstrating the effectiveness of the procedure of propensity score matching applied to balance the sample and reduce the selection bias.

Table 3 shows the standardised differences of covariate means between treated and control participants before and after matching. The normalised differences are in almost all cases lower than the same differences calculated before

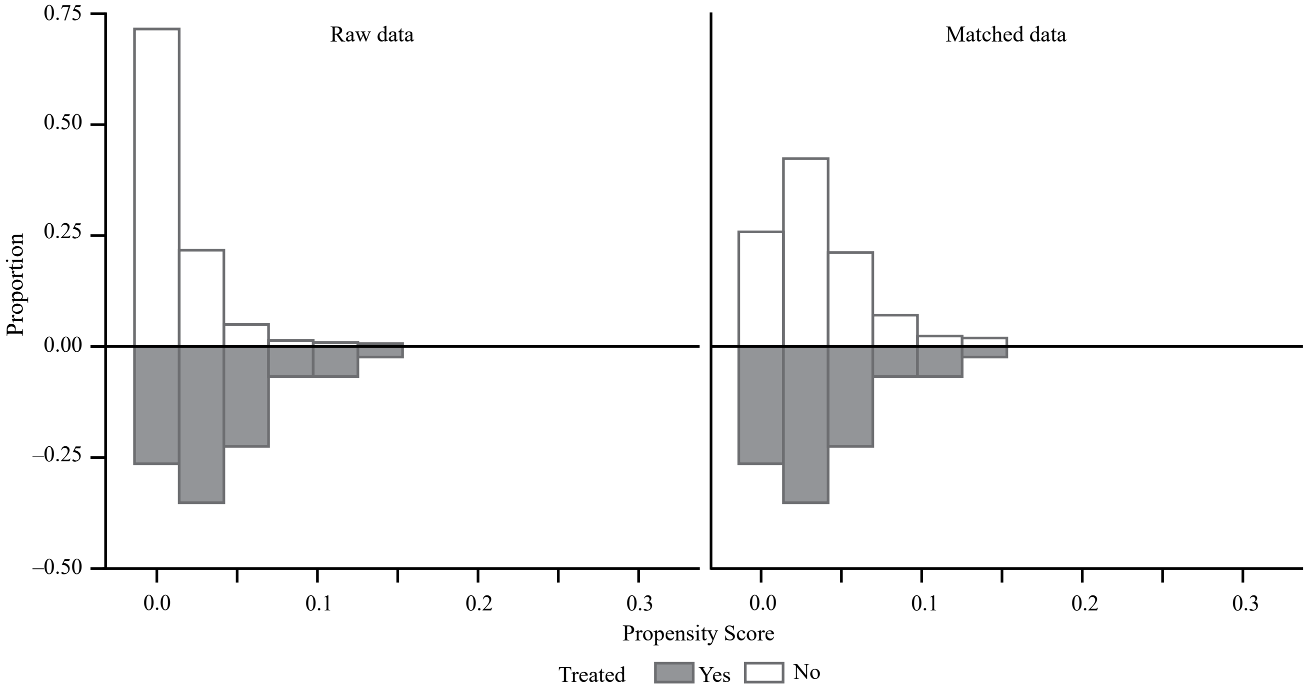

Figure 1: Distribution of the propensity scores calculated for treated and untreated subjects before and after matching.
Source: Authors' elaborations on Italian FADN data

Table 3: Group means and standardised differences of means between treated and untreated subjects before and after propensity score matching.

Variables	All data			Matched data		
	Treated (n=45)	Untreated (n=3204)	Std. diff.	Treated (n=45)	Untreated (n=225)	Std. diff.
Gender (Female = 1; Male = 0)	0.09	0.11	–0.076	0.09	0.07	0.062
Age (Young = 1; Old = 0)	0.13	0.06	0.221	0.13	0.10	0.092
Education (High-medium level = 1; Low level = 0)	0.64	0.44	0.419	0.64	0.63	0.037
Access to RDP (Yes = 1; No = 0)	0.73	0.47	0.602	0.73	0.74	–0.020
Altitude (Mountains = reference)						
Flat	0.64	0.70	–0.124	0.64	0.69	–0.093
Hills	0.22	0.23	–0.030	0.22	0.19	0.086
Typology (Organic = 1; Conventional = 0)	0.29	0.12	0.371	0.29	0.25	0.078
Legal form (Others = reference)						
Individual holding	0.40	0.73	–0.673	0.40	0.43	–0.054
Company	0.60	0.27	0.675	0.60	0.57	0.054
Productive specialisation (Mixed = reference)						
Arable	0.11	0.35	–0.745	0.11	0.07	0.127
Horticulture	0.00	0.07	–0.284	0.00	0.00	0.000
Permanent	0.36	0.30	0.115	0.36	0.39	–0.065
Livestock	0.44	0.18	0.530	0.44	0.47	–0.054
Diversification (share of revenues)	0.03	0.02	0.073	0.03	0.02	0.089
Land (ha of UAA)	68.64	37.12	0.417	68.64	53.91	0.195
Livestock (units)	246.57	50.51	0.293	246.57	170.82	0.113
Labour (AWH)	8,176.98	4,489.03	0.336	8,176.98	7,305.84	0.079
Family work (share of AWU)	0.75	0.84	–0.311	0.75	0.73	0.057
Machinery (kW)	381.33	277.22	0.166	381.33	377.64	0.006

Note: the variables related to time and period of observation, which are used for exact matching, are not shown. The relevant standardised differences are zero after matching.
Source: Authors' elaborations on Italian FADN data

Table 4: % of subjects that experience positive variations of monetary and quantitative outcome indicators and results of logistic regressions for estimating the ATT of participation in EIP-AGRI OGs.

	Matched data		Regression results		
	% Treated	**% Untreated**	**Coefficient**	**Robust std. error**	**Odds ratio**
Monetary outcome indicators					
Output (euro/ha)	64.4	64.0	0.019	0.341	1.020
Variable costs (euro/ha)	46.7	56.4	–0.393	0.337	0.675
Fertilisers (euro/ha)	46.7	59.1	–0.502*	0.305	0.605
Pesticides (euro/ha)	68.9	49.8	0.804**	0.353	2.234
Water, energy, and fuels (euro/ha)	84.4	71.6	0.769**	0.391	2.158
Farm income (euro/ha)	84.4	68.9	0.897**	0.447	2.452
Quantitative outcome indicators					
Fertilisers (q/ha)	53.3	62.2	–0.365	0.308	0.707
Nitrogen (q/ha)	53.3	57.3	–0.162	0.278	0.805
Phosphorus (q/ha)	44.4	59.6	–0.610*	0.337	0.524
Pesticides (q/ha)	62.2	55.1	0.294	0.341	1.342
Crops per farm (no.)	24.4	42.7	–0.833**	0.390	0.435

* Statistically significant at 10%; ** Statistically significant at 5%; *** Statistically significant at 1%.
Source: Authors' elaborations on Italian FADN data

matching. Moreover, they are below the suggested rule of thumb of 0.25 standard deviations (in absolute value). Therefore, these results support the conclusion that the matching procedure performs well, also at level of single covariates, in eliminating possible sources of bias.

Table 4 reports the percentages of treated and untreated subjects that experience positive variations concerning a set of economic and environmental indicators as well as the ATT derived by regressing outcomes on the participation of farmers in OGs.

As regards monetary indicators, the majority of participants is characterised by increases in output and farm income and decreases in variable costs and fertiliser expenditure per hectare. However, 69% and 84% of treated subjects increase expenditure on pesticides and expenditure on water, energy, and fuels, respectively. Control group exhibits outcome variations having similar directions about output, expenditure on water, energy, and fuels, and farm income. The main differences concern pesticide expenditure, which decreases in a half of observations, and variable costs and fertiliser expenditure, which, conversely, increase in 56% and 59% of observations, respectively.

Comparing treated with untreated subjects, from regression analysis it turns out that the coefficient associated with fertiliser expenditure is significant and negative. This means that it is more probable that fertiliser expenditure decreases in farmers participating in OGs. The relevant odd ratio indicates that there is an approximately 40% reduced probability that fertiliser expenditure increases in treated subjects compared to control units.

A further significant coefficient is the one concerning expenditure on pesticides. In this case, the coefficient reveals that the participants in OGs have a larger likelihood to experience an increase in this kind of expenditure in comparison with control units. The corresponding odds ratio indicates that treated subjects have a probability of increasing pesticide expenditure that is 2.2 times the odds of nonparticipants.

The coefficient related to water, energy, and fuels expenditure is also positive and significant. Thus, it is more likely that this expenditure increases in farmers participating in OGs. The probability that water, energy, and fuels expenditure increases is, similarly to pesticide expenditure, 2.2 times higher in treated than in untreated subjects, as the relevant odds ratio shows.

A last significant coefficient among monetary indicators is the one related to farm income. The relevant value suggests a higher probability that farm income increases in treated rather than in untreated units. According to the relevant odds ratio, this probability is 2.5 times higher.

With reference to quantitative indicators, results show that a slightly higher percentage of participants in OGs have increased the overall quantity of fertilisers and the quantity of nitrogen contained in fertilisers, while most participants have used a reduced quantity of phosphorus per hectare. The use of pesticides has increased in 62% of participants and about 75% have decreased the number of crops cultivated. Control group shows more contrasting results. Compared to the participants in OGs, the use of fertilisers and the quantity of nitrogen increase to a larger extent, i.e., in 62% and 57% of observations, respectively. Furthermore, the used quantity of phosphorus increases in a higher share of units (60% against 44%), while the use of pesticides and the number of crops per farm increase in a lower share of subjects, respectively in 55% and 43% of observations.

Looking at regression results, a significant and negative coefficient related to the used quantity of phosphorus per hectare can be observed. Therefore, in treated there is a lower propensity to increase the use of phosphorus. The relevant odds ratio is around 0.5. The probability that the use of phosphorus increases in participants is thus about 50% lower compared to control units.

The coefficient associated with the number of crops per farm is also significant and negative. This implies that in treated subjects there is a higher tendency to decrease the number of crops cultivated. The odds ratio being around 0.4, the probability that the number of crops cultivated increases in treated is therefore about 60% lower compared to control units.

Discussion

Impacts and policy implications

OGs were designed to meet the objectives of increasing productivity and sustainability in agriculture, which, for a farm, could translate into an increase in profitability levels. The results obtained in this study show that OGs may have allowed the participating farms an improvement in fertiliser management that has given rise to decreases in the fertiliser expenditure, a possible substitution of fertilisers with products having environmental lower impact, and increases in income. This could be indicative of the effectiveness of the projects to rationalise the use of fertilisers that fall within the scope of those analysed.

However, these positive impacts are accompanied by negative dynamics that run counter to the environmental objectives of EIP-AGRI in line with what other studies have highlighted (Eckerberg *et al.*, 2023). In fact, the results show, compared to nonparticipants, a higher expenditure on water, energy, and fuels, a greater expenditure on pesticides and a higher increase in the level of specialisation with possible and well-known negative consequences on water quality, health, biodiversity, soil fertility, and climate change. The used quantity of pesticides also increases, although with no significant differences compared to non-participants. These variations are unexpected in consideration of the projects financed, which include those aimed at rationalising the use of water, reducing pesticides, and increasing biodiversity.

A first reason for these results can be the different degree of involvement of participants. Maziliauskas *et al.* (2018) warned that there is the risk that there could be partnerships that are only formal. This implies that not all partners are involved in the same way. The consequence is that any positive impacts will be concentrated only on a part of the farms and that the impact assessment focused on a different sample will not be able to highlight these impacts.

A further reason can relate to the nature of the projects. In the 2014-2020 programming period, several projects providing only feasibility studies and monitoring activities were funded in addition to those intended for actual experimentation and introduction of agricultural innovations. These studies produce contextual analyses depicting the current situation and provide methodological guidelines to lead other farms or the participating farms themselves towards paths of greater sustainability and productivity. Consequently, the effects will only be seen in the future provided that the results of the monitoring are concretely used for the benefit of a more virtuous management and that the guidelines developed are put into practice. However, this could be a great limitation of OGs. Having funded surveys and methodological studies without providing conditions of effective applicability might in fact compromise the effectiveness of OGs and public spending to finance them.

Other factors underlying the results could be linked to the trade-off between objectives and to the selection criteria of the partners. Projects by their nature tend to focus on certain aspects of farm management. This means that all other aspects could be neglected. In this case, the risk is that a farmer that has been selected, for example, to experiment with the use of by-products for energy purposes deriving from the production of arable crops could be able to reduce the consumption of non-renewable energy but could specialise in certain productions (the degree of specialisation measured by the number of cultivated crops therefore increases) and could continue to make extensive use or even increase the use of pesticides and other inputs. This raises issues relating to both the link between the choice of partners and the type of project and the consistency with the objectives of EIP-AGRI during the phases of project preparation and selection. The choice of the partners by the OG, first, and then, the evaluation of the project's fundability by policy makers should in fact consider the characteristics of farmers, the impacts deriving from the current management, and any potential changes resulting from the implementation of the project. In the example given above, a project aimed at reducing potentially harmful inputs to the environment or at introducing new varieties in favour of biodiversity would have been more suitable for that type of farm. It is also true that the calls for selection of OG projects published by the Emilia-Romagna managing authority already included the consistency between the composition of the partnership and the objectives of the project among the evaluation criteria (Regione Emilia-Romagna, 2020). However, this criterion, like others, is not a necessary condition but contributes to the determination of an overall score and is not among the criteria that produce the highest scores. In addition, the criterion is rather generic and susceptible to discretionary evaluations based on the statements provided by those presenting the project.

There is therefore the need to revise governance processes to improve the effectiveness of OGs as other studies have stressed (Giarè and Vagnozzi, 2021). Based on the considerations made above, the managing authorities of RDPs are first called to be more selective by excluding projects that do not explicitly provide for experimentation and the introduction of innovations. This means that projects including only feasibility studies and monitoring activities should be rejected. Furthermore, the managing authorities should require farmers, during the project proposal presentation phase, to clearly indicate the management situation by providing quantitative and verifiable data on the current impacts to allow the evaluation of the coherence between the project objectives and farm characteristics and, therefore, the opportunity to admit that farm into the partnership. In addition to the indicators of the current management situation, participants should also be required to quantify the results achieved, as part of the necessary and constant monitoring of activities (Maziliauskas *et al.*, 2018). This enables both the OG and policy evaluators to calculate variations and thus measure the effects deriving from the application of innovations. Knowledge of the impacts, which could be checked on a sample basis with on-site checks, would not only help to improve the effectiveness and orient the future setting of OGs but could also be a reason for reducing public contributions in the event of unjustifiable results and in contrast with the initial objectives. The provision of possible penalties associated with results could in turn act as a strong incentive for OGs to be more tailored and selective during its constitution by presenting projects and forming partnerships that are more involved and more consistent with the aims of EIP-AGRI.

The policy framework that is proposed here responds to the principles of the performance-based approach adopted by the 2023-2027 CAP. This approach, also called New Delivery Model, gives more emphasis to policy performance compared to the previous programming period. Basically, it provides for the verification of the level of achievement of predefined target indicators at level of Member States, the requirement of an action plan in the event of excessive discrepancies between targets and realisations and the suspension of payments if the action plan is not submitted or manifestly insufficient (art. 128–129 of Regulation EU No. 2021/2115).

However, the approach suggested here presents four main differences. First, it would be applied at level of single projects. Second, two list of indicators could be drawn up according to the implementation phase of the project. A first one can be broader and consider different aspects of farm management. i.e., economic, social, and environmental aspects. These indicators can be used for initial selection. In fact, their value calculated at an early stage for potential farms applying to participate in the project could be compared with those of farms having similar characteristics. The aim is to measure the impact of farm management, relatively to the competitive context in which farms operate, and to evaluate the real need for innovation and the opportunity to include them within the partnership. This is because marked differences with the comparison group could signal management criticalities that can be resolved through the application of economic, environmental, or social innovations. In this regard, the FADN data could be effectively used to identify a battery of possible indicators and make comparisons as Arzeni *et al.* (2021) showed.

A second list could contain a selection of all indicators initially identified and based on the type of project. These indicators would be employed after the project has been approved for monitoring and final assessment. For instance, in the case of projects aimed at reducing the used quantity of water, indicators such as the incidence of both the amount of water used and the expenditure for water consumption per hectare could be monitored. Third, the action plan is represented here by all the corrective actions that the OG undertakes during the implementation of the project following the constant monitoring activities in order to reduce the gap between objectives and results. Fourth, penalties are applied once the project is completed under two hypotheses. One occurs if the plan is not implemented as established. This situation was already contemplated by the managing authorities of RDP. The other circumstance would occur if the opposite effects were produced with respect to the initial objectives. In the example above, they would be applied if the ratio of used quantity of water to hectares increased rather than decreased. This is to avoid the application of sanctions in situations where innovations, even if correctly applied according to the plan, are neutral, i.e., they do not produce significant effects as expected because of external and unpredicted factors.

The official guidelines for measuring the progress of the OGs financed under the 2023-2027 CAP substantially confirm the previous ones (European Commission, 2016). The main focus is on the need to classify projects rather than improve their performance (Annex VI of Commission Implementing Regulation EU 2022/1475). The risk, therefore, is that the distortions highlighted by this study will not only be removed but even exacerbated. However, thanks to the greater flexibility attributed by the reformed CAP at national level, Member States can decide to integrate the current monitoring and controlling system, in compliance with the general principles, in order to increase the effectiveness of OGs. The framework proposed here could be a possible option in this direction.

Data implications

The results of this study may be influenced by the data used. A first source of influence can be the size of the sample analysed. This study focuses in fact on a regional case, the Emilia-Romagna region, and on a small percentage of farmers that participated in concluded OGs (about 6%). This depends on the characteristics of FADN, which collects data from a representative but still limited sample, and on the fact that there are several OGs still not concluded or that are not officially concluded.

Another source of influence concerns the construction of the sample and the methods of calculating outcomes. The farmers analysed are observed during the treatment, i.e., during the implementation period of the OG projects. Due to a planned turnover of the units observed, FADN does not always allow for each farm an analysis of the periods preceding and following that in which the project was implemented. One of the requirements of propensity score matching techniques is that the treatment should not influence the confounders analysed, otherwise endogeneity problems could arise. This is what can happen if variables are measured during treatment. In this regard, one of the assumptions of this study is that an inverse relationship between treatment and confounders does not exist or is so weak that it does not affect the results. This assumption can be considered plausible as the OG projects have a limited scope and do not alter the main characteristics of farms, which form the basis for the construction of the control group. A further assumption underlying this study is that the potential effects measured in terms of impact direction occur during the implementation of the project and can therefore already be measured without waiting for a certain period to pass from the end of the project. This assumption can also be reasonably accepted in all those cases in which a practical application of innovations is provided and considering that the last period of the project is generally dedicated to the dissemination of results, while practical activities of experimentation and application of agricultural innovations are carried out in the initial and especially in the intermediate phases.

It is evident that a more accurate analysis of the impacts produced by OGs on farmers will be possible as soon as all projects are completed. Nevertheless, this does not affect the usefulness of this study, which, in addition to representing a first attempt to analyse the impact of a sample of OGs in a given regional context for the benefit of the programming that has just begun, provides some useful practical and methodological indications for setting future analyses. Nor can the usefulness of the FADN data be called into question. This is because FADN, besides making a large amount of information available about a sample of farms that is representative of regional (and national) agriculture, makes it possible to compare the farms participating in OGs with similar farms that did not take part in them, without having to resort to further and costly investigations.

Concluding remarks

This paper analysed the possible impacts produced by a sample of OGs of EIP-AGRI on the economic and environmental performance of the participating farms. The focus was on the Italian region of Emilia-Romagna. This is an ideal laboratory to test the effectiveness of OGs because it is the region with the highest number of financed OGs and for the prevailing characteristics of regional agriculture that make innovations capable of reducing environmental impacts highly desirable.

This study compared a group of farmers participating in OGs with one of nonparticipating farms having very similar characteristics and selected by using a propensity score matching technique applied to the FADN data. This research reveals conflicting results. On the one hand, it turns out that there are possible improvements in fertiliser management and in profitability levels induced by participation in OGs, while, on the other hand, there emerge an increased consumption of water, energy, and fuels, an increased use of pesticides, and a greater loss of crop diversity. Overall, it can therefore be argued that the OGs analysed may have provided the participating farmers with a comparative economic advantage. This is a strong incentive for participation and therefore represents an encouraging result for the future of this political instrument. However, those OGs may not have achieved some of the most important objectives related to environmental sustainability. Possible factors that could be at the origin of this result are the financing of projects that only provide for feasibility studies, a different degree of involvement of farmers, and the selection criteria of OGs, which may be not very effective in ensuring full consistency between the characteristics of agricultural partners, the type of project, and the objectives of EIP-AGRI. For these reasons, policy makers are advised to condition projects on the actual experimentation and implementation of agricultural innovations and introduce a performance-based system of indicators for the assessment of the ex-ante and ex-post impacts of farm management.

Acknowledgements

This research was supported by the Italian National Rural Network Programme 2014-2020 (CREA Project Sheet 25.1 – Support for the development of AKIS).

References

Abadie, A. and Spiess, J. (2022): Robust Post-Matching Inference. Journal of the American Statistical Association, **117** (538), 983–995. https://doi.org/10.1080/01621459.2020.1840383

Altieri, M.A. (1999): The ecological role of biodiversity in agroecosystems. Agriculture, Ecosystems & Environment, **74** (1–3), 19–31. https://doi.org/10.1016/S0167-8809(99)00028-6

Arzeni, A., Ascione, E., Borsotto, P., Carta, V., Castellotti, T. and Vagnozzi, A. (2021): Analysis of farms characteristics related to innovation needs: a proposal for supporting the public decision-making process. Land Use Policy, **100**, 104892. https://doi.org/10.1016/j.landusepol.2020.104892

Austin, P.C. (2010): Statistical Criteria for Selecting the Optimal Number of Untreated Subjects Matched to Each Treated Subject When Using Many-to-One Matching on the Propensity Score. American Journal of Epidemiology, **172** (9), 1092–1097. https://doi.org/10.1093/aje/kwq224

Austin, P.C. (2011): An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behavioral Research, **46** (3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Austin, P.C., Grootendorst, P. and Anderson, G.M. (2007): A comparison of the ability of different propensity score models to balance measured variables between treated and untreated subjects: A Monte Carlo study. Statistics in Medicine, **26** (4), 734–753. https://doi.org/10.1002/sim.2580

Brookhart, M.A., Schneeweiss, S., Rothman, K.J., Glynn, R.J., Avorn, J. and Stürmer, T. (2006): Variable selection for propensity score models. American Journal of Epidemiology, **163** (12), 1149–1156. https://doi.org/10.1093/aje/kwj149

Caliendo, M. and Kopeinig, S. (2008): Some practical guidance for the implementation of propensity score matching. Journal of Economic Surveys, **22** (1), 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Cenzer, I., Boscardin, W.J. and Berger, K. (2020): Performance of matching methods in studies of rare diseases: a simulation study. Intractable & Rare Diseases Research, **9** (2), 79–88. https://doi.org/10.5582/irdr.2020.01016

Cisilino, F., Bodini, A. and Zanoli, A. (2019): Rural development programs' impact on environment: An ex-post evaluation of organic farming. Land Use Policy, **85**, 454–462. https://doi.org/10.1016/j.landusepol.2019.04.016

Cristiano, S. and Proietti, P. (2018): Do EIP interactive innovation approaches interact each other? International Journal of Agricultural Extension, **6** (3), 53–63.

D'Agostino, R.B. (1998): Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Statistics in Medicine, **17** (19), 2265–2281. https://doi.org/10.1002/(SICI)1097-0258(19981015)17:19<2265::AID-SIM918>3.0.CO;2-B

Eckerberg, K., Bjärstig, T. and Miljand, M. (2023): Steering 'green' innovation policy toward sustainability? Lessons from implementing EIP-AGRI in Sweden. Environmental Innovation and Societal Transitions, **48**, 100732. https://doi.org/10.1016/j.eist.2023.100732

European Commission (2014): Commission implementing regulation (EU) No 808/2014 of 17 July 2014 laying down rules for the application of Regulation (EU) No 1305/2013 on support for rural development by the European Agricultural Fund for Rural Development (EAFRD). Brussels, Belgium.

European Commission (2016): EIP-AGRI Common format for interactive innovation projects. Brussels, Belgium.

European Commission (2019): Building stronger agricultural knowledge and innovation systems (AKIS) to foster advice, knowledge and innovation in agriculture and rural areas. Brussels, Belgium.

Gehrlein, U. and von Kutzleben, N. (2016): First Experiences in Implementation and Evaluation of the EIP Approach in two Federal States of Germany. In Proceedings of the 12th European IFSA Symposium 12-15 July 2016. UK: Harper Adams University, 1–14.

Giarè, F. and 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, **76** (3), 41–52. https://doi.org/10.36253/rea-13206

Guo, S., Fraser, M. and Chen, Q. (2020): Propensity Score Analysis: Recent Debate and Discussion. Journal of the Society for Social Work and Research, **11** (3), 463–482. https://doi.org/10.1086/711393

Harrahill, K., Macken-Walsh, Á., O'Neill, E. and Lennon, M. (2022): An Analysis of Irish Dairy Farmers' Participation in the Bioeconomy: Exploring Power and Knowledge Dynamics in a Multi-actor EIP-AGRI Operational Group. Sustainability, **14** (19), 12098. https://doi.org/10.3390/su141912098

Ibáñez-Jiménez, Á., Jiménez-Olivencia, Y., Mesa-Pedrazas, Á., Porcel-Rodríguez, L. and Zimmerer, K. (2022): A Systematic Review of EU-Funded Innovative Agri-Food Projects: Potential for Transfer between Territories. Land, **11** (4), 519. https://doi.org/10.3390/land11040519

Imbens, G.W. (2004): Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. Review of Economics and Statistics, **86** (1), 4–29. https://doi.org/10.1162/003465304323023651

Imbens, G.W. and Wooldridge, J.M. (2009): Recent Developments in the Econometrics of Program Evaluation. Journal of Economic Literature, **47** (1), 5–86. https://doi.org/10.1257/jel.47.1.5

Knotter, S., Kretz, D. and Zeqo, K. (2019): Operational Groups Assessment 2018. Bruxelles: IDEA Consult nv.

Lin, B.B. (2011): Resilience in Agriculture through Crop Diversification: Adaptive Management for Environmental Change. BioScience, **61** (3), 183–193. https://doi.org/10.1525/bio.2011.61.3.4

Liu, X., Herbert, S.J., Hashemi, A.M., Zhang, X. and Ding, G. (2006): Effects of agricultural management on soil organic matter and carbon transformation – a review. Plant, Soil and Environment, **52** (12), 531–543. https://doi.org/10.17221/3544-PSE

Maziliauskas, A., Baranauskienė, J. and Pakeltienė, R. (2018): Factors of effectiveness of European innovation partnership in agriculture. Management Theory and Studies for Rural Business and Infrastructure Development, **40** (2), 216–231. https://doi.org/10.15544/mts.2018.21

McCarthy, J., Meredith, D. and Bonnin, C. (2021): Actor motivations to engage with collaborative agri-environmental policy: An assemblage based exploration. Journal of Rural Studies, **87**, 88–98. https://doi.org/10.1016/j.jrurstud.2021.08.025

Menta, C., Bonati, B., Staffilani, F. and Conti, F.D. (2017): Agriculture Management and Soil Fauna Monitoring: The Case of Emilia-Romagna Region (Italy). Agricultural Research & Technology: Open Access Journal, **4** (5), 86–89. https://doi.org/10.19080/ARTOAJ.2017.04.555649

Molina, N., Brunori, G., Favilli, E., Grando, S. and Proietti, P. (2021): Farmers' participation in operational groups to foster innovation in the agricultural sector: An Italian case study. Sustainability, **13**, (10), 1–27. https://doi.org/10.3390/su13105605

OECD (2001): Measuring Productivity - OECD Manual: Measurement of Aggregate and Industry-level Productivity Growth. Paris: OECD.

Parzonko, A., Wawrzyniak, S. and Krzyżanowska, K. (2022): The role of the innovation broker in the formation of EIP-AGRI operational groups. Annals of the Polish Association of Agricultural and Agribusiness Economists, **24** (1), 194–208. https://doi.org/10.5604/01.3001.0015.7252

Piñeiro, V., Nieto-alemán, P., Marín-Corbí, J. and Garcia-Alvarez-Coque, J.-M. (2021): Collaboration through EIP-AGRI Operational Groups and their role as innovation intermediaries. New Medit, **20** (3), 17–32. https://doi.org/10.30682/nm2103b

Pretty, J. (2008): Agricultural sustainability: concepts, principles and evidence. Philosophical Transactions of the Royal Society B: Biological Sciences, **363** (1491), 447–465. https://doi.org/10.1098/rstb.2007.2163

Purvis, B., Mao, Y. and Robinson, D. (2019): Three pillars of sustainability: in search of conceptual origins. Sustainability Science, **14** (3), 681–695. https://doi.org/10.1007/s11625-018-0627-5

Rassen, J.A., Shelat, A.A., Myers, J., Glynn, R.J., Rothman, K.J. and Schneeweiss, S. (2012): One-to-many propensity score matching in cohort studies. Pharmacoepidemiology and Drug Safety, **21**, 69–80. https://doi.org/10.1002/pds.3263

Regione Emilia-Romagna (2020): Reg. CE n. 1305/2013 - Avviso pubblico per l'attuazione del tipo di operazione 16.1.0 - Sostegno per la costituzione e la gestione dei Gruppi Operativi del PEI in materia di produttività e sostenibilità dell'agricoltura. https://bur.regione.emilia-romagna.it/dettaglio-inserzione?i=1ab25951a47e4ddc9a4ad20b438df6e1

Rosenbaum, P.R. (1989): Optimal Matching for Observational Studies. Journal of the American Statistical Association, **84** (408), 1024–1032. https://doi.org/10.1080/01621459.1989.10478868

Rosenbaum, P.R. and Rubin, D.B. (1983): The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika, **70** (1), 41–55. https://doi.org/10.2307/2335942

Rubin, D.B. (1973): Matching to Remove Bias in Observational Studies. Biometrics, **29** (1), 159–183. https://doi.org/10.2307/2529684

Schreuder, R., Peratoner, G., Goliński, P. and Van den Pol-van Dasselaar, A. (2022): EIP-AGRI: EU initiatives for the transfer and co-creation of innovations on and for grassland. In Delaby, L., Baumont, R., Brocard, V., Lemauviel-Lavenant, S., Plantureux, S., Vertès, F. and Peyraud, J.L. (eds.), Grassland at the Heart of Circular and Sustainable Food Systems. Paris, France: The Organising Committee of the 29th General Meeting of the European Grassland Federation, INRAE, 767–780.

Stone, C.A. and Tang, Y. (2013): Comparing propensity score methods in balancing covariates and recovering impact in small sample educational program evaluations. Practical Assessment, Research and Evaluation, **18** (13), 2–12. https://doi.org/10.7275/qkqa-9k50

Stuart, E.A. (2010): Matching Methods for Causal Inference: A Review and a Look Forward. Statistical Science, **25** (1), 754–768. https://doi.org/10.1214/09-STS313

Van Oost, I. and Vagnozzi, A. (2021): Knowledge and innovation, privileged tools of the agro-food system transition towards full sustainability. Italian Review of Agricultural Economics, **75** (3), 33–37. https://doi.org/10.13128/REA-12707

Velten, S., Leventon, J., Jager, N. and Newig, J. (2015): What is sustainable agriculture? A systematic review. Sustainability, **7** (6), 7833–7865. https://doi.org/10.3390/su7067833