Estimating the Effects of Agri-Environmental Measures Using Difference-in-Difference Coarsened Exact Matching

D. Bertoni[1], D. Curzi[1], G. Aletti[1], A. Olper[1,2]*

[1]*University of Milan, Department of Environmental Science and Policy, Italy*

[2]*LICOS Center of Institution and Economic Performance, KU Leuven, Belgium*

[*]Corresponding author: daniele.curzi@unimi.it[1]

Abstract

This paper studies the effect of agri-environmental measures (AEMs) in improving greener farming practices. We focus on the quantification of the cost-effectiveness of AEMs implemented in the Rural Development Program of the Lombardy Region, during the 2007-2013 programming period. Our work attempts to address the well-known potential failures of these kinds of policy instruments, such as adverse selection effects, by relying on an innovative matching procedure, the coarsened exact matching (CEM). This methodology presents a number of advantages over other matching methodologies, by allowing better control for selection bias. Our empirical analysis focuses on three AEM schemes, which promote arable crop diversification, grassland maintenance and organic farming. From a comparison between CEM and propensity score matching (PSM) using our data, CEM proves, first, to exploit more the heterogeneity of farms in the control group. Second, CEM presents a lower level of imbalance between treated and control farms. Third, it provides, more importantly, lower heterogeneity in the results. Overall, our results suggest that AEMs were apparently effective in improving the farms' environmental performance. However, our cost-benefit analysis highlights that the costs of implementing this policy, when compared to the results obtained, tend to be quite large.

Keywords: Agri-environmental measures, coarsened exact matching, difference-in-difference, cost-benefit analysis.

---

[1] Acknowledgments. This study has been supported by Fondazione Cariplo, within the research project "Evaluation of CAP 2015-2020 and taking action – CAPTION" (Project Id: 2017-2513)

1. Introduction

Agri-environmental measures (AEMs) are policy instruments to support environmentally-friendly farming methods and to improve biodiversity in the rural areas. AEMs provide payments for EU farmers who adopt, on a voluntary basis, green farming practices that go beyond the mandatory environmental quality standards defined by the European Union (EU) and, in particular, by the EU Common Agricultural Policy (CAP). Since 1999, AEMs constitute a relevant part of EU Rural Development Policies under the Second Pillar of the CAP. The AEMs absorb about 25% of the CAP budget.

The objective of this paper is to empirically assess the cost-effectiveness of the application of AEMs by farmers. This was done by giving particular emphasis to the detection of well-known potential failures that typically characterize these kind of policy instruments and which, in turn, may lead to a decrease in their actual effect (Canton et al., 2009). One of these (negative) undesired effects is *adverse selection*. This represents a recurring problem documented in several studies concerning AEMs (e.g. Evans and Morris, 1997; Falconer, 2000; Fraser, 2005; Hart and Latacz-Lohmann, 2005; Baylis et al., 2008; Canton et al., 2009; Unay-Gailhard and Bojnec, 2015; Gómez-Limón et al., 2018). Adverse selection may occur when farmers, whose usual farming practices already satisfy AEMs' commitments – or are close to accomplishing them. These farmers are more likely to participate in the program than farmers who are far from achieving the AEMs' environmental requirements; although the latter would represent the real target of the policy. As a consequence, adverse selection results in a selection bias, as the probability of participation is not randomly distributed between participants and non-participants, but differs for some unknown farm's and farmer's characteristics.

This paper tries to address this issue by empirically estimating the effects of AEMs on green farmers' practices using (Difference-in-Difference-DiD) CEM. This is an innovative matching methodology developed by Iacus et al. (2012; 2017). Previous works in the literature dealing with similar research questions have tried to address these identification problems by relying on PSM (see, e.g., Pufahl and Weiss, 2009; Chabé-Ferret and Subervie, 2013; Arata and Sckokai, 2016). Generally speaking, matching methodologies are particularly suitable for these types of analyses. The existence of a selection bias issue makes the selection of the counterfactual the crucial step in the correct quantification of the average treatment effect. Our choice of using CEM is motivated by the fact that this method improves over existing matching approaches in the estimation of causal inference, by reducing imbalance in the covariates between the treated and control units. CEM incorporates properties of the exact matching procedure, but has a key advantage over other matching methods. It allows the choice of the balance between the treated and control groups *ex-*

ante, rather than having to discover it *ex-post*. In short, data are initially temporally coarsened by the user. Then an exact matching is run on the coarsened data. Finally, the analysis is run on the un-coarsened matched data. In addition, CEM is straightforward to use and conceptually easy to understand. It requires fewer assumptions and possesses more attractive statistical properties for many applications than existing matching methods.

The importance of properly addressing the above-mentioned selection bias stems from the fact that a reliable assessment of the effect of this policy should consider adverse selection. In AEMs, this may result in two main interconnected effects (Ferraro and Pattanayak, 2006; Mante and Gerowitt, 2007; Engel et al., 2008; Chabé-Ferret and Subervie, 2013):

1. the lack of additional effects obtained from the overall participation in the measure, as larger effects are expected from farmers with lower environmental quality practices
2. the windfall effects, that arise when farmers are paid for practices that they would have implemented irrespective of their participation in the policy programme.

In brief, in the presence of adverse selection, the policy implementation may lead to the over-compensation of farmers and limited additional environmental effects (Uthes and Matzdorf, 2013). For the reasons discussed above, existing studies often face recurring methodological difficulties in directly quantifying the (real) effects of these policies. In our contribution, the use of CEM is focused on assessing the effect of the adoption of AEMs on greener farming practices by exploiting the properties of this methodology. This allows a more precise matching of the farms participating in the AEMs with their counterfactuals. As a consequence, our analysis will provide a more reliable quantification of the effects of the policy implementation.

Using CEM we quantify the additional and windfall effects of AEMs implemented in the Rural Development Programme (RDP) of the Lombardy Region, during the period 2007-2013. The choice of Lombardy as a relevant case study, has several justifications. First, Lombardy is the main Italian region in terms of the value of agricultural production and the value added per farm worker. Second, Lombardy is characterized by very intensive farming practices, mostly based on livestock production (milk and meat) and maize monoculture. These, in turn, may determine a considerable environmental pressure. In this framework, AEMs are intended to reduce the environmental effects of agriculture by providing an incentive to farmers who implement low-intensity farming practices. Finally, and perhaps most importantly, our data cover the universe of farmers in the Lombardy Region. Hence, by considering all treated and untreated farms (potentially) involved in AEMs, the analysis may provide an important contribution to better understanding the overall effect of this policy on the farms' agri-environmental outcomes. Indeed, most of the previous studies are based on the analysis of small-scale samples. This leads to a general lack of evidence on large-scale samples, which, in turn, may benefit a more general assessment of the effectiveness of AEMs.

Our paper delivered three main results. First, from a methodology point of view, our assessment of the AEMs through the CEM estimator clearly confirms its interesting properties vis-à-vis the standard matching estimators used in previous literature, in terms of the reduce imbalance and the selection of counterfactual units. Second, the average treatment effect on the treated (ATT) estimates showed that, overall, AEMs had an effect on agri-environmental outcomes that goes in a direction consistent with the policy expectations. Thus they improve greener farming practices in Lombardy. Third, and more importantly, the estimated additional effects are often quite limited when compared with the total payments received by farmers adopting AEMs. Thus, our results provide support for the existence of significant *windfall effects* when agri-environmental policy schemes are applied by farmers.

The remainder of the paper is organized as follows. Section 2 explains the applied methodology. Section 3 describes its implementation in Lombardy and provides an overview on the data and variables used in the analysis. Section 4 summarises the results and Section 5 discusses the main findings and suggests further developments.

2. Methodology

Matching methods are powerful non-parametric approaches used for causal inference. They are very popular and widely used by applied researchers, as they are relatively straightforward to apply and conceptually simple to understand. Theories of statistical inference in the literature, on which are rooted the application of matching estimators, are based on the axiom of simple random sampling. According to this, each individual in the population has the same probability of being treated (Abadie and Imbens, 2006). However, this approach is theoretically appropriate only when relying on an exact matching, where treated and control units thus have the same values for all the pre-treatment covariates or the same propensity score. Unfortunately, this condition is unlikely to be met, as applied researchers usually work with continuous variables and finite data. So the use of exact matching would lead to the loss of most (or even all) of the available observations. In practice, empirical analyses, by employing various typologies of approximate matching estimators (e.g. nearest neighbour matching, radius matching, kernel matching, etc.) regularly violate this exact matching requirement. Thus they do not satisfy the theoretical axiom on which they are based. In particular, this occurs as these matching methodologies operate a simple random sampling by stratifying the sample *ex-post* based on the propensity score. In practice, these methodologies approximate matching within each stratum as if it were an exact matching.

Against this background, Iacus et al. (2012; 2017) propose a theory where they show that by replacing simple with stratified sampling in the way they suggest, matching methodologies become coherent with the theoretical axiom on which they are based. Specifically, they include an assumption about the *ex-ante* stratification of the data (rather than burring it *ex-post*). This assumption formulates an alternative axiom on the data generating process, which follows a stratified sampling framework. All these assumptions that, according to this axiom, are necessary to operate a valid causal inference, are then made explicit in the model. According to this theory all the strata are defined *ex-ante*, and thus, working on the original variables, instead of doing that *ex-post* on more complicated variables, which are retrieved from the matching procedure, like the propensity score or the Mahalanobis distance. As a consequence, the properties of the matching estimators based on this theory satisfy the theoretical axiom on which they are based. This theory is based on the fact that most of the data used by applied researchers are characterized by continuous variables that are featured by natural meaningful breakpoints well known by data analysts.

In short, our empirical analysis is based on the use of the CEM to assess the causal effect of the AEMs on a number of outcome indicators. Unlike other widely used matching estimators that regularly violate the theoretical axiom on which they are based (i.e. simple random sampling), the properties of CEM allow this methodology to be consistent and coherent with its theoretical underling axiom developed by Iacus et al. (2017) on stratified sampling. As a consequence, the use of CEM, rather than other matching methodologies, gives further reliability and credibility to our empirical analysis.

In our exercise we compare variations in the outcome variables between the treated and the selected controls in a pre-treatment and post-treatment period. Thus, as in previous contributions (e.g. Pufahl and Weiss, 2005; Chabé-Ferret and Subervie, 2013; and Arata and Sckokai, 2016), we work with a DID matching estimator. In what follows, after the introduction of the CEM properties, we briefly introduce how this methodology is applied in our context.

2.1 CEM and identification

Assume a sample of n units, which are randomly taken from a population N (with $n \leq N$). Consider now a unit i, which receives a treatment T (then denoted as $T_{i,}$). $T_{i,} = 1$ if i receives the treatment, otherwise $T_{i,} = 0$. In this setting, the outcome variable of interest Y assumes the value of zero ($Y_i = 0$) if the unit i does not receive the treatment, while it assumes the value of 1 ($Y_i = 1$), if unit i receives the treatment. The final outcome then is given by $Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0)$. The problem is that $Y_i(0)$ cannot be observed if i receives the treatment, and similarly, $Y_i(1)$ cannot be observed if i does not receive the treatment, i.e. the standard problem of causal inference.

The matching estimator proposes a solution for this observational data problem, where the treated and control groups are not perfectly identical in the pre-treatment period and not the same, on average, given the treatment not being randomly assigned. The main idea behind the matching approach is to control for pre-treatment individual characteristics in order to address the observational data problem highlighted before. Consider a h-dimensional dataset denoted by $\mathbf{X} = (X_1, X_2, \ldots X_h)$, with X_j being a vector of observed values for a covariate j before the treatment, for all the n observations in the sample.

The treatment effect (TE) for a given unit i, given by the difference between $Y_i(1)$ and $Y_i(0)$ [$TE_i = Y_i(1) - Y_i(0)$] is unobserved. Several causal quantities of interest are given by the average of TE_i, over different sub-samples of units, and so have to be estimated.

The sample average treatment effect on the treated (SATT) is given by the following relation:

$$SATT = \frac{1}{n_T} \sum_{i \in T} TE_i$$

where n_T is given by $\sum_{i=1}^{n} t_1$ and $T = \{1 \leq i \leq n{:}T_i = 1\}$. According to the matching methodology, in order to determine the quantity being estimated, the algorithm can select the control and/or the treated units.

In the same vein of the "no omitted variable bias" assumption in standard regression analysis, the procedure is based on the standard ignorability assumption (Iacus et al., 2012): conditional on **X**, the treatment variable is not dependent to the potential outcomes: i.e. $T_i \perp (Y_i(0), Y_i(1))$. The main objective of the matching algorithm is to non-parametrically control for all the potential confounding elements in the pre-treatment period, for all the selected control variables in the data. The main purpose of the matching is, thus, to make the distributions of the covariates (X) as much alike as possible between the treated and the control groups. This is achieved by reducing the available observations so that the remaining units are better balanced in the two groups. In the case of exact matching (i.e. when the treated and control groups are exactly balanced), a simple difference of the means would lead to a proper causal estimation of the treatment on the outcome, since a further control for X is not necessary given that X is not related to the treatment.

However, exact matching, by requiring the same value for all the control variables in the treatment and control groups, leads to an important reduction in observations, thus, often undermining the possibility to do valid inference. In order to avoid this important loss in the data, other matching methodologies are available, which are not exact matching, but they match treated and control units exploiting a certain degree of similarity. Such methodologies include nearest neighbour matching, radius matching, kernel matching and stratification matching. An important shortcoming of the widely diffuse matching algorithms is that the size is set *ex-ante* the matching, and the balance is then checked *ex-post* the matching algorithm that has worked. If the balance is not satisfactory, then the entire procedure is run again, until an acceptable balance is reached.

The CEM is part of a class of matching methods called monotonic imbalance bounding, which allows bounding the higher level of imbalance in some characteristics of the distribution through an *ex-ante* choice (the coarsening, in the case of CEM). Besides the *ex-ante* choice of the imbalance, the CEM has three main features (see Iacus et al, 2012). First, the congruence principle, according to which data and analysis spaces should be the same, is met in the CEM. Meeting this principle allows us to have a much higher knowledge of the data, which is essential to obtaining a good match. Second, while other approximate matching methodologies need, before the matching, a prior step to restrict the data to a region of common empirical support, the CEM algorithm does not require this procedure. Indeed, the CEM automatically considers only data within a coarsened stratum, where treated and control units are present, while the other observations are dropped. Finally, the computational efficiency of CEM, makes it suitable also for use with very large data sets.

The CEM overcomes these limitations, by providing an exact matching with a reduced loss of observations in the treated and control groups, by choosing the balance between the groups *ex-ante* and not *ex-post,* as in the other matching methodologies. The basic idea is to temporally coarsen each control variable in meaningful groups (which may be either of the same size or not), and then to run an exact matching on these values. Afterwards, the original values (un-coarsened) of the matched data are retained. In addition, one important advantage of the CEM is that it allows the users to choose how to coarsen the variables, in order to preserve meaningful information. In case the user decides not to choose to personally coarsen the variable, the CEM algorithm provides its own way to temporally coarsen the variables.

The CEM algorithm basically works using the following main steps. First, it makes a copy of the set of covariates chosen to make the matching (X^*); second, the variables X^* are then coarsened in different strata, either according to user choice, or automatically through the CEM algorithm. Thirdly, a unique stratum for each observation of X^* is created and each observation is then placed in a stratum. The created strata are reassigned to the original set of data X and any stratum which does not contain at least one treated and one control unit is dropped. The treatment effect is, thus, based on the matching provided by the algorithm, since the difference between treated and control units is based on the difference of the outcome variable between units belonging to the same strata. It is worth noting, that the higher the coarsening (higher number of strata), the lower will be the number of matches provided by the CEM, as well as the lower will be the imbalance.

Because there may be systematic differences between treated and untreated outcomes, even after conditioning on observables, we follow Heckman et al. (1997) in applying a conditional DID CEM procedure. This strategy controls for unobserved heterogeneity and selection bias and thus improves the matching procedure. Introducing a time dimension, with t representing a time period after the programme's starting date (2012) and $t^{'}$ a time period before the program (2005), the conditional DID estimator can be written as:

$$E\left(Y_{it}(1) - Y_{it^{'}}(0)\,\middle|\,T = 1, X\right) - \left(Y_{it}(0) - Y_{it^{'}}(0)\,\middle|\,T = 0, X\right).$$

This estimator has the key advantage over the standard matching procedure of controlling also for unobserved time invariant factors. Clearly, this comes at some costs because in so doing we are assuming that the outcome variables of interest of the treated and control units, absent any treatment, should display the same growth path, namely the parallel trends assumption of the DID method. Thus, in our implementation, DID matching is obtained by applying the CEM procedure to the outcome variables differenced with respect to the pre-treatment period. Finally, note that, other than the parallel trends assumption, our identification strategy still relies on the hypothesis of no spill over effect between treated and control units, i.e. the stable unit treatment value assumption.[2]

3. Data and variables

AEMs represented the main policy measure of the Rural Development Program in Lombardy for the period 2007-2013, accounting for 28.4% of the total public expenditure (around EUR 291 million). Consider, for instance, the year 2012, more than 200,000 ha of utilized agricultural area (UAA) were under agri-environmental commitments. This corresponded to about 20% of regional UAA. Also around 8,000 farms (about 16% of Lombardy farms) were involved in at least one agri-environmental scheme. During the 2007-2013 period, AEMs encompassed a set of 10 different schemes. In this paper we focus on the implementation of three of them, that, taken together, account for about 60% of the total UAA covered by AEMs and involved about 6,600 farms.

The first AEM considered is related to *crops diversification*, a scheme introduced for the first time in the 2007-2013 programming period. The second AEM is on *grassland maintenance*, while the third is on *organic farming*.[3] Note that, the last two schemes have their own equivalent in the

[2] More precisely, the SUTVA assumption has two components (see Rubin, 1974): i) units do not interfere with each other, meaning that treatment applied to one unit does not affect the outcome for another unit and ii) there is only a single version of each treatment level (potential outcomes must be well defined).
[3] With reference to the Lombardy RDP the above mentioned AEMs are identified as: *214_a – crops diversification - 2014_c – grassland maintenance –* and *214_e – organic farming*.

previous AEMs programming periods.[4] It is worth noting that participation in *organic farming* was not compatible with participation in the other schemes. Detailed information about the objectives, characteristics, requirements, eligibility criteria and payments of the three AEMs analysed in the present study are reported in Table 1.

In order to estimate the additional effect of the farms' participation in the AEMs, we used data extracted from the SIARL dataset. SIARL is the electronic system by which Lombardy Region manages farm demands for all CAP subsidies (first and second pillars). In particular, using the SIARL data it is possible to identify, for each year of the 2007-2013 period and for each scheme, those farms participating in AEMs, their UAA under agri-environmental contracts and the agri-environmental payments received. In addition, SIARL contains information at the farm level about the farm structure, farmers' characteristics and the farms' crops and livestock production. SIARL data were used to define, at farm level, dummy variables related to the farms' participation in each of the AEMs, farm environmental performance indicators (used as outcome variables) and a set of control variables to define a correct matching of treated and non-treated farms.

3.1 Participation variables

To identify farms participating in each of the AEM schemes considered, we chose a reference year during the 2007-2013 period. Since the implementation of AEMs in Lombardy began in 2008, with the first farms adopting this policy, and AEMs commitments are plurennial (5-year contracts),[5] we decided to select 2012 as the *reference year* to distinguish between participants and non-participants in the 2007-2013 period. This choice sought to capture the highest number of farms participating in the AEMs, including farms which began their commitments in 2008 and those enrolled in the following years. Moreover, 2012 was the last year in the 2007-2013 period in which new entrants could take part in that policy programme. Hence, 2012 has been selected to be the most representative year to depict farm participation, because it shows the peak of participation during the 2007-2013 period. However, as shown in Table 2, most of 2012 participants enrolled in AEMs scheme in the first years of the 2007-2013 programming period. Hence, we may argue that by the end of the programming period (i.e. 2012), most of the farms had already consolidated the farming practices as required by each scheme.

Based on the farms' participation in 2012, we built three participatory binary variables (*part1_2012, part2_2012, and part3_2012)*. These variables take the value 1 when the farm participated in an agri-environmental scheme and a value of 0 when the farm was potentially

[4] Specifically, the 1992-1999 period was governed by the 2078/92 EC Regulation and the 2000-2006 period was governed by the 1257/99 EC Regulation.

[5] Only the *organic farming* scheme started in 2007 with a 7-year contract for farms entering it in 2007, a 6-year contract for farms entering in 2008 and a 5-year contract for entry from 2009.

eligible for the measure, but did not participate. The potential eligibility was evaluated for each farm and for each scheme based on the presence of admissible crops and an altimetry criterion. As a consequence, the number of eligible farms was different, depending on the scheme (from about 24,000 potentially eligible farms for the *crops diversification* measure to about 30.000 for *organic farming*). Furthermore, we built an additional variable, *part_others_2012,* to capture a farms' participation in at least one of the other schemes of the Lombardy AEMs. As explained below, this variable will be used to estimated cross-over effects between measures.

3.2 Outcome variables

In order to estimate the additional effect of the participation in AEMs we have identified a set of potential outcome variables, based on the stated objectives of each scheme. The selected outcome variables quantified the environmental effect of the farm's participation in a specific AEM. To estimate the treatment effect of participation, outcome variables were calculated for each farm. Each outcome variable was the difference between the value of the outcome in the year 2012 (designated as the reference year for treatment) and the value of the outcome for the farms in 2005 (i.e. pre-treatment status), prior to the considered programming period starting. As the three selected AEMs, and particularly the *crops diversification* and *grassland maintenance* schemes, pay farmers depending on the cultivated crops, the main part of the outcome variables is focused on farmland use. In this regard SIARL provides complete information on the use of each of the 2 million agricultural land parcels of the Lombardy Region (259 different uses) and the farms they belong to. Starting from these data we were able to calculate a set of outcome variables for each farm potentially eligible for AEMs.

We are aware that farmland use represents only the outcomes of the analysed AEMs and not their (actual) environmental effect. That, instead, would be captured by indicators related, for instance, to water quality, nitrogen losses, soil depletion, greenhouse emissions, energy consumption or biodiversity (see Peerlings and Polman, 2008; Wrbka et al., 2008; Pacini et al., 2015; Galler et al., 2015). However, these environmental indicators are generally not available for large samples and at a detailed parcel or farm level (Primdahl et al., 2003). In contrast, information on land use, in our case, is available for the entire universe of farms within the Lombardy Region. Our strategy is supported by the fact that, according to the RDP managing authority, the three AEMs considered in our analysis are meant to reach their environmental achievements by providing incentives for preserving some crops/land uses and by introducing rotations instead of monoculture. In addition, the change in land use by the participating farmers is the most direct outcome of the AEM itself. If we had considered an environmental achievement, this could be also the results of other (simultaneous) shocks attributable for, example, to weather conditions and their respective interaction with land characteristics, all elements quite difficult to control in our exercise.[6] Hence, as the objective of this paper is to estimate the effect of participation in AEMs of a large sample of farms, variables on farm land use represent a valuable proxy for capturing the (potential) environmental effects of the selected schemes.

Table 3 summarises the selected outcome variables for each of the investigated AEMs. For instance, since the AEM on *crops diversification* had as its main objectives to increase biodiversity, to preserve soil structure and organic substance, and to reduce nitrogen losses, by diversifying arable crops, its outcome variables are related to the quantification of crop variety on the farm. In particular, for each farm, we computed the area covered by the main arable crop and the share of the main arable crop on total farm arable crop area, the number of arable crops and their heterogeneity. Furthermore, we calculated the share of nitrogen-fixing crops and non soil-depleting crops on farms arable crop area. For farms engaged in *crops diversification*, the presence in the 5-year rotation of a crop belonging to these two last categories was mandatory, given their positive agronomic and environmental properties. Nitrogen-fixing crops encompass 22 different crops, while non soil-depleting ones include 135 crops. This latter count includes, besides nitrogen-fixing crops, other cereals (maize, sorghum), vegetable crops and other crops useful for their beneficial effects on soil structure. Also set aside land, was included in this last wide category, that basically excluded only winter cereals. It is worth mentioning that rice and temporary grassland have been not used in calculating the outcome variables for the *crops diversification* scheme, because the regulations for these two crops were different from those of the other crops, making a comparison among farms difficult. Moreover, we have not considered other minor land uses, such as flowers, plant nurseries and kitchen gardens. Overall, we accounted for 147 potential arable crop land uses.

For the *grassland maintenance* AEM scheme, we computed two outcome variables quantifying the grassland area in absolute terms and as a share of the farm's UAA (excluding mountain pasture – as the scheme was not applicable in the mountains – and other minor land uses). Finally, for the *organic farming* AEM, we considered as outcome variables the total farms' UAA converted to organic farming and that under conversion.

3.3 Control variables

To estimate the treatment effect for farms participating in the AEMs, we built the counterfactual scenario by using farms not participating in them. For farms belonging to these two groups we observed the differences in all outcome variables between 2005 (pre-treatment scenario) and 2012 (treatment scenario). However, a simple comparison of the selected outcome variables

[6] Note in addition, that the selected outcome variables are in line with those used in previous studies evaluating, at the territorial or farm level, the environmental effects of the different CAP payments (see for instance Arata and Sckokai, 2016; Paracchini et al., 2015; Desjeux et al., 2014; Chabé-Ferret and Subervie, 2013; Beltrán-Esteve et al., 2012).

between participants and non-participants could be misleading. In fact, the probability of the farms' participation in the measures is not likely to be randomly assigned among farms.

To overcome the selection bias and to build a correct counterfactual scenario, we compared treated and untreated outcomes of similar farms using the CEM, by matching farms with similar characteristics on the basis of a set of control variables. The choice of the control variables to be used to match treated and control farms was based on a consideration of both the relevant previous literature dealing with AEMs and the availability of data in our sample. Previous works in the literature highlighted that the probability of a farm being engaged in AEMs mostly depends on factors related to the farms' profitability (e.g. income per hectare), the farms' labour structure (hired versus family labour), the farms' typology, farmers' characteristics (e.g. age, education), the farms' location and the incidence of subsidies on the farms' income (see for instance Vanslembrouck et al., 2002; Defrancesco et al., 2008; Pufahl and Weiss, 2009; Hynes and Garvey, 2009; Bertoni et al., 2011; Chabé-Ferret and Subervie, 2013; Arata and Sckokai, 2016). The use of the SIARL database allows us to consider some of this list of indicators. The database does not contain any information on economic or financial indicators (except for subsidies and an economic dimension expressed in economic standard units - ESU) and labour structure. However, SIARL provides very detailed information on the universe of farms in the Lombardy Region concerning their structure, some farmers' characteristics, the farms' typologies, and their participation in AEMs.

Based on data availability, we selected the counterfactual using the following farm level characteristics in 2005:

- farm size, expressed as the UAA;
- farm location (i.e. mountain, hills or plain);
- type of farming (15 categories)
- the share of the most important crops on the UAA
- livestock density
- farmer's age
- previous participation in AEMs during the 2000-2006 programming period.[7]

We are aware of the existence of a recurring and ongoing discussion about the number of variables it would be better to include when estimating the probability of participating in a treatment when using matching methodologies (in particular PSM). While some authors argue that the use of a small and relevant set of variables is preferable (see Bryson et al., 2002; Person and

[7] Arata and Sckokai (2016), working with five EU member states, estimated the probability of participation in AEMs also using macroeconomic indicators (i.e. gross domestic product per capita and the share of agricultural value added over the total value added of the region) at NUTS1 or NUTS2 levels (NUTS – nomenclature of territorial units for statistics). As we work with only one NUTS2 region, we did not consider macroeconomic indicators as potential variables of interest to estimate the probability of participation in AEMs.

Tabellini, 2003), others hold the opposite view (Rubin and Thomas, 2003). Our choice is limited by data constraints to a set of variables that we really think can allow us to match similar farms. In addition, it is worth noting that, as the level of imbalance in CEM is chosen *ex-ante*, rather than discovered *ex-post,* as in PSM, CEM does not require, as a first step, estimation of the probability of the farm participating in the policy. This, then, exempts the selection of only those variables that significantly affect the policy adoption.

3.4 Sample description

Descriptive statistics of the selected outcome variables for the three considered schemes are reported in Appendix A (see Tables A.1, A.2 and A.3). Figures were calculated for both the full sample and sub-samples of participants and non-participants in each AEM, reporting the values of the outcome variables before the treatment (in 2005). Farms for which we do not have data for 2005 have been excluded from the analysis, as it was not possible to quantify their pre-treatment status. Hence, the sample for the *crops diversification* scheme, consists of about 26,000 farms. Of these, 1,113 participated in the AEM in 2012 and represent the treated units. For other schemes, these numbers are, respectively, 29,478 with 1,668 treated farms for *grassland maintenance*, and 33,789 with 272 treated farms for *organic farming*. The sample related to *AEM_others* is the same as that for the *organic farming*, but the number of participants was much higher (3,141 treated farms).[8]

For *crops diversification,* the value of the outcome variable 'main arable crop (ha)' was initially higher in farms that would have successively participated in the scheme, but this depended mainly on the average farm size. This was 55% higher among participants (38.6 ha versus 24.8 ha for non-participants). However, the share of the main arable crop, the number of arable crops and the heterogeneity of the arable crops were already higher among future participants. This clearly confirms the presence of adverse selection among the farmers.

The same considerations are valid for the other two AEM schemes. Indeed, the average area covered by grassland and the share of grassland on the farm's UAA, and the organic UAA, were found to be significantly higher among participants versus non-participants, again suggesting a huge adverse selection effect.

4. Results

As discussed above, in our empirical analysis we employed an innovative methodology, the CEM. This, to the best of our knowledge, has never before been used in this kind of study and that

[8] The main parts of other AEMs have eligibility criteria similar to those for the *organic farming* scheme

may provide several advantages over the use of other matching methodologies. In what follows, we first test the properties of the CEM in comparison with the PSM. Then, we show, in detail, the results of our empirical analysis, where we assess the effect of participation in AEMs in affecting green farming practices.

4.1 CEM versus PSM: A comparative analysis

In order to test the properties of the CEM, we conducted a direct comparison with the PSM. This, so far, has been one of the most widely used methodologies by other researchers to evaluate the effect of AEMs. As an illustrative example, we considered the analysis of the effect of a farm's participation in the *crops diversification* measure on one of the selected outcome variables – the number of hectares dedicated to the main arable crop. Specifically, we compare CEM (based on the same specification as used in the empirical analysis) with the nearest neighbour PSM.[9] [10]

Our first test looks at the construction of the counterfactual without considering the outcome of the analysis. This will be considered in the second test. In particular, we look at the extent to which the sample of control farms is involved in building the counterfactual of the treatment units. To do that, we look at the distribution of weights that are assigned by CEM and PSM to the control units constituting the counterfactual. For this purpose, we use the Lorenz curve, a graph that is usually employed to represent income inequality or wealth distribution. The Lorenz curve, developed by the American economist Max Lorenz in 1905, plots quantiles of the population on the abscissa axis according to income or wealth. It plots cumulative income or the wealth function on the ordinate axis, so that, for instance, an abscissa value of 0.34 and an ordinate value of 0.025 would suggest that the bottom 34% of the population controls 2.5% of the total income or wealth. We use this representation to plot the weight inequality among the control group in the PSM and in the CEM cases. Accordingly, our Lorenz curves plot quantiles of the control group on the abscissa axis according to increasing weight. On the ordinate axis, the corresponding cumulative function for the weights of the control group is calculated, so that an abscissa value, for instance, of 0.34 and an ordinate value of 0.025 would mean that 34% of the control group will have only 2.5% of the total weights in the matching.

The results presented in Figure 1, show a remarkable difference between the two methodologies. Considering CEM in the left panel of the figure, we observe that about 90% of the

[9] Note that in all the comparative analyses shown in this section, both CEM and PSM are carried out using the same set of control variables.

[10] Among the various matching estimators our choice fell on the use of a nearest neighbour matching with one neighbour, with replacement. Our choice is motivated by the fact that, given that CEM is based on exact matching, we argue that this propensity score matching estimator allows the most reliable comparison with CEM, as it matches any treated farm with the closest one, based on the estimated propensity score.

control units accounts for about 20% of the weights assigned by the algorithm to farms representing the counterfactuals of treated farms. The remaining 10%, having more weights in building the counterfactuals of treated farms, accounts for about 80% of the overall weight distribution. In contrast, the Lorenz curve for PSM, right panel of Figure 1, suggests that the whole weights are absorbed by only 10% of the control unit sample.

We also estimate in the two cases the Gini coefficient, which is used in our analysis to measure the level of inequality represented in the figure. The Gini coefficient can range from 0 (or 0%) to 1 (or 100%). Complete equality, in which every unit of the control group has the exact same weight as the match, corresponds to a Gini coefficient of 0. A Gini coefficient of 1 means that only one unit of the control group has been matched. The lower the Gini coefficient, the greater the number of control groups that have to be considered in its diversity. In our analysis the Gini coefficient proved to be lower for CEM than for PSM. Overall, the results of this test suggest that CEM allows a higher exploitation of the data in the sample and thus it considers the greater the heterogeneity of the farms in the counterfactual sample.

We next present our second test, that compares the level of imbalance obtained through the use of CEM and PSM, together with the estimated ATT. For this purpose, we rely again on the estimate of the effect of the farms' participation in the c*rops diversification* measure on the number of hectares dedicated to the main arable crop. In particular, we follow Iacus et al. (2017), by analysing our demonstrative example through a simulation of 500 CEM and 500 PSM. We then compare the results obtained in terms of their level of imbalance and ATT.[11] For the CEM, this simulation is carried out by creating 500 random stratifications of the control variable space, and, in particular, by breaking-down the support of each covariate into a random number of strata.[12] The simulation for the PSM in based on 500 matching solutions obtained through a nearest neighbour PSM, where the results are obtained through a random selection of propensity score models and their respective calipers.

Figure 2 plots the results of this test, where on the ordinate axis we show the ATT obtained with our 500 CEM and PSM simulations, while on the abscissa axis we have their respective measures of imbalance (as measured by the L1 distance).[13] The imbalance represents a measure of the difference between the multivariate empirical distribution of the pre-treatment covariates for the treated group and the matched control group. We use here the 'multivariate imbalance measure' as

---

[11] This analysis has been carried out using the *spacegraph* command, available in the software R. Note that within this command, nearest neighbour matching can be implemented considering one neighbour only.

[12] The simulation chooses uniformly-based intervals for an integer that varies from 1 to 15.

[13] The Minkowski or L1 distance represents a comprehensive measure of global imbalance that basically measures the difference between the multivariate histogram in the treated group with the multivariate histogram in the control group. For more details on the computation of the L1 distance see Iacus et al. (2008).

defined by Iacus et al. (2011). Matching methods are designed to find the best balance between the distribution of covariates in the treatment and control samples. From this perspective, the lower the measured imbalance, the lower is the bias in the matching of the treated and their respective control units. By comparing the results of our simulations plotted in Figure 2, some important differences between the two methodologies emerge. Estimations obtained with CEM (black crosses) present higher heterogeneity in terms of imbalance with respect to the PSM estimations (red circles). However, about half of the CEM estimations show a level of imbalance that is lower than that obtained with PSM. These are more concentrated around the values of 0.4-0.5. From this perspective, our results suggest that, on average, CEM can perform better than PSM in finding a match in the distribution of covariates between the treated and control groups, and thus lead to a lower bias in the estimation. However, the most striking results emerging from this analysis concern the ATT estimations. The variability of the results obtained with our randomised CEM (included in the black circle) is much lower than that obtained with the PSM (included in the red circle). From this perspective, this finding suggests that, in this specific case, the high variability of the results obtained with the PSM may lead the ATT estimation, relative to the effect of the farm's participation in the *crops diversification* measure on the number of hectares dedicated to the main crop, to be either positive or negative, depending on how the PSM is specified. In contrast, the CEM provides results that are very close to each other. In this case it suggests a uniquely negative ATT, no matter how the model is specified. Clearly, the last one, is the key advantage of the CEM versus the PSM estimator.

To summarise the findings emerging from our two tests, we argue that CEM has shown some relevant characteristics that should be worth considering. First, CEM proved to better exploit the heterogeneity of farms included in the control groups than PSM. Second, CEM shows the capability of obtaining a lower biased matching between the distributions of covariates in the treated and control groups. Third, and perhaps, more importantly, the results obtained with CEM show a much lower heterogeneity than those obtained with PSM, thus limiting the discretionary use of the results because of the model specification. Since the ultimate objective of our empirical analysis is a policy evaluation, we cannot overlook this important property of CEM.

Despite our tests highlighting some important characteristics of CEM, we are not arguing that the use of CEM is preferable to other matching methodologies in absolute terms. The same comparisons presented in this section when used in a different case study may provide the opposite results. What emerges from our analysis is that, when considering the use of matching methodologies for empirical analysis, it is worth doing a preliminary assessment in order to get more information on which methodology enables the most reliable results to be obtained.

4.2 Assessing the effectiveness of AEMs through DID CEM

Table 4 provides information concerning the number of matched and unmatched farms in the treated and control groups using CEM. Two sets of estimates are considered, one with the full set of covariates and one where participation in the AEMs in the previous programming period is excluded from the set of covariates. Looking at the CEM results obtained using the full set of covariates, the share of matched, non-treated farms is around 40-50% for all the considered AEMs, except for *organic farming*. In this instance the share of matched non-treated farms is only 14.5%. Considering the treated farms, the share of those matched is much higher than those non-treated, varying from around 70% to 75%. Again the exception is for *organic farming*, where only 30.9% of the farms have been matched. When excluding previous participation in AEMs, and thus using a less-restrictive matching criteria, the number of matched farms, both in the treated and untreated samples grows considerably, in particular when considering *organic farming*. As shown by the figures reported in Table 4, both sets of estimates present a number of matched non-treated farms considerably higher than the treated farms, and so ensures a good representativeness for the non-treated (counterfactual) sample.

Tables 5, 6, 7 and 8 present the results of estimating the ATT of the different AEMs considered on a set of outcome variables by using the (DID) CEM. Figures in the tables show the average variation for each outcome variable over the period 2005-2012, for both samples of treated and non-treated farms matched with CEM. This information can be useful in interpreting the results, as they suggest the average trends in the two groups over the period 2005-2012. The ATT is then obtained by subtracting the average variation in each outcome variable for the sample of non-treated matched farms from that relative to the sample of treated matched farms.

The estimated (positive or negative) ATT may arise from the difference in variations of the same sign, or of a different sign. In the first case, the treatment strengthens (or limits) a trend that is common between treated and non-treated farms. In the second case, it inverts the trend that would have occurred without treatment. In the latter event the observation of the absolute values of the estimated differences shows whether the net effect of the measure on treated farms prevails, or not, on the attenuation of an opposite trend among non-treated farms. Notably, for each of the AEMs considered, the mean differences among groups and the related ATT have been computed. This has been done not only for the outcome variables that are supposed to be the target of the relative measure (figures in bold), but also for the other outcome variables that should be considered as the target of other AEMs. This makes it possible to assess the potential existence of cross-over effects (figures in italic).

All Tables 5-8 present in the first three columns the results for the mean differences between treated and non-treated farms, and the ATT estimated with CEM, including a full set of control variables, as discussed in Section 2.3. As a robustness check, we present in the last three columns the previous results obtained excluding from the set of control variables referring to farms with previous participation in AEMs.[14]

At first sight, it emerges that qualitatively the estimated ATT are robust to the use of the two sets of control variables. For the sake of simplicity, as we consider the estimations obtained using the full set of control variables (columns 1) more reliable, in what follows we focus the discussion on these results.

Table 5 presents the results concerning the *crops diversification* measure. Farms participating in this measure show a reduction in the main arable crop area in terms of the number of hectares with respect to non-treated farms (ATT = – 0.46 ha), although the difference is not statistically significant. However, when the same outcome is expressed as a percentage, instead of number of hectares, the ATT is significant at the 1% statistical level. Specifically, farms participating in the measure show a reduction in the share of the main arable crop area of 11.35% with respect to the counterfactual, thus not an irrelevant effect. Nevertheless, in this case the measure slightly counterbalances a general trend that sees a strong percentage increase in the share of the main arable crop for non-treated farms (10.38%). The treated farms actually reduced their share of the main arable crop area by just 1% between 2005 and 2012. Treated farms show a significant increase in the number of arable crops (ATT = 0.68) and a sizable increase in the arable crop heterogeneity index (ATT = 12.9). For the former outcome, the ATT is the result of a difference computed between an average arable crop reduction of 0.835 among the non-treated farm sample and one of 0.158 in the treated farm sample, over the period 2005-2012. In this case, the measure attenuates the broad trend, which, however, still persists among the treated farms. With reference to the latter outcome, treated farms slightly increased their arable crop heterogeneity, compared to an evident decrease of the outcome in the non-treated group. The proportion of nitrogen-fixing crops slightly decreases on treated farms (1.08%) and to a lesser extent on non-treated farms (3.03%). The ATT is estimated with less precision (p-value < 0.1). Farms participating in this measure also reduced significantly more the share of non soil-depleting crops (ATT = 3.5%). Finally, concerning potential cross-over effects, treated farms show also an increase in the number of hectares of grassland (ATT = 0.56 ha) and their share of the farmland (ATT = 1.44%).

[14] Note that we also run the same estimations presented in Tables 5 to 8, using both local linear PSM and nearest neighbour propensity score matching. Overall, the results prove to be consistent with those shown in these tables. However, as in a previous section, we proved that, in our case, CEM performs better than PSM, and we believe it is more coherent to present the results obtained with CEM only.

Table 6 considers the ATT estimates for *grassland maintenance*. Farms participating in this scheme show a positive effect in the grassland area, both expressed as number of hectares (ATT = 2.3 ha) and as the proportion over UAA (ATT = 5.8%); both significant at the 1% statistical level. Quantitatively, these are relevant effects, even if they derive mainly from the counterbalancing of a dominant trend of decreasing values in permanent grassland rather than a net increasing effect in treated farms. As regards potential cross-over effects, participation in such a scheme seems to lead farmers to reduce somewhat the organic area and to weakly increase the number of hectares of the main arable crop.

Table 7 presents the results concerning the *organic farming* measure. Not surprisingly the results are stark. Indeed, treated farms show a statistically significant increase in organic farming expressed both as the number of hectares (ATT = 5.8 ha) and as the number of hectares under conversion (ATT = 4.4 ha). In this case, the increasing effect of the organic farming area among treated farms predominates over the decreasing trend detected for non-treated farms. Cross-over effects for this AEM are generally not significantly different from zero, except for the number of hectares of the main arable crop. This latter shows an increasing trend on farms participating in this AEM.

Finally, Table 8 presents the results concerning the ATT of the participation in at least one of the other schemes of AEMs on the set of outcome variables considered as targets for the other three AEMs. The main results suggest that, first, treated farms show a slight increase in the number of arable crops. This occurs in both arable crops heterogeneity and in the percentage of non soil-depleting crops, while the share of leguminous crops decreases by 1.4%. Second, treated farms show a reduction in the percentage of grassland area. Finally, the participation in *other AEMs* leads farms to a higher increase in the organic farming area, both expressed as number of hectares and under conversion.

When estimating the ATT using CEM the results are still robust and qualitatively similar to the ones discussed above. And this without considering previous participation in AEMs as control variables, thus using less-restrictive criteria to match treated and untreated farms. Focusing on the outcome variables that we consider as the main targets of the different schemes, the only result that changes significantly is the ATT of *crops diversification* with reference to the percentage of non-soil-depleting crops. In this case the estimated effect is no longer significant. In general, estimations that do not control for previous participation, show results similar to those previously discussed from a qualitative point of view, but less so from a quantitative one. The ATT is always lower in magnitude, but still significant. From this perspective, we may argue that the higher effect detected by CEM, considering previous participation in AEMs within the set of controls, is because the estimated additional effect is likely to be higher when treated farms not participating in previous AEMs are matched with non-treated farms that did not participate in previous AEMs as well. Indeed, if previous participation is not considered in the set of controls, the outcomes related to treated farms not participating in previous AEMs may be compared with those of farms that actually participated, and, thus, lead to a likely underestimation of the additional effects.

5. Additional effects or windfall effects: money for nothing?

When assessing the effectiveness of a policy that foresees a payment to the beneficiaries, in addition to the estimation of the actual effect (in our case expressed as ATT), it is important to associate the results obtained with their cost of implementation. In what follows, starting from the estimations presented above, we have performed a straightforward cost-benefit analysis. This allows for a better understanding of the extent to which the moneys spent for the implementation of the policy have produced the results obtained.

From a theoretical point of view, the social welfare effect of each AEM should be measured accounting for its effect on changes in the consumer surplus (*CS*), producer surplus (*PS*) and taxpayer cost (*TX*). Because our analysis does not quantify the AEMs effect on *PS*, as we do not have information on the AEM effects on farm profit or job creation/maintenance. Following Chabé-Ferret and Subervie (2013), we adopt a taxpayer's view, by computing the variation in CS net of taxpayer's cost.[15] In practice, variation in *CS* is related to the magnitude of the AEMs' additional effects, i.e. our estimated *ATT,* while the taxpayers' costs are related to the amount of the AEM payments spent to get them.

Our cost-benefit analysis consists of two main steps. First, we estimate the windfall effect, by quantifying what would have been the value of the outcome variables in event of the farms not having participated in the policy. This is done, by subtracting from the outcome mean of the treated farms, the estimated ATT. The windfall effect, thus, gives a flavour of the effectiveness of the policy. Second, we assess the effectiveness of the policy in the light of its implementation costs. For this purpose, we consider the average farm cost for policy implementation, relating it to the estimated additional effects.

Table 9 presents the results of our cost-benefit analysis. In column (a) we report for treated matched farms the average value of each outcome variable after the treatment (in 2012). Following Chabé-Ferret and Subervie (2013), we compute windfall effects (column d) as the difference between the post-treatment observed level of the outcome variables in treated matched farms (column a) and the estimated ATTs (column c). In a second step, we considered for each scheme

[15] For additional details on the approach, see equations 21 and 22 of Chabé-Ferret and Subervie (2013, p. 23-24).

the average annual farm payment received by treated matched farms (column e) and their average area under agri-environmental commitments in 2012 (column d). Using these values, we compute the average payment per hectare under the constraints (column f) and the average payment per unit of additional effect (column g).

As the outcome variables are expressed in different units of measurement and the direction of the effect could be both negative or positive, the interpretation of windfall effects and their comparison with additional effects may not be straightforward. The easier case is represented by outcome indicators directly tied to the financed practice (e.g. area of grassland or organic farming). For example, with reference to *grassland maintenance* we observe that the average grassland area in treated farms after the treatment is 13.25 ha, while the ATT is 2.32 ha. As a consequence, treated farms, under the hypothesis of no-treatment, would have kept 10.93 ha of grassland anyway irrespective of the payment. That is the windfall effect. From an economic point of view, it follows that the *grassland maintenance* annual average farm payment (EUR 3,021 per farm), generated by voluntary uptake of the measure on 13.25 ha, in reality has contributed to obtain an additional effect of only 2.32 ha, with a payment for additional units of ATT of EUR 1,303/ha, against a payment of EUR 228/ha under the constraint.

Similarly, considering the case of the *organic farming* scheme the interpretation of windfall and additional effects is quite immediate. Given an *ex-post* average farmland area under organic farming among treated matched farms of about 17 ha, the ATT has been quantified as 5.78 ha, with a resulting windfall effects of 11.31 ha. This result determines that the average payment per unit of additional effect rises to EUR 770/ha against an average payment per area under constraint of EUR 260/ha. Interestingly, payments for conversion to organic farming in practice determine only the additional effects.

In summary, from this stylised cost-benefit analysis the main conclusion that can be drawn is that, the investigated AEMs prove to exert a significant effect in a direction consistent with public policy. However, at the same time the burden of the taxpayer costs for reaching these effects appears to be, by far, extraordinarily large. Hence, the economic rationale of this policy could only be defended if other relevant hidden benefits or spill over effects, such as job creation effects (not considered in the present analysis), result from the policy's implementation. Current empirical evidence on the economic effects of the Pillar II policy, however, gives only weak support to this compensating effect (see Olper et al., 2014; Garrone et al., 2018). Indeed, the impact on job creation of agri-environmental measures, though often positive, is not enough to reverse our cost-benefit analysis conclusion.

6. Conclusions

This paper provides an *ex-post* evaluation of the direct and cross-over effects of three AEMs implemented in the 2007-2013 RDP program in the Lombardy Region of northern Italy. In order to assess whether these policy measures determine additional environmental effects or not, we exploited data from a big dataset of about 50,000 farms, representing the universe of farms in the Lombardy Region. Starting from this dataset we built for each scheme a sample of farms potentially eligible for these policy instruments, calculating for each of them a set of environmental outcome indicators before and after the period of policy implementation. The empirical analysis was based on the use of an innovative matching procedure called CEM. To the best of our knowledge this is the first time such a matching procedure has been implemented in the assessment of an agri-environmental related policy. This methodology allows the causal effect of AEMs to be assessed for the selected outcome indicators. Additionally, it leads to better controls for selection bias with respect to other matching procedures. Finally, we observed for each measure whether additional or windfall effects prevail through a stylized cost-benefit analysis that considers the amount of public funds spent for policy implementation.

From a methodological perspective, our analysis highlights some important properties of CEM, which should make it worthy of consideration when deciding how to empirically deal with a policy evaluation, such as the implementation of AEMs. In our exercise, when comparing CEM with PSM, the former proved to exploit more deeply the heterogeneity of farms in the sample. And, perhaps more importantly, allows for having, on average, a lower level of imbalance and more stable results.

As regards the effect of AEMs on green farming practices, our results provide evidence that the considered AEMs were apparently effective in improving the environmental performance on farms participating in these policy schemes. Most of the selected outcome variables proved to be affected by the implementation of the policies in a direction consistent with the policy-makers' expectations. However, the results are more nuanced when cross-over effects are considered. Yet the magnitude of the estimated effects are quite limited, with the notable exception of the *organic farming* scheme. Furthermore, the effects found for the *crops diversification* and *permanent grassland maintenance* schemes arise mainly from an attenuation, or a slight counterbalancing, of general farming intensification trends rather than to a net effect observed among the treated farms during the observation period. Consequently, the cost-benefit analysis highlights that the cost of the implementation of this policy, if compared to the results obtained, appears to be extraordinarily large.

This paper provides interesting insights both from the methodological and policy implication points of view. Our results suggest that CEM may represent a suitable instrument to assess CAP payment effects. The future CAP framework, that will be based, most likely, on national/local flexibility and which will probably attach high relevance to the achievement of environmental goals, will make policy-evaluation analysis increasingly relevant.

Some limitations of this approach are worth mentioning. For the AEMs analysed in this paper, the implementation of CEM, and of other matching estimators, strictly needs data for both treated and non-treated farms – better if for a wide sample. However, not all the CAP measures have this property, especially when considering those schemes paying for farming practices. For instance, sod-seeding or minimum tillage measures present outcomes that are available only for treated farms, as data for non-compliant farms are not collected. Moreover, the assessment of a precise treatment effect would necessarily need primary data gathered from field experiments on restricted samples of farms. Clearly, in this case CEM cannot be considered an option for a policy-evaluation analysis.

References

Arata, L., Sckokai, P., 2012. The Impact of Agri-environmental Schemes on Farm Performance in Five E.U. Member States: A DID Matching Approach. Land Econ 92(1), 167-186. https://doi.org/10.3368/le.92.1.167

Baylis, K., Peplow, S., Rausser, G., Simon, L., 2008. Agri-environmental policies in the EU and United States: A comparison. Ecol Econ 65(4), 753-764. https://doi.org/10.1016/j.ecolecon.2007.07.034

Beltrán-Esteve, M., Gómez-Limón, J.A., Picazo-Tadeo, A.J., 2012. Assessing the impact of agri-environmental schemes on the eco-efficiency of rain-fed agriculture. Span. J. Agric. Res. 10(4), 911-925. https://doi.org/10.5424/sjar/2012104-3088

Bertoni, D., Cavicchioli, D., Pretolani, R., Olper, A., 2011. Agri-environmental measures adoption: New evidence from Lombardy region, in Sorrentino, A., Severini, S., Henke, R. (Eds.), The Common Agricultural Policy After the Fischler Reform: National Implementations, Impact Assessment and the Agenda for Future Reforms, Ashgate Publishing, Surrey, UK, pp. 275-294.

Canton, J., De Cara, S., Jayet, P.A., 2009. Agri-environmental schemes: Adverse selection, information structure and delegation. Ecol Econ 68(7), 2114-2121. https://doi.org/10.1016/j.ecolecon.2009.02.007

Chabé-Ferret, S., Subervie, J., 2013. How much green for the buck? Estimating additional and windfall effects of French agro-environmental schemes by DID-matching. J Environ Econ Manage 65(1), 12-27. https://doi.org/10.1016/j.jeem.2012.09.003

Defrancesco, E., Gatto, P., Runge, F., Trestini, S., 2008. Factors affecting farmers’ participation in agri-environmental measures: A Northern Italian perspective. J Agric Econ 59(1), 114-131. https://doi.org/10.1111/j.1477-9552.2007.00134.x

Desjeux, Y., Dupraz, P., Kuhlman, T., Paracchini, M.L., Michels, R., Maigné, E., Reinhard, S., 2015. Evaluating the impact of rural development measures on nature value indicators at different spatial levels: Application to France and The Netherlands. Ecol Indic 59, 41-61. https://doi.org/10.1016/j.ecolind.2014.12.014

Engel, S., Pagiola, S., Wunder, S., 2008. Designing payments for environmental services in theory and practice: An overview of the issues. Ecol Econ 65(4), 663-674. https://doi.org/10.1016/j.ecolecon.2008.03.011

Falconer, K., 2000. Farm-level constraints on agri-environmental scheme participation: a transactional perspective. J Rural Stud 16(3), 379-394. https://doi.org/10.1016/S0743-0167(99)00066-2

Ferraro, P.J., Pattanayak, S.K., 2006. Money for nothing? A call for empirical evaluation of biodiversity conservation investments. PLoS Biol 4(4), e105. https://doi.org/10.1371/journal.pbio.0040105

Fraser, R., 2005. Moral Hazard and Risk Management in Agri-environmental Policy. J Agric Econ 53(3), 475-487. https://doi.org/10.1111/j.1477-9552.2002.tb00033.x

Unay-Gailhard, İ., Bojnec, Š., 2015. Farm size and participation in agri-environmental measures: Farm-level evidence from Slovenia. Land use policy 46, 273-282. https://doi.org/10.1016/j.landusepol.2015.03.002

Galler, C., von Haaren, C., Albert, C., 2015. Optimizing environmental measures for landscape multifunctionality: Effectiveness, efficiency and recommendations for agri-environmental programs. J Environ Manage 151, 243-257. https://doi.org/10.1016/j.jenvman.2014.12.011

Garrone, M., Emmers, D., Olper, A. Swinnen, J. 2018. Jobs and Agricultural Policy: Impact of the Common Agricultural Policy on EU Agricultural Employment. Paper presented at 30th ICAE Conference, International Association of Agricultural Economists, July 28-August 2, 2018, Vancouver, Canada.

Gómez-Limón, J.A., Gutiérrez-Martín, C., Villanueva, A.J., 2018. Optimal Design of Agri-environmental Schemes under Asymmetric Information for Improving Farmland Biodiversity. J Agric Econ (in press). https://doi.org/10.1111/1477-9552.12279

Hart, R., Latacz-Lohmann, U., 2005. Combating moral hazard in agri-environmental schemes: a multiple-agent approach. Eur. Rev. Agric. Econ. 32(1), 75-91. https://doi.org/10.1093/erae/jbi002

Heckman, J., Ichimura, H., Todd, P.E., 1997. Matching as an econometric evaluation estimator: evidence from evaluating a job training program. Rev. Econ. Stud. 64, 605–654. https://doi.org/10.2307/2971733

Hynes, S., Garvey, E., 2009. Modelling Farmers' Participation in an Agri-environmental Scheme using Panel Data: An Application to the Rural Environment Protection Scheme in Ireland. J Agric Econ 60(3), 546-562. https://doi.org/10.1111/j.1477-9552.2009.00210.x

Iacus, S.M., King, G., Porro, G., 2011. Multivariate Matching Methods That Are Monotonic Imbalance Bounding. J Am Stat Assoc 106(493), 345-361. https://doi.org/10.1198/jasa.2011.tm09599

Iacus, S.M., King, G., Porro, G., 2012. Causal Inference Without Balance Checking: Coarsened Exact Matching. Political Anal., 20(1), 1-24. https://doi.org/10.1093/pan/mpr013

Lankoski, J., Cattaneo, A., 2010. Guidelines for Cost-effective Agri-Environmental Policy Measures. OECD Publication Centre, Paris Cedex, France, pp. 51-62, ISBN (978-92-64-08665-4)

Mante, J., Gerowitt, B., 2007. A survey of on-farm acceptance of low-input measures in intensive agriculture. Agron. Sustain. Dev. 27(4), 399-406. https://doi.org/10.1051/agro:2007038

Olper, A., Raimondi, V., Cavicchioli, D. Vigani, M., 2014. Do CAP payments reduce farm labor migration: A panel data analysis across EU regions. Eur. Rev. Agric. Econ. 41(5), 843–873. https://doi.org/10.1093/erae/jbu002

Pacini, G.C., Merante, P., Lazzerini, G., Van Passel, S., 2015. Increasing the cost-effectiveness of EU agri-environment policy measures through evaluation of farm and field-level environmental and economic performance. Agr Syst 136, 70-78. https://doi.org/10.1016/j.agsy.2015.02.004

Paracchini, M.L., Bulgheroni, C., Borreani, G., Tabacco, E., Banterle, A., Bertoni, D., Rossi, G., Parolo, G., Origgi, R., De Paola, C., 2015. A diagnostic system to assess sustainability at a farm level: The SOSTARE model. Agr Syst 133, 35-53. https://doi.org/10.1016/j.agsy.2014.10.004

Peerlings, J., Polman, N., 2008. Agri-environmental contracting of Dutch dairy farms: the role of manure policies and the occurrence of lock-in. Eur. Rev. Agric. Econ 35(2), 167-191. https://doi.org/10.1093/erae/jbn022

Primdahl, J., Peco, B., Schramek, J., Andersen, E., Oñate, J.J., 2003. Environmental effects of agri-environmental schemes in Western Europe. J Environ Econ Manage 67(2), 129-138. https://doi.org/10.1016/S0301-4797(02)00192-5

Pufahl, A., Weiss, C.R., 2009. Evaluating the effects of farm programmes: results from propensity score matching. Eur. Rev. Agric. Econ 36(1), 79-101. https://doi.org/10.1093/erae/jbp001

Rubin, D., 1974. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. J Educ Psychol. 66 (5), 688–701. http://psycnet.apa.org/doi/10.1037/h0037350

Uthes, S., Matzdorf, B., 2013. Studies on Agri-environmental Measures: A Survey of the Literature, Environ Manage 51, 251-266. https://doi.org/10.1007/s00267-012-9959-6

Vanslembrouck, I., Huylenbroeck, G., Verbeke, W., 2002. Determinants of the Willingness of Belgian Farmers to Participate in Agri-environmental Measures. J Agric Econ 53(3), 489-511. https://doi.org/10.1111/j.1477-9552.2002.tb00034.x

Wrbka, T., Schindler, S., Pollheimer, M., Schmitzberger, I., Peterseil, J., 2008. Impact of the Austrian Agri-Environmental Scheme on diversity of landscapes, plants and birds. Community Ecol 9(2), 217-227. https://doi.org/10.1556/ComEc.9.2008.2.11

Figures

Figure 1: Lorenz curve for cumulative distribution of weights in CEM and PSM estimations

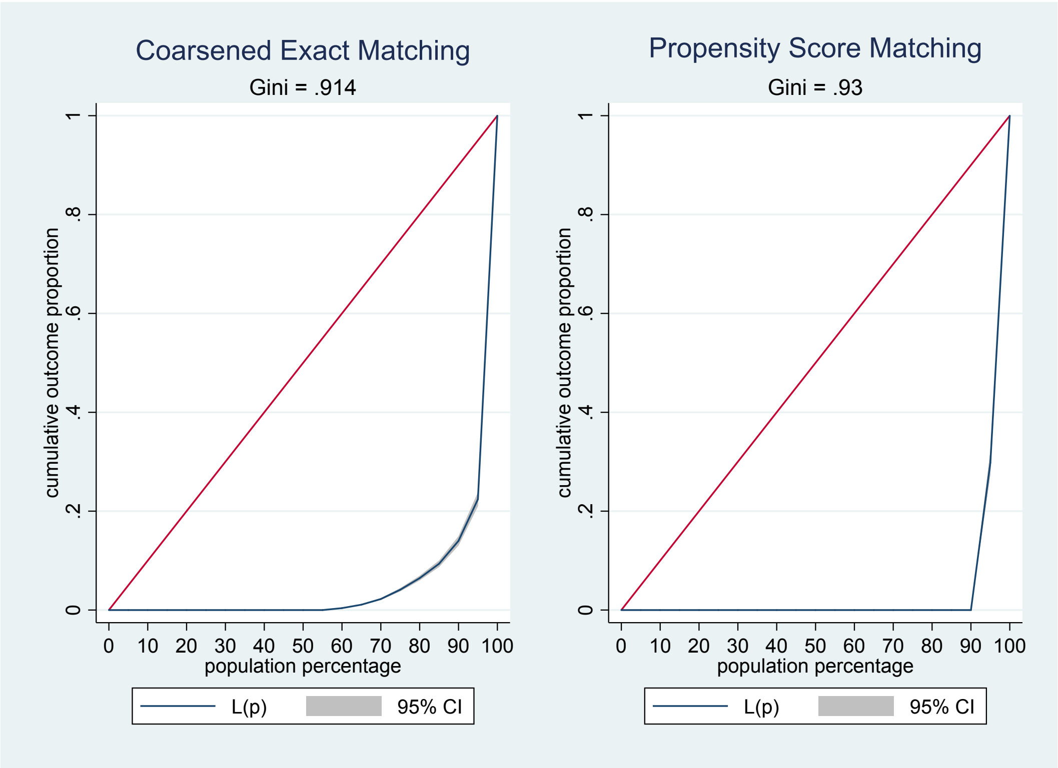
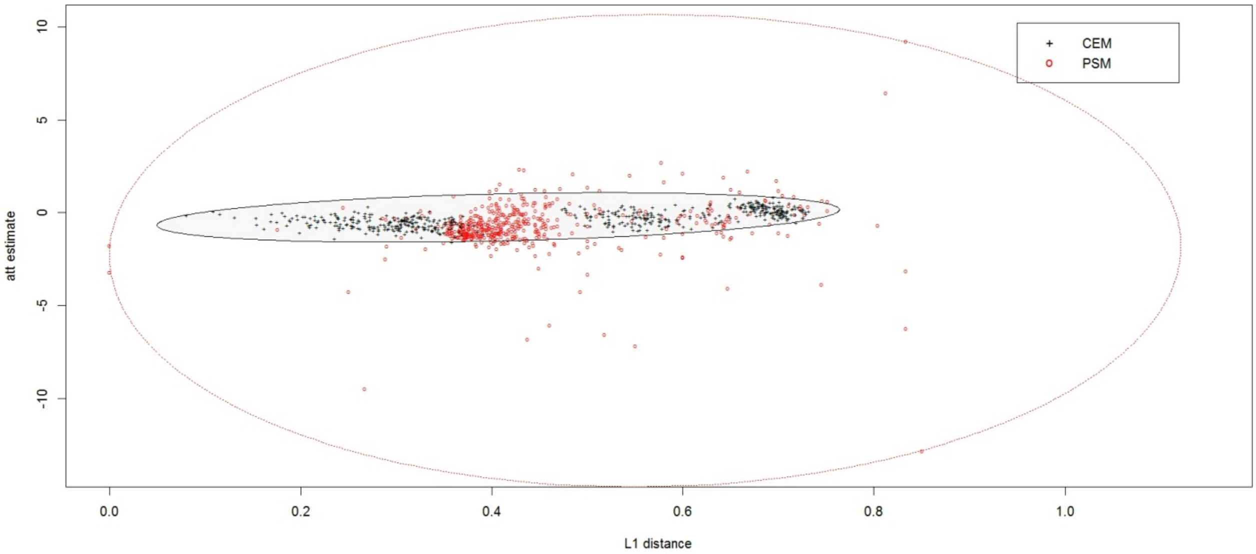


Figure 2: ATT and imbalance on 500 simulated CEM and PSM



Note: The figure plots ATT and the level of imbalance (L1 distance) for 500 simulations of CEM and PSM. CEM results are identified with black crosses, while red circles show the results for PSM. CEM results are included in the black ellipse, while PSM results are shown within the red ellipse.

Tables

Table 1 – Description of Lombardy agri-environmental schemes

Measure	Requirements	Eligibility criteria	Payments
Crops diversification	a) to cultivate at least 3 arable crops in a 5-years rotation on the same parcel; b) at least one crop in the 5-year rotation has to be a nitrogen-fixing crop or another arable crop (excepted for winter cereals and rice) or a 1-year set-aside; c) an arable crop cannot be cultivated on the same field for two years consecutively; d) it is not permitted to grow any winter cereals on the same field for two years consecutively; e) to adopt a balanced fertilization plan (with soil analysis)	a) at least 2 hectares of arable crops in the plain; b) at least 1 hectare of arable crops in the mountains and in the hills	169 EUR/ha
Grassland maintenance	To maintain grassland on the same area for 5 years (not applicable for mountain pastures)	a) at least 1 hectare of grassland in the plain; b) at least 0.5 hectares of grassland in the hills; c) not available in the mountain area	270 EUR/ha
Organic farming	a) To respect the provision of EU Regulation 834/2007, on the entire farm UAA for 5 years; b) to convert all the farm UAA to the organic farming within 5 years since the entry in the scheme	a) at least 1 hectare of UAA in the plain; b) at least 0.5 hectares of UAA in the mountains and the hills	160-570 EUR/ha for maintenance (depending on the crop); 174-620 EUR/ha for conversion (depending on the crop)

Notes: Other agri-environmental schemes were *214_b* (low input production), *214_f* (landscape conservation), *214_i* (biodiversity in paddy fields), *214_l* (biodiversity in mountain pastures), *214_m* (sod seeding and minimum tillage), *214_h* (livestock biodiversity).

Table 2 – AEMs implementation in Lombardy in 2012

Measure	*Crops diversification*	*Grassland maintenance*	*Organic farming*	*Other AEMs*
Farm participants in 2012 (n.)	1,433	1,926	371	4,962
- of which entered in 2007			*112*	
- of which entered in 2008	*571*	*1,000*	*107*	*1,210*
- of which entered in 2009	*342*	*544*	*75*	*748*
- of which entered in 2010	*256*	*201*	*27*	*367*
- of which entered in 2011	*212*	*137*	*38*	*2,011*
- of which entered in 2012	*52*	*44*	*12*	*626*
Area enrolled in 2012 (hectares)	45,846	20,410	8,528	147,247
Payments 2012 (EUR)	7,804,933	5,510,735	2,149,083	31,415,703

Source: Authors calculation based on SIARL (See text).

Table 3 – Description of outcome variables

Measure	Outcome variable	Definition	Unit of measurement
Crops diversification	Main arable crop (ha)	The area covered by the main arable crop in the farm	Number of hectares
	Main arable crop (%)	The share of the main arable crop on total farm arable crop area	%
	Number of arable crops	The number of arable crops in the farm	Number
	Arable crops heterogeneity	Th Gini index of heterogeneity of arable crops in the farm	Number
	Nitrogen-fixing crops	The share of nitrogen-fixing crops on total farm arable crop area	%
	Non soil-depleting crops	The share of nitrogen fixing crops, other arable crops different from winter cereals and set-aside on total farm arable crop area	%
Grassland maintenance	Grassland (ha)	The area covered by grassland in the farm	Number of hectares
	Grassland (%)	The share of grassland on total farm UAA	%
Organic farming	Organic farming (ha)	The UAA under organic farming in farm	Number of hectares
	Organic farming - under conversion (ha)	The UAA in conversion to organic farming in the farm	Number of hectares

Table 4 - Number of matched and unmatched farms in the treated and control groups with CEM

		Full Set of Covariates			**Excluding Previous Participation**		
		Non Treated	Treated	Total	Non Treated	Treated	Total
Crops diversification	All	24,960	1,113	26,073	24,960	1,113	26,073
	Matched	10,557	796	11,353	15,078	1,010	16,088
	Unmatched	14,403	317	14,720	9,882	103	9,985
	Share matched	**42.3%**	**71.5%**	**43.5%**	**60.4%**	**90.7%**	**61.7%**
Grassland maintenance	All	27,810	1,668	29,478	27,810	1,668	29,478
	Matched	11,016	1,208	12,224	16,905	1,545	18,450
	Unmatched	16,794	460	17,254	10,905	123	11,028
	Share matched	**39.6%**	**72.4%**	**41.5%**	**60.8%**	**92.6%**	**62.6%**
Organic farming	All	33,517	272	33,789	33,517	272	33,789
	Matched	4,861	84	4,945	11,615	225	11,840
	Unmatched	28,656	188	28,844	21,902	47	21,949
	Share matched	**14.5%**	**30.9%**	**14.6%**	**34.7%**	**82.7%**	**35.0%**
Other AEMs	All	30,648	3,141	33,789	30,648	3,141	33,789
	Matched	14,547	2,381	16,928	17,298	2,803	20,101
	Unmatched	16,101	760	16,861	13,350	338	13,688
	Share matched	**47.5%**	**75.8%**	**50.1%**	**56.4%**	**89.2%**	**59.5%**

Source: Authors calculation based on data describe in the text.

Table 5 – Estimated additional effects for crops diversification measure

Crops diversification measure	CEM with previous participation			CEM without previous participation		
	Mean Difference in Treated Farms	*Mean Difference in non-Treated Farms*	ATT	*Mean Difference in Treated Farms*	*Mean Difference in non-Treated Farms*	ATT
Main arable crop(ha)	1.522	1.985	**-0.462** **(0.499)**	1.318	1.292	**0.0255** **(0.472)**
Main arable crop(%)	-0.964	10.384	**-11.35***** **(0.839)**	-0.487	5.565	**-6.052***** **(0.696)**
Number of arable crops	-0.158	-0.835	**0.677***** **(0.0571)**	-0.152	-0.504	**0.352***** **(0.0492)**
Arable crops heterogeneity	0.690	-12.205	**12.90***** **(0.929)**	0.248	-7.204	**7.453***** **(0.793)**
Leguminous crops(%)	-1.077	-3.033	**1.956*** **(1.124)**	-0.852	-0.550	**-0.302** **(0.879)**
Non soil-depleting crops(%)	-2.414	1.075	**-3.490***** **(1.013)**	-1.615	-1.253	**-0.362** **(0.942)**
Grassland(ha)	0.307	-0.252	0.558*** (0.103)	0.305	-0.193	0.498*** (0.111)
Grassland(%)	1.078	-0.358	1.436*** (0.278)	1.026	-0.152	1.178*** (0.264)
Organic farming(ha)	0.019	0.421	-0.402* (0.243)	-0.370	0.040	-0.410** (0.181)
Organic farming - under conversion(ha)	0.023	-0.016	0.0396 (0.118)	-0.056	-0.014	-0.0425 (0.0618)

Notes: Figures in the tables show for each of the considered outcome variable the mean difference over the period 2005-2012 in both the samples of treated and non-treated matched farms, as well as the resulting ATT. In bold are reported direct estimated effects, in thin cross-over effects. Estimations are based on d-i-d CEM implemented in STATA13 (command cem). Standard errors are in parentheses. ***, **, or * denote statistical significance at 1%, 5% and 10% level, respectively.

Table 6 – Estimated additional effects for grassland maintenance measure

Grassland maintenance measure	CEM with previous participation			CEM without previous participation		
	Mean Difference in Treated Farms	*Mean Difference in non-Treated Farms*	ATT	*Mean Difference in Treated Farms*	*Mean Difference in non-Treated Farms*	ATT
Grassland(ha)	0.582	-1.737	**2.319***** **(0.185)**	0.473	-0.221	**0.694***** **(0.118)**
Grassland(%)	1.863	-3.941	**5.804***** **(0.540)**	1.633	0.269	**1.364***** **(0.393)**
Main arable crop(ha)	1.837	1.384	0.453 (0.367)	1.843	1.592	0.251 (0.379)
Main arable crop(%)	-2.127	-1.452	-0.676 (0.832)	-1.457	1.386	-2.843*** (0.696)
Number of arable crops	-0.131	-0.251	0.120*** (0.0368)	-0.129	-0.325	0.196*** (0.0339)
Arable crops heterogeneity	-0.482	-1.538	1.056 (0.650)	-0.927	-4.120	3.193*** (0.563)
Leguminous crops(%)	1.673	1.927	-0.254 (0.589)	1.655	0.483	1.172** (0.545)
Non soil-depleting crops(%)	-1.214	-2.745	1.531 (0.931)	-0.837	-1.981	1.144 (0.779)
Organic farming(ha)	-0.140	0.077	-0.217*** (0.0802)	-0.517	0.133	-0.650*** (0.187)
Organic farming - under conversion(ha)	0.000	0.011	-0.0106 (0.0474)	0.035	0.032	0.00308 (0.0522)

Notes: Figures in the tables show for each of the considered outcome variable the mean difference over the period 2005-2012 in both the samples of treated and non-treated matched farms, as well as the resulting ATT. In bold are reported direct estimated effects, in thin cross-over effects. Estimations are based on d-i-d CEM implemented in STATA13 (command cem). Standard errors are in parentheses. ***, **, or * denote statistical significance at 1%, 5% and 10% level, respectively.

Table 7 – Estimated additional effects for organic farming measure

Organic farming measure	CEM with previous participation			CEM without previous participation		
	Mean Difference in Treated Farms	*Mean Difference in non-Treated Farms*	ATT	*Mean Difference in Treated Farms*	*Mean Difference in non-Treated Farms*	ATT
Organic farming(ha)	4.715	-1.065	**5.780***** **(1.180)**	1.17	0.032	**1.138***** **(0.389)**
Organic farming - under conversion(ha)	3.198	-1.245	**4.443***** **(0.631)**	2.275	-0.012	**2.289***** **(0.134)**
Main arable crop(ha)	2.943	1.170	1.772 (1.246)	1.505	-0.15	1.656** (0.728)
Main arable crop(%)	0.499	-2.306	2.805 (4.458)	5.799	2.769	3.030 (2.597)
Number of arable crops	-0.167	-0.255	0.0887 (0.153)	-0.12	-0.196	0.0769 (0.0884)
Arable crops heterogeneity	-6.654	-4.587	-2.067 (2.672)	-4.978	-3.059	-1.919 (1.553)
Leguminous crops(%)	-1.058	-2.363	1.305 (4.266)	-1.119	-0.612	-0.507 (1.911)
Non soil-depleting crops(%)	-4.588	-7.101	2.514 (5.085)	2.763	1.702	1.061 (2.880)
Grassland(ha)	-0.029	-0.041	0.0119 (0.236)	0.39	-0.169	0.560*** (0.174)
Grassland(%)	0.905	1.433	-0.528 (1.454)	1.454	0.588	0.866 (0.837)

*Note*s: Figures in the tables show for each of the considered outcome variable the mean difference over the period 2005-2012 in both the samples of treated and non-treated matched farms, as well as the resulting ATT. In bold are reported direct estimated effects, in thin cross-over effects. Estimations are based on d-i-d CEM implemented in STATA13 (command cem). Standard errors are in parentheses. ***, **, or * denote statistical significance at 1%, 5% and 10% level, respectively.

Table 8 – Estimated additional effects for other AEMs

Other AEMs	CEM with previous participation			CEM without previous participation		
	Mean Difference in Treated Farms	*Mean Difference in non-Treated*	ATT	*Mean Difference in Treated Farms*	*Mean Difference in non-Treated*	ATT
Main arable crop(ha)	0.813	1.024	-0.211 (0.414)	0.818	0.899	-0.0802 (0.361)
Main arable crop(%)	4.798	4.175	0.622 (0.930)	4.698	6.367	-1.669** (0.790)
Number of arable crops	-0.137	-0.393	0.256*** (0.0286)	-0.129	-0.275	0.147*** (0.0264)
Arable crops heterogeneity	-2.643	-5.852	3.209*** (0.450)	-2.570	-4.663	2.093*** (0.417)
Leguminous crops(%)	-0.831	0.577	-1.408*** (0.546)	-1.621	0.700	-2.321*** (0.495)
Non soil-depleting crops(%)	3.167	0.754	2.413** (1.006)	3.244	3.710	-0.467 (0.862)
Grassland(ha)	0.022	-0.034	0.0561 (0.0709)	0.010	-0.062	0.0720 (0.0672)
Grassland(%)	-0.046	0.754	-0.800*** (0.226)	-0.057	0.687	-0.743*** (0.228)
Organic farming(ha)	0.545	0.070	0.474*** (0.105)	0.167	0.268	-0.100 (0.129)
Organic farming - under conversion(ha	0.146	-0.009	0.155*** (0.0390)	0.142	0.008	0.134*** (0.0429)

Notes: Figures in the tables show for each of the considered outcome variable the mean difference over the period 2005-2012 in both the samples of treated and non-treated matched farms, as well as the resulting ATT. In bold are reported direct estimated effects, in thin cross-over effects. Estimations are based on d-i-d CEM implemented in STATA13 (command cem). Standard errors are in parentheses. ***, **, or * denote statistical significance at 1%, 5% and 10% level, respectively.

Table 9 - Cost-benefit analysis for treated matched farms after the treatment (2012)

Variable	Outcome variables means for treated farms	Average farmland under costraints (ha)	ATT	Windfall effects	Average AEM payment per farm (€)	AEM payments per hectare under costraints (€/ha)	AEM payments per unit of ATT (€)
	(a)	(b)	(c)	(d=a-c)	(e)	(f=e/b)	(g=e/c)
Main arable crop(ha)	19.148	38.56	-0.462	19.609	6,413	166	13,881
Main arable crop(%)	58.741		-11.35***	70.090			565
Number of arable crops	3.687		0.677***	3.010			9,472
Arable crops heterogeneity	51.351		12.90***	38.451			497
Nitrogen-fixing crops(%)	42.052		1.956*	40.096			3,278
Non soil-depleting crops(%)	75.293		-3.490***	78.782			1,837
Grassland(ha)	13.246	13.25	2.319***	10.927	3,021	228	1,302
Grassland(%)	49.579		5.804***	43.775			520
Organic farming(ha)	17.091	17.09	5.780***	11.310	4,448	260	769
Organic farming - under conversion(ha)	4.057	4.06	4.443***	-0.386	1,255	309	282

Appendix A

Table A.1: Outcome variables in the pre-treatment scenario (2005) – Crops diversification

Crops diversification	All farms			Not participants	Participants	Difference
Number of farms			26,073	24,960	1,113	
	Mean	Min	Max	Mean	Mean	t-test
Main arable crop (ha)	**12.70**	**0.00**	**453.88**	**12.52**	**16.59**	**-7.57*******
Main arable crop (%)	**78.67**	**18.71**	**100.00**	**79.47**	**60.81**	**26.81*******
Arable crops heterogeneity	**28.18**	**0.00**	**87.65**	**27.24**	**49.42**	**-29.58*******
Number of arable crops	**2.37**	**1.00**	**15.00**	**2.31**	**3.75**	**-25.05*******
Nitrogen-fixing crops (%)	**13.76**	**0.00**	**100.00**	**12.33**	**45.80**	**-33.67*******
Non soil-depleting crops (%)	**82.36**	**0.00**	**100.00**	**82.55**	**78.21**	**6.64*******
Grassland (ha)	1.67	0.00	268.77	1.72	0.71	4.52***
Grassland (%)	6.65	0.00	100.00	6.89	1.31	24.48***
Organic farming (ha)	0.35	0.00	569.14	0.35	0.47	-0.58
Organic farming - under conversion (ha)	0.02	0.00	114.73	0.02	0.07	-0.92

Table A.2: Outcome variables in the pre-treatment scenario (2005) – Grassland maintenance

Grassland maintenance	All farms			Not participants	Participants	Difference
Number of farms			29,478	27,810	1,668	
	Mean	Min	Max	Mean	Mean	t-test
Main arable crop (ha)	11.20	0.00	453.88	10.98	14.95	-7.37***
Main arable crop (%)	68.52	0.00	100.00	68.50	68.82	-0.37
Arable crops heterogeneity	24.66	0.00	87.65	24.87	21.23	6.14***
Number of arable crops	2.07	0.00	15.00	2.08	1.99	2.29**
Nitrogen-fixing crops (%)	11.10	0.00	100.00	11.49	4.71	18.47***
Non soil-depleting crops (%)	71.31	0.00	100.00	71.19	73.39	-2.25**
Grassland (ha)	**1.78**	**0.00**	**393.09**	**1.14**	**12.41**	**-39.6*******
Grassland (%)	**10.88**	**0.00**	**100.00**	**8.47**	**50.98**	**-56.77*******
Organic farming (ha)	0.31	0.00	569.14	0.29	0.61	-1.15
Organic farming - under conversion (ha)	0.02	0.00	114.73	0.02	0.03	-0.37

Table A.3: Outcome variables in the pre-treatment scenario (2005): Organic farming

Organic farming	All farms			Not participants	Participants	Difference
Number of farms			33,789	33,517	272	
	Mean	Min	Max	Mean	Mean	t-test
Main arable crop (ha)	9.96	0.00	453.88	9.98	7.85	1.60
Main arable crop (%)	63.89	0.00	100.00	63.99	50.98	5.93***
Arable crops heterogeneity	22.15	0.00	87.65	22.08	30.54	-4.67***
Number of arable crops	1.88	0.00	15.00	1.88	2.35	-3.40***
Nitrogen-fixing crops (%)	11.19	0.00	100.00	11.07	26.30	-7.67***
Non soil-depleting crops (%)	66.46	0.00	100.00	66.51	61.08	2.19**
Grassland (ha)	1.55	0.00	393.09	1.53	4.00	-2.09**
Grassland (%)	9.52	0.00	100.00	9.53	8.73	0.63
Organic farming (ha)	**0.29**	**0.00**	**569.14**	**0.15**	**17.54**	**-5.39*** **
Organic farming - under conversion (ha)	**0.02**	**0.00**	**114.73**	**0.01**	**1.09**	**-3.10*** **

Source: Authors calculation based on data describe in the text.