Exploring Agri-environmental effectiveness using counterfactual analysis

Cisilino[1] F., Bodini[1] A., Zanoli[2] A., Lasorella[1] M.V.

[1]*CREA, Council for Agricultural Research and Economics, Research Centre for Agricultural Policies and Bio-economy*
federica.cisilino@crea.gov.it; antonella.bodini@crea.gov.it; mvalentina.lasorella@crea.gov.it
[2]*ISTAT, Italian National Institute of Statistics agostina.zanoli@istat.it*

Paper prepared for presentation for the 162nd EAAE Seminar
The evaluation of new CAP instruments: Lessons learned and the road ahead

April 26-27, 2018
Corvinus University of Budapest
Budapest, Hungary

Abstract
The CAP has introduced a higher allocation of the budget to the Agri-environmental schemes and Organic Farming over the years. In order to estimating impacts at regional level the evaluation of a Rural Development Programme single sub-measure have been considered. A FADN panel sample of Marche Region has been selected for years 2013, 2015. The application of a non-parametric matching model and the Difference-in-Difference estimator is presented.
The environmental performance of organic farms is statistically different from conventional ones, therefore policy intervention has reached a positive effect. Instead, results on detected economic variables shown no differences between the two groups.

Keywords: Statistical Matching, Difference-in-Difference, Environmental Farm Performance, Rural Development Policies, Organic Farming

1. Introduction

In recent years, the agricultural sector in the European Union is facing new challenges. The increasing attention to the relationship between agriculture and the environment and the rise in price volatility on agricultural markets has led to a new emphasis on Agri-environmental policies as well as to a search for new risk management strategies for the farmer. These trends are strengthened by the evolution of the Common Agricultural Policy (CAP) towards a higher allocation of the EU budget to the Agri-environmental schemes (AESs) and Organic Farming over the years and towards a liberalization of agricultural markets in the EU. Nowadays, EU farmers are required to combine the provision of environmental benefits with the production of agricultural commodities, whose market price is hardly predictable. In this evolving contest, this study draws attention to the Second Pillar Policies' effectiveness, taking into account 2007-2013 programming period using an ex-post evaluation method. The scientific debate on Agricultural Policy evaluation methods has developed broadly. The European Commission has put a lot of effort on monitoring the results of Common Agricultural Policy, as European funds represent the main source of grating subsidies to Agriculture in most of the European countries. Conventionally, quantitative evaluation methods are applied in labour and business analysis; nevertheless recently they have been increasingly used to assess the Rural Development Programmes (RDP). The logic of public intervention in this case is essentially based on the concept of compensating for a particular disadvantage for farms that are not in a position to realize agri-environmental projects, but which would have the capacity to develop them - if in the appropriate economic conditions - generating a positive impact not only for the farm itself, but also for the economy as a whole (Sisti, 2007). Farms receiving public subsidies are those coming from a two steps selection process (figure 1): a self-selection stage (farms that apply for funds have a certain ability to carry out projects) which involve a selection bias, and a second step related to worthiness (farms meet the eligibility requirements established by the public administration) (Cagliero et al. 2010). The evaluation of rural development programmes comes out from single measures' evaluation, comparing changes observed in treated and non-treated using different tools and approaches both quantitative and qualitative (for instance indicators monitored over time or using direct surveys data). But only when the concept of causality between hypothesis and empirical evidence is introduced (causal nexus), between intervention and change in specific farm performance, then the counterfactual analysis is turned on. The approach is the so-called counterfactual paradigm and it is based on the following assumption: the effect of a policy comes as an evidence from the difference between what it is observed when funding occur and what it is observed during the same period and for the same subjects, in the absence of subsidies (Smith and Todd, 2005; Martini, 2006). To apply quantitative methods researchers need a wide range of data and reliable sources of information and this is crucial but not so easy to get when the agricultural sector is under study (Rettore *et al.*, 2002). The purpose of this paper is to identify a counterfactual group within the Farm Accountancy Data Network (FADN) sample of an Italian region and applying the Statistical Matching Method and the DiD estimator.

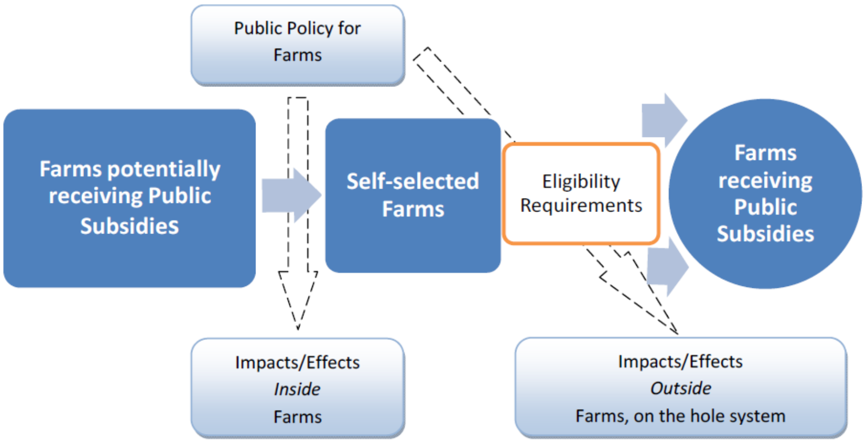

Figure 1 - The Logic of Public Policy when farms are involved.
Source: authors' processing.

As far as the RDPs are concernedthe main problem arising is linked to the fact that it is difficult to distinguish objectively the changes brought about by the programme, which take place in the time called for in programming, and those related to other exogenous factors. Identifying the net contribution of a particular measure is the cognitive goal of the counterfactual approach. However, this faces at least two types of problems: the first involves the changes recorded over time and the difficulty in isolating the components due to subsidies from other independent exogenous variables. The second issue is the definition of counterfactual itself and the inability to observe and therefore to measure an effect on subjects which cannot be both beneficiaries and non-beneficiaries of a given policy. Since a direct effect is not observable it is necessary to estimate it through experimental and non-experimental methods based on the building of a control group. The present paper addresses these issues by applying a non-parametric Logit model through the Propensity Score Matching approach (PSM), using a Difference-in-Difference estimator (DiD). The PSM analysis has been applied in the last three decades as a programme's evaluation methods and recently applied also to environmental issues (Pufahl and Weiss, 2009; Chabé-Ferret and Subervie, 2011; Arata and Sckokai 2016).
The analysis of this paper is focused on the effects of Organic Agriculture subsidies, a sub-measure of Marche region Agri-Environmental Schemes (AES-OF) for 2007-2013 period, aiming at evaluating its impact among treated and non-treated. In order to measure group's differences a panel sample with FADN has been selected for years 2013 and 2015.

2. Background

In the European Union (EU), the Common Agricultural Policy (CAP) has introduced environmental measures in order to reduce the impact of negative environmental externalities and to promote positive externalities of agricultural activities. For example, negative externalities are sanctioned through a reduction in direct income payments if cross-compliance is not respected, while some ad hoc directives have been implemented for addressing some specific problems (i.e. the nitrates directive). Positive externalities are encouraged by some Rural Development measures which promote environmentally sustainable farming practices, through payments that compensate farmers for the provision of environmental goods that the market does not reward. One of these measures are the Agri-environmental schemes (AESs) introduced in the late 1980s as an option to be applied by Member States. Since 1992, with the Mac Sharry reform, the development of Agri-Environmental Programmes has become compulsory for all Member States with the Regulation 2078/92, while their application by farmers was still voluntary. Since 1999, with the Agenda 2000 CAP reform (EEC n. 1257/99) the AESs have become a section of the Rural Development Programs (RDPs) which represents Pillar II of the CAP. Agri-environmental contracts are voluntary contracts, of at least 5 years, stipulated between the farmer and the government; under these contracts, the farmer provides environmental goods that go beyond the minimum requirements of the cross compliance and of the European and national compulsory environmental regulations and they receive a fixed per hectare payment to face the additional costs and the loss of income linked to these commitments. The main objective of the AESs consists of reducing the agricultural pollution risks as well as protecting biodiversity and landscape. AESs payments are co-financed by Member States and they represent a large share of the public budget for Rural Development. They can be designed at national, regional or local level and this allows to take into account the heterogeneity of the natural characteristics and agricultural systems throughout the EU Member States. The Agri-environmental and animal welfare Programmes are the only measures that are compulsory in all RDPs and the AESs are the most significant measure in term of EU funding for Rural Development (23.6%), followed by 'Modernization of agricultural holdings' (12%) and 'Payments to farmers in areas with development disadvantages, marginal areas other than mountain areas(7.6%). In some EU countries the AESs are mainly defined at national level with little decision power at regional level (e.g. France and the Netherlands), while in other countries they are defined and implemented at regional or sub regional level (e.g. Italy, Germany, Spain). There is quite a lot of literature about AESs; most of it tries to analyze the factors affecting farmer's participation to Agri-environmental contracts (Vanslembrouck *et al.*, 2002; Defrancesco *et al.*, 2007). Another widely studied topic concerns the analysis of the environmental effectiveness of farmers' environmental practices; most of these studies outline the importance of accounting for farm heterogeneity, by applying more farm-specific measures (Aakkula *et al.*, 2011). A few studies analyse the effects of AESs on farms' practices and economic results. Sauer *et al.* (2012) argued that "only a few studies so far have attempted to empirically measure the actual impact of being subject to AESs on producer behavior at individual farm level using statistical or econometric tools". The expectation is that farmers are heavily affected by participation to AESs, which may lead to a deep reorganization of the farm and to a change in the sources of income. The Propensity Score matching has been used as ex-post analysis tool to analyse the effects of agricultural policy measures on farm's performances and have been widely developed in the last thirty years as a program evaluation method based on observational data in a broad range of disciplines, such as medicine, epidemiology, psychology, social sciences, education. More recently it has been applied also to environmental economics and to the analysis of some measures of RDPs. In particular other researches have applied PS matching to analyze farm performances related to participation to AESs at member state level (Arata and Sckokai, 2016; Pufahl and Weiss, 2009).

3. Materials and Methods

3.1. The area under study

Marche is a rural region located in the Central-Eastern part of Italy and its extension represents 3.2% of the national territory. It is a mainly rural region and it is characterized by the absence of large urban centres. From the morphology point of view the territory under study is mostly hilly and mountainous with high likelihood of hydrogeological instability. The nitrate vulnerable zones are about 11% of territory and 21% of total utilized agricultural area (UAA), and are located along river courses and the regional coastal strip (Regione Marche, 2015).

According to the latest agricultural census data, the number of agricultural holdings in Marche region account for 44.000 units, corresponding to 3% of total Italian farms and cover almost 471 thousand hectares, which equal 60% of regional land. Among them 4% are organic farms on a surface of 52 thousand hectares, equivalent to 11% of regional UAA whereas on average in Italy the UAA for organic production is 4% only. From the labour force point of view organic farms employ 10% of the working days over the total of agricultural working days, this results is above the national average which is around 6%.

The regional agriculture is characterized by arable crops, Mediterranean crops such as vineyards, olive groves and orchards, and sheep breeding. This specialization reflects also in organic farming of the region.

The regional intervention on Axis II Rural Development Programmes 2007-2013 is of 69 million of euros. Out of it Organic Farming account for65 Million of euros, in the equivalent to 95% of the total financial availability of regional EAFRD. In terms of beneficiaries (treated) the number of holding participating are roughly 15.000, of which 95% are organic, whereas in terms of territorial extension the surface interested by Organic Framing (OF) beneficiaries covers 96% of the UAA interested by Agri-Environmental Schemes (AES) over all. Agri-environmental programs account for about one fourth of total regional rural development expenditure, highlighting the huge amount of financial resources allocated within the RDP and CAP framework to promote sustainable production methods from both economic and environmental point of view (figure 2).

Therefore, Marche region is one of the most interesting Italian region to analyze as far as Organic Farming is concerned.

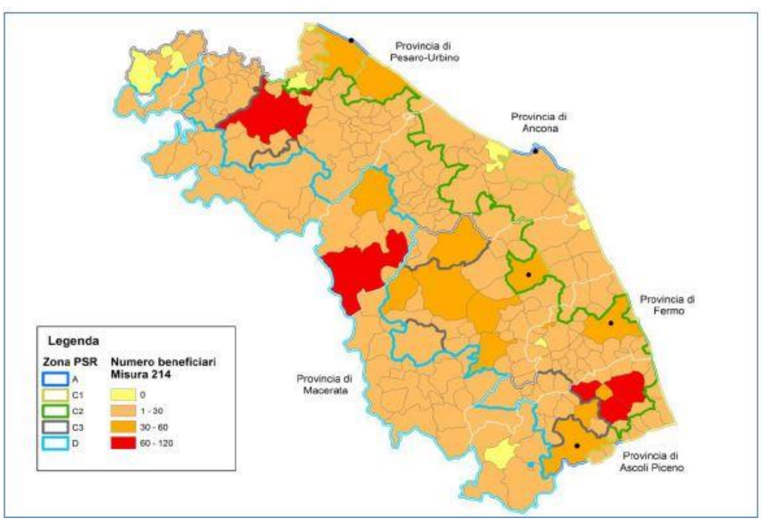

Figure 2 – Marche RDP 2007-2013: Regional distribution of treated farms
Source: Lattanzio Evaluation Report (2016)

The evaluation of RDP sub-measure AES-OF (Agri-Environmental-Scheme-Organic Farming sub-measure) was developed specifically to allow assessment of differences in the environmental and economic performance of farms according to their participation in AES-OF in particular. The evaluation question "does the AES-OF sub-measure work?" has been investigated by the identification of a set of outcomes variables. Two are the main issues to be addressed:

1) Economic impact on Farm performance: Are Organic Farming convergent to Conventional Farming after the implementation of AES-OF (a priori expectation)?

2) Environmental Impact on soil fertility and water determinants: to what extent is Organic Farming performing better than Conventional Farming (a priori expectation)?

3.2. The Counterfactual Analysis: evaluating the impact of AES-OF with DiD approach

The debate on methods for evaluating public policies is always open and lively. The main goal is how to assess the impact of a given policy and how to measure its effectiveness and efficiency, or how to establish if and how much it has helped in changing a particular situation facing new challenges with a positive outcome. For this reasons it becomes crucial to produce an estimate of what would have been achieved in the absence of intervention (Smith and Todd, 2005). The counterfactual approach is based on the effect of a policy/measure as the difference between what it is empirically observable after the policy implementation and what we would have observed, in the same period and for the same subjects, in the absence of intervention (Martini *et al.*, 2006). In order to estimating impacts and analysing the causal link determined by the implementation of a given policy and identifying its net contribution, it would be necessary to observe a group of individuals who are both treated and non-treated. What is missing is the information related to those treated if they had not been treated (counterfactual). The effect is not observable in the reality (difference between a real value and a hypothetical one), but this can be estimated by substituting the counterfactual value with an empirical and credible estimate (Martini *et al.*, 2006; Martini Sisti, 2010). One of the main issue is the selection bias. This happens when the access to a policy intervention is not random. The simple difference does not allow to obtain an unbiased estimate of the policy effect, because of the likely relationship between the outcome and selection into treatment. Thus, it is crucial to take into account sources of selection biases when estimating ex-post policy impacts.

When estimating a policy impact we shall consider the difficulty to distinguishing changes caused by the programme by those attributable to exogenous factors. The literature on these issues is well-established, and the number of applications to RDP's based on applied statistical methods is increasing in the last years. Within member states reports on monitoring systems some evidence are periodically issued by the Advisory Public Services in order to assess the level of implementation of the regional RDP, by comparing results between those expected and achieved.

Since this paper aims at evaluating Agri-environmental Rural Development Policies, providing an ex-post evaluation of the effects coming from a regional RDP, the choice made within this work is the application of a Propensity Score Matching (PSM) model and the Difference-in-Difference estimator (DiD) to account for potential disparities in the starting conditions of the subjects and potential "spontaneous" trends in the outcome variables which are independent from the policy intervention. Matching estimators work by matching each treated individual with one or more non-treated individuals that have similar observed characteristics, the covariate X, and interpreting the difference in their outcomes as the effect of the treatment (Smith and Todd, 2005). Furthermore, the PS matching estimator is based on conditioning the Probability P(X) of being treated such that the conditional distribution of X - given P(X) - is independent of the

treatment assignment (Rosembaum and Rubin,1984) The first step is the implementation of a Logit/Probit model that allows to define the so-called Propensity Score (PS) as a basis to discriminate units excluded from the programme (control group). The basic idea of the matching is to combine each treated with one or morenon-treated as similar as possible. The pairing (between treated and non-treated) is based on the PS, which is the conditional probability that a farm is treated - given its observable pre-treatment characteristics (covariates).
The carried out analysis was developed under the following a priori assumptions:

- a common support was considered: the values given by the control variables of treated and non-treated are similar (if this is not the case, those units were excluded from the analysis);
- the property of balance was respected: the units having the same PS should have the same distribution in the observable characteristics, regardless of treatment. This assumption ensures that, given the PS, the exposure to the treatment is random. The selected non-treated units are defined as the control group, and so they represent the situation that would be observed for treated in the absence of intervention. This property is verified by stratifying all units on the basis of PS values and then applying some tests on means of treated and non-treated within each stratum. The property of balance is important to understand the goodness of the model used. If not satisfied, a different model specification should be used.

The most common parameter that give evidence of the difference in the mean outcomes of the treated and the matched control group is the Average Treatment effect on Treated (ATT).

$$ATT = \left\{E\left(Y^1|D = 1, P(X)\right) - E\left(Y^0|D = 0, P(X)\right)\right\} \quad (1)$$

Where Y^1 is the is the value of the outcome variable for a unit in presence of treatment; whereas Y^0 is the value of outcome variable for the same unit without treatment. A dummy variable D denote the involvement (D=1) or the non-involvement (T=0) to the program (1). When data about the outcome variables are available for both the treated and non-treated individuals, both before and after the interventions, it is possible to apply to treated and matched group the *Differences-in-Differences* (DID) method, taking into account both the ex-post and ex-ante differences between the two groups (Smith and Todd, 2005)[1].
Combining the two approaches the counterfactual is constructed matching treated and non-treated by the propensity score, and then the twofold difference is performed among treated and untreated and before and after treatment.

In this case the average effect of the treatment across the treated individuals (ATT) will be:

$$ATT = \left\{ E(Y_t^1 - Y_s^0 \mid D = 1, P(X)) - E(Y_t^0 - Y_s^0 \mid D = 0, P(X)) \right\} \quad (2)$$

[1] The *Differences-in-Differences* approach measures the effect of the policies between treated and non-treated, before and after measuring the difference in outcome by excluding common time trends.

Where *t* is a time period *after* the programme's starting date while *s* is a time period *before* the implementation of the programme (2).

Summarizing, these are the main methodological steps followed to develop the analysis :

i. Identifying the treated group: a panel dataset has been defined (2013 and 2015) using the Administrative database, farms funded by the Marche RDP, measure 214, sub-measure Organic Agriculture matched with Organic farms included in FADN regional dataset.
ii. Identifying the control group: non-treated farms belonging to FADN regional dataset (Organic farms excluded).
iii. Matching treated and non-treated using the Propensity Score.
iv. Evaluating differences by Differences in Differences (DiD) method looking at a set of outcome variables: treated and non-treated farms' performance and main environmental impacts.

3.3. Dataset and variables used

The main characteristics of agricultural farms located in Marche region and subsidied by measure 214 are: specialization in cereals, permanent crops (mainly olive and wine groves) and livestock (sheep mainly).

Data processed for the application of the counterfactual analysis come from the regional RICA/FADN database. In particular, surveyed farms with an economic size higher than 8 thousand Euros of Standard Output were considered (European Regulation No. 1291/2009). The analysis was conducted on a balanced panel data set of 182 farms, observed in two years 2013 and 2015, respectively before and after intervention with 2007-2013 Marche RDP, Organic Farming sub-measure. The agricultural sector taken into account were: farms specialized in arable crop, permanent crop and livestock farms, excluding others types of farming. According with Administrative Database 28 farms out of 182 are OF subsidied (treated). The evaluation would be more accurate if we analyze separately the three farm types. The small number of treated farms found in the FADN balanced panel do not allow us to make this distinction.

The selection of variables for the identification of the control group is one of the most important features of the matching approach. In fact, it is assumed that, conditional to these variables, the exposure to the treatment is random. It is essential to find the optimum combination of available variables so that it would possible to select the most similar control group to treated group. Ideally, every treated should have a non-treated homologous with equal (or very similar) PS value. This situation is difficult to achieve, especially for high values of PS, since, the number of non-treated is often low (Dehejia and Wahba, 2002).

The FADN variables used for the PS estimation are structural and related to the localization of farms (Cisilino *et al.*, 2011): the *farm type* and the *economic dimension* as defined by the reg. (CE) N. 1242/2008, the altimetry zone, the Utilized Agricultural Area (UAA), the Family Agricultural Working Units on Agricultural Working Units (FAWU/AWU), the machine power expressed by KW on Utilized Agricultural Area (KW/UAA). Uaa and FAWU/AWU have been considered as continuous variables, whereas the others are categorical. Furthermore, the presence/absence of Livestock units and female agricultural manager were also taken into account using dummies. Analysis were carried out under the assumption of a common support and the property of balance was respected. This excluded from the control group many farms, non-comparable with the treated, as showed by the PS variables means in the different groups (Table 1).

Table 1- Main descriptive statistics of treated and non-treated (as whole and as non-treated matched).

Variables		n.	Mean	Std. Dev.
FAWU/AWU	Non-treated total	154	0.960577	0.118286
	Non-treated matched	95	0.944539	0.143647
	Treated	28	0.89861	0.204553
KW/UAA	Non-treated total	154	12.82181	13.06958
	Non-treated matched	95	8.65202	6.41665
	Treated	28	6.479776	5.554682
UAA	Non-treated total	154	31.28234	40.53954
	Non-treated matched	95	40.33558	47.96918
	Treated	28	75.09929	97.94211

Source: own data processing (Marche region FADN Database 2013, 2015).

4. Results and discussion

Results of the PS estimation in Figure 3 shows that treated and non-treated have similar probability to access the AES-OF sub-measure, except that for the highest PS value (right part of graph).

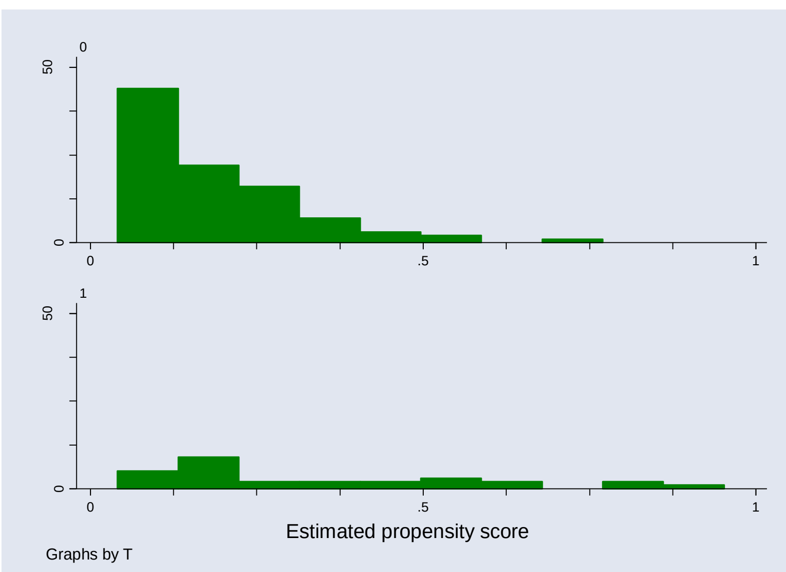

Figure 3 – The Propensity Score for non-treated (0) and treated (1)

Note. The variables used to perform the matching: altimetry zone, UAA, FAWU/AWU, KW/UAA, farm type, economic dimension, presence/absence of Livestock units and female agricultural manager.

Source: own data processing (Marche region FADN Database 2013, 2015).

The other key step concerns the choice of the matching algorithm. The more widely employed in the policy evaluation literature are the Nearest Neighbor Matching, the Radius Matching, the Kernel Matching and the Stratification Matching. Each method has advantage and disadvantage

(Caliendo, 2008), the performance varies case-by-case and depends on the data structure (Zhao, 2000). We processed data using different methods achieving similar results. In this paper we show results obtained with Kernel Matching. In this case all non-treated individuals (included in the common support) are exploited to build the counterfactual. Each treated is matched with all of the non-treated, using a weight which is inversely proportional to the distance from the propensity scores.

The average effect estimation of treatment across the treated individuals (ATT) applying Differences in Differences (DiD) method to treated and non-treated matched by Kernel matching gave a measure of the effectiveness of Agri-environmental scheme adoption on Organic Farming: the ATT (Average Treated Treatment) estimator shows the difference in the average growth (or decrease) of the outcomes between treated and the control group[2].

The results, shown in table 3, highlight a positive (negative) ATT whenever the outcome of treated group is larger (smaller) than the control group (3 years considered). For each outcome the average difference over the period 2013-2015 in the two groups (treated and control group) is presented.

The outcomes considered for Difference in Difference analysis (DiD) are both related to economic and environmental issues . As far as the economic one the results refers to farm income per hectare and per AWU, and costs per hectare while as the Agri-environmental issues the use of pesticide, phosphorus and nitrogen per hectare are shown (table 2). Phosphorus and nitrogen used at farm level included also the organic component because especially in Organic Farming the use of manure as natural fertilizer is relevant, even though the actual provision of nitrogen and phosphorus is very low (0,05% and 0,02% per quintal of bovine manure).

Table 2 – Average treatment Effect on Treated (ATT) of AES-OF Participation

Outcome variables	Nr treated	Nr control group	Mean difference of the treated group	Mean difference of the matched control group	ATT(1)	Standard error
Economic indicators						
Output/UAA (euro/ha)	28	95	-93.34	-291.00	197.65	224.59
Farm Income/AWU (euro/AWU)	28	95	5,845.00	3,506.08	2,338.92	6,384.68
Variable costs/Output (%)	28	95	0.02	0.00	0.02	0.04
Environmental indicators						
Quintals of Pesticides /UAA	28	95	-0.03	3.43	-3.46	3.00
Pesticides expenditure/UAA (euro/ha)	28	95	-2.79	11.95	-14.74	9.80
Quintals of Nitrogen/UAA (q/ha)	28	95	-0.19	0.01	**-0.19(**)**	0.09
Quintals of Phosphorus/UAA (q/ha)	28	95	-0.17	0.02	**-0.18(**)**	0.08
LU/UAA (n)	28	95	0.01	-0.01	0.02	0.03
Number of crops per farm (n)	28	95	0.79	-0.06	**0.85(*)**	0.45

(1) Bootstrap Standard Error

*, ** Statistical significance at 10%, and 5% respectively, of a t-test on the equality of mean differences in the treated and the control group.

Source: own data processing (Marche region FADN Database 2013, 2015).

[2] The Propensity score matching and the Difference-in-Difference comparison were run using STATA package.

To give evidence of the evaluation a priori-expectation some economic and environmental indicators were considered. As it is known the support given by Organic Farming sub-measure generally aims at enhancing environmentally-friendly agriculture that is also economically viable. In particular, the measure aims at compensating lower income due to the farm adoption to particular farming practices. For these reasons, we expected positive effects on environmental indicators, decrease of input use, and, at the same time, the economic results shall be equivalent or at least showing a convergence between organic and conventional farms.
As expected the level of nitrogen and phosphorus used per hectare in the treated group was lower with a significance level of 5%. Farms participating to 214-measure decrease the use of nitrogen by 0.19 quintals per hectare, whereas the increases by 0,01 quintals. The phosphorus use drops in organic farms by 0,17 quintals per hectare, whereas the conventional ones rise the use by 0,02 quintals per hectare. The Average Treatment Effect on Treted (ATT) is 0.18. This results are probably due to the lower need in fertilizer in farms that differentiate more crops cultivated and that employ organic manure more than the conventional farms (control group).
Both agro-chemicals expenditure and quantity are not significant. This is due to the fact inputs allowed in organic practices might be more expensive and the interventions required are more frequent due to the fact that products are slightly less persistent and effective towards the pathogens. As known the basic principle of Organic Farming is the preventive intervention rather than curative one.
On diversification index, calculated as the number of different crops cultivated per year by each farm, the treated farms resulted to be significantly different from the control group, with positive index, thus with higher number of species cultivated. As literature demonstrated (Campanelli and Canali, 2012) farms diversifying crop production enhance soil fertility and prevent soil erosion, organic farms seems to put more effort than conventional ones into crop diversification not only because by introducing legumes into the crop rotation raise soil organic matter, but also because rotating crops over the field is entitled to control and contrast pests and weeds. On the contrary conventional farming tend to intervene against antagonists, either weed or microorganisms, rather than maintaining a good natural balance within the agrarian ecosystem.
As for the economic indicators the difference in income is not statistically significant, neither at 10% nor at 5%. The variable Farm Income includes subsidies both from first and second Pillar of CAP. A convergent result was expected as organic farms accomplish lower yields and the market does not always pays a premium price, thus the economic viability might be worsen also by certification costs and procedures limitations.

Some of the main limitations that affect the analysis are related to data availability and time run. On one hand a larger number of observations would be preferable in order to obtain robust results; on the other hand applying the method on a wider period of time would be positive to better capture changes over time. Moreover, being able to deepen in greater detail the levels of subsidies received by farmers could allow a more detailed assessment of the differences between the two groups.
One of the issue that is worth to be highlighted is the programming capacity of the competent bodies (Managing Authorities). In fact, it comes out how important it would be to foresee terms and needs of the evaluation process at the beginning of the seven-year programming period. The FADN satellite samples could help in realizing this purpose. This is an “integrative system of samples” to the FADN sample. The satellite samples are selected from farms subsided by a specific measure using FADN methodology taking the following steps: a) Analysing the distribution of farms belonging to administrative list of treated farms (selection criteria); b)

Analysng the distribution of treated farms included in the FADN sample; c) Fixing the satellite sample's number of farms, taking into account the analysis requirements; d) Selecting the satellite sample by selecting farms from the list of beneficiaries included in the administrative database.

5. Conclusions

The Farm Accountancy Data Network (FADN) has demonstrated to be a valuable source of information for the evaluation of Rural Development Policies (Cisilino, 2010), as also increasing number of studies have attempted to. The problem of achieving target groups and representative results have been solved using both administrative data for the identification of treated on one side and on the other extracting a group of non-treated farms from FADN sample. In this study the representative sample of treated farms stems from the Marche Region RDP database (measure 214), and FADN data have been used in order to select a group of non-treated farms.

The empirical evidence has showed some critical issues, such as the different aims of databases. However the structural information of agricultural holdings is common to administrative database and to FADN and those are the variables used for matching. The crucial gap between the two sources of information lay on economic data (such as total output, inputs or labour or work-to-thirds expenditure) that is (completely) missing because not requested among the eligibility criteria by the management body. As the administrative database is lacking on economic data, the evaluation of intervention would be possible only considering the number and UAA of farms participating to AES-OF, unless ad hoc surveys were ranby advisory services.

Nevertheless the evaluation process was accomplished and brought to an economic assessment of farm results by analysing treated and non-treated farms within the FADN database. As a matter of fact treated farms seem to reach higher performance in terms of environmental sustainability than farms not participating in Agri-Environmental Schemes. In particular, organic farms not only have diminished the environmental pressure on soil by using less pesticides and fertilizers, but also they improved biodiversity by diversifying crop rotation. Economic performance was also evaluated and turned to be not different from non-treated, due to policy intervention.

As demonstrated also by other studies (Pufahl and Weiss, 2009; Arata and Sckokai, 2016) to differentiate the effects between farm types would be a very important challenge to better understand the effectiveness of the intervention on different farming systems. Unfortunately for this study we did not have the chance to process enough observations to split the sample of treated individuals. Furthermore to better accomplish the ex-post evaluation it is important to have reliable data in terms of number of observations and on representativeness of the phenomenon under investigation for a certain geographical area. On this behalf next to the FAND sample, *satellite samples* might be implemented and economic data might be assembled according to FADN methodology. This could allow an integrative system of samples where a relatively large FADN regional sample is used as a benchmark. When the number farms treated by a specific RDP measure included in the FADN sample is low, in order to achieve appreciable and consistent results, according to a satellite design, other samples could be selected from the administrative lists and revolving around the regional FADN sample. Such samples shall take into account a specific objective and be sufficiently numerous and split into farm type to have a good regional estimate of the specific farm behaviour or policy intervention effect. In fact, different farm productions might affect the outcome of policy intervention. The identification of appropriate satellite samples may produce a double benefit since it could produce the developing of both spatial and dynamic analysis. These purpose aim to describe evolution of farms behavior

over time and monitoring any changes, or giving evidence of their structural and economic performance.
The lesson learned with this study is that farms' structural variables are reliable enough to link different databases, but are lacking in a second stage when sustainability dimensions are to be considered. Economic, environmental and social issues are relevant issues to be evaluated, and need an integrative set of information that can be fulfilled by the FADN variables.
We can conclude that, in line with a recent Italian study on Marche region developed using FADN (Bonfiglio *et al.*, 2017) that Agri-Environmental Schemes affects farms' ecological behavior. As far as the economic farms' performance, at this stage the results (not significant) do not allow to produce a discussion. With the advent of the new CAP tools introducing the greening payments all farms should reorient their farming systems toward diversification and maintaining Ecological Focus Areas.Within the new RDP 2014- 2020 the Marche region has decided to pay farms adopting Agri-environmental practices, namely adopting of maintaining organic agriculture, a higher amount of subsidies per hectare than the previous programming period, (+ 20% to +30% depending on the crop productions). This incentive enhances even more the attention paid by this region for environmental concerns and economic viability of organic farms that sustain higher costs and obtain lower income.
Further research comparing the effect of policy intervention on organic farms in different regions might be challenging to detect areas of improvements and strength of the measure itself. In fact, considering different regions might give evidence that different strategies at regional level need to be implemented to promote Organic Farming independently by the support itself.
Also the RDP measures need to be taken into account with a higher degree of detail so that the effect of single measures can be isolated from the other interventions. Details on RDP payments are available in the Italian FADN dataset however in a very limited number if considering single regions.
Given data availability the Propensity Score Matching combining with DiD estimator is a useful method for assessing RDP programmes. Indeed Propensity Score Matching has the desirable characteristic of allowing comparisons on the outcome variable between treated and control individuals (which are as similar as possible). Policy impact estimation can then be obtained by computing the differences between the values of the outcome variables within matched individuals. This should allow to avoid systematic biases in terms of policy response. To further the research it would be interesting to deepen the analysis for the economic issues and testing the method on other Italian and European regions.

References

Aakkula J., Kuussaari M., Rankinen K., Ekholm P., Helio¨la¨ J., Hyvo¨nen T., Kitti L., Salo T. (2012),Follow-up Study on the Impacts of Agri-environmental Measures in Finland. In Evaluation of Agri-environmental Policies: Selected Methodological Issues and Case Studies, 111–28. Paris: OECD Publishing.

Arata L., Sckokai P. (2016), The Impact of Agri-environmental Schemes on Farm Performance in Five E.U. Member States: A DID-Matching Approach, *Land Economics 92 (1) 167-186*.

Bonfiglio A., Arzeni A., Bodini A. (2017), Assessing eco-efficiency of arable farms in rural areas, *Agricultural Systems 151, 114-125*.

Cagliero R., Cisilino F., Scardera A. (2010), Evaluating Rural Development Programmes Using FADN Data, Italian ENRD 2007-2013, Ministry of Agriculture.

Cisilino F. (2010), “La valutazione delle politiche locali” in Bassi I., Cisilino F. (a cura di) I dati RICA per la valutazione di piani e programmi di Sviluppo rurale: il caso del PSR del Friuli Venezia Giulia 2000-2006, Analisi Regionali, INEA, Roma.

Cisilino F., Zanoli A., Bodini A. (2011), Stimare gli effetti delle politiche di sviluppo rurale: un’applicazione dello Statistical Matching, *Agriregionieuropa, Anno 7 n. 27, pp. 52-55.*

Campanelli G., Canali S. (2012), Crop Production and Environmental Effects in Conventional and Organic Vegetable Farming Systems: The Case of a Long-Term Experiment in Mediterranean Conditions (Central Italy), *Journal of sustainable agriculture, 2012 - Taylor & Francis.*

Caliendo M., Kopeinig S (2008), Some practical guidance for the implementation of propensity score matching, *Journal of economic surveys 22 (1), 31-72.*

Chabé-Ferret S., Subervie J. (2011), Matching vs differencing when estimating treatment effects with panel data: The example of the effect of job training programs on earnings - Unpublished manuscript, IRSTEA, 2012.

Defrancesco, E., Gatto, P., Runge, F. and Trestini, S. (2007), Factors Affecting Farmers’ Participation in Agri-environmental Measures: a Northern Italian perspective. *Journal of Agricultural Economics 59(1): 114-131.*

Dehejia R. H., Wahba S. (2002), Propensity Score Matching Methods for Non Experimental Causal Studies. *The Review of Economics and Statistics 84(1): 151-161.*

Lattanzio (2016), Rapporto di valutazione ex post del Programma di Sviluppo Rurale 2007-2013. Regione Marche. Roma.

Martini A., Sisti M. (2010), Valutare il successo delle politiche pubbliche. Il Mulino.

Martini A., Costabella L., Sisti M. (2006), Valutare gli effetti delle politiche pubbliche. Metodi ed applicazioni al caso italiano, Roma, Formez, page 308.

Pufahl A., Weiss C.R. (2009), Evaluating the effects of farm programmes: results from propensity score matching, *European Review of Agricultural Economics 36, 1, pp. 79-101*.

Rettore E., Trivellato U., Martini A. (2002), *La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati?* - WP Dipartimento di Scienze Statistiche, Università di Padova, Dipartimento di Politiche Pubbliche e Scelte Collettive, Università del Piemonte orientale.

Rettore E., Trivellato U., Martini A. (2003), La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati? *Politica Economica XIX (3) 301-342.*

Trivellato U. (2009), La valutazione degli effetti di politiche pubbliche: paradigma e pratiche. IRVAPP Discussion Paper series.

Rosenbaum P.R., Rubin D.B. (1984), Reducing bias in observational studies using sub-classification on the propensity score, *Journal of the American Statistical Association 79, 516-524*.

Regione Marche (2015), Rural Development Programme 2007-2013 (Regional) - Marche .

Sauer J., Walsh J., Zilberman D. (2012),The Identification and Measurement of Behavioural Effects from Agri-environmental Policies: An Empirical Analysis. Paper presented at the 14th Annual BIOECON, Kings College Cambridge, U.K., September 18–20.

Sisti M. (2007) La valutazione delle politiche di aiuto alle imprese: alcune prospettive d'analisi a confronto. I regimi di aiuto alle imprese 2007-2013. Quaderni Formez, n. 59/2007

Smith J. A., Todd P.E. (2005), Does Matching Overcome LaLonde's Critique of Non-experimental Estimators?, *Journal of Econometrics 125 (1-2), 305-53.*

Vanslembrouck I., Huylenbroeck G., Verbeke W. (2002) Determinants of the Willingness of Belgian Farmers to Participate in Agri-environmental Measures, *Journal of agricultural economics 53 (3) 489-511*.

Zhao, Z.(2000) Data Issues of Using Matching Methods to Estimate Treatment Effects: An Illustration with NSW Data Set, Working Paper, China Centre for Economic Research.