Estimating the effects of agri-environmental measures using difference-in-difference coarsened exact matching

Danilo Bertoni[a], Daniele Curzi[a,*], Giacomo Aletti[a,b], Alessandro Olper[a,c]

[a] *Department of Environmental Science and Policy, University of Milan, Italy*
[b] *ADAMSS Center, University of Milan, Italy*
[c] *LICOS Center of Institution and Economic Performance, KU Leuven, Belgium*

ARTICLE INFO

ABSTRACT

This paper studies the effect of agri-environmental measures (AEMs) in improving greener farming practices. We focus on the quantification of the effectiveness of AEMs implemented in the Rural Development Programme 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 interesting properties that are worth considering in policy-evaluation analyses. Our empirical analysis focuses on three AEM schemes protecting and enhancing the environment, Crops diversification, Grassland maintenance and Organic farming. Overall, our results suggest that AEMs were apparently effective in improving the farms' environmental performances. However, our preliminary cost-benefit analysis highlights that the costs of implementing this policy, when compared to the additional results obtained, tend to be quite large.

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 have constituted a relevant part of EU Rural Development Policies under the Second Pillar of the CAP. During the 2007–2013 programming period, AEMs absorbed about 22% of the Rural Development Policies' expenses in the entire EU; this corresponds to about EUR 20 billion (European Commission, 2019). 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 kinds 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 – 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 and farmer characteristics.

This paper tries to address this issue by empirically estimating the effects of AEMs on green farming practices using the (Difference-in-Difference [DID]) CEM. This is an innovative matching methodology developed by Iacus et al. (2012, 2019). Previous works in the literature dealing with similar research questions have tried to address these identification problems by relying on propensity score matching (PSM) (see, e.g., Arata and Sckokai, 2016; Chabé-Ferret and Subervie, 2013; Mennig and Sauer, 2019; Pufahl and Weiss, 2009). 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

---

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

Danilo Bertoni[1], Daniele Curzi[1], Giacomo Aletti[1], Alessandro 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 effectiveness of AEMs implemented in the Rural Development Programme 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, coarsened exact matching (CEM). This methodology presents a number of interesting properties that are worth considering in policy-evaluation analyses. Our empirical analysis focuses on three AEM schemes protecting and enhancing the environment, Crops diversification, Grassland maintenance and Organic farming. Overall, our results suggest that AEMs were apparently effective in improving the farms' environmental performances. However, our preliminary cost-benefit analysis highlights that the costs of implementing this policy, when compared to the additional results obtained, tend to be quite large.

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

---

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 have constituted a relevant part of EU Rural Development Policies under the Second Pillar of the CAP. During the 2007-2013 programming period, AEMs absorbed about 22% of the Rural Development Policies' expenses in the entire EU; this corresponds to about EUR 20 billion (European Commission, 2019). 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 – 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 and farmer characteristics.

This paper tries to address this issue by empirically estimating the effects of AEMs on green farmers' practices using the (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 propensity score matching (PSM) (see, e.g., Pufahl and Weiss, 2009; Chabé-Ferret and Subervie, 2013; Arata and Sckokai, 2016; Mennig and Sauer, 2019). 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 has been designed by the authors to provide an improvement over existing matching approaches in the estimation of causal inference, by reducing any imbalance in the covariates between the treated and control units. CEM incorporates properties of the exact matching procedure, but has a peculiar characteristic that distinguishes it from the other matching methods. CEM indeed allows a 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.

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): *i.* 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; *ii.* 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 should 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, representing one of the largest samples ever used for evaluating the effects of AEMs. 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 were 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.

The contribution of our paper is three-fold. First, from a methodology point of view, to the best of our knowledge, this is the first paper in the agricultural economics literature using CEM to estimate the causal effect of a policy participation. Our assessment of the AEMs through the CEM estimator shows some interesting properties of this estimator, which could make it worthy of consideration for future application in this field. 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, our results provide support for the existence of significant *windfall effects* when agri-environmental policy schemes are applied by farmers. As a consequence, the

estimated additional outcomes are often quite limited, especially when compared with the total payments received by farmers adopting AEMs.

The remainder of the paper is organized as follows. Section 2 provides a literature overview on evaluations of the AEMs' environmental and economic effects Section 3 describes the implementation of AEMs in Lombardy and provides an overview on the data and variables used in the analysis. Section 4 explains the applied methodology. Section 5 summarises the results and Section 6 provides a preliminary cost-benefit analysis and discusses the main findings and suggests further developments. Finally, Section 7 draws some conclusions.

2. Related literature

Our paper is related to a large literature investigating AEMs. In particular, given their broad diffusion among EU farmers, AEMs have been widely investigated from many different perspectives since their introduction within the CAP framework, in 1992.[2] Among various strands of literature, a considerable number of studies focused on the empirical quantification of the AEMs' environmental effects (Uthes and Matzdorf, 2013). More specifically, some of these studies considered, in particular, biodiversity aspects (e.g. Kleijn and Sutherland, 2003; Walker et al., 2007; Lindenmayer et al., 2012; Ansell et al., 2016) and the AEMs' impacts on soil (e.g. Marriot et al., 2005; Marconi et al., 2015) and water quality (e.g. Parrot and Burningham, 2008; Balana et al., 2011; Poole et al., 2013).

The estimation of the *ex-post* environmental effects of AEMs is not straightforward. Existing studies are usually affected by methodological difficulties and limitations because of the complex nature of interactions between the uptake of AEMs by farmers and related environmental outcomes

---

[2] Uthes and Matzdorf (2013) group articles related to AEMs as follows: i) individuating factors influencing farmers' willingness to take up AEMs; ii) analysing the institutional and political environment in which AEMs are implemented; iii) developing methods for monitoring and evaluating AEMs; iv) estimating citizens' willingness to pay for environmental services provided by the AEMs; v) quantifying farmers' participation costs; vi) evaluating environmental benefits resulting from the AEMs' implementation; and vii) evaluating the cost-effectiveness of the AEMs; viii) designing more efficient agri-environmental contracts. Our paper is related especially with the last three categories. In what follows, we propose a summary of the main relevant literature, focusing especially on the specific research designs and their main results.

(Primdahl et al., 2010; Carey et al., 2003). Primdahl et al. (2003) state that problems in measuring the AEMs' environmental effects arise from the "lack of linearity and immediacy of environmental effects, unequivocal causalities and often high costs of measurement."

The lion's share of *ex-post* environmental evaluations relies on field experiments or farm surveys (Primdahl et al., 2010), which are usually based on the analysis of small-scale samples (see for a review Kleijn and Sutherland, 2003; OECD. 2005; Matzdorf and Lorenz; 2010; Balana et al. 2011; Ansell et al., 2016). This leads to a general lack of evidence on large-scale samples, which may benefit a more general assessment of the effectiveness of AEMs. A notable exception is represented by Chabé-Ferret and Subervie (2013), who evaluated the environmental effects of AEMs on a large sample of about 3,500 French farms.

Related to the analysis, other studies focus on the AEMs' impacts on economic outcomes. These works, which only marginally consider environmental effects, have mostly exploited large datasets and considered different type of outcomes. The effect of AEMs has been analysed considering, for instance, farm incomes (Arata and Sckokai, 2016), farm productivities (Mennig and Sauer, 2019; Mary, 2013; Salhofer and Streicher, 2005; Lansink et al., 2002), the use of farm inputs (Laukkanen and Nauges, 2014; Pufahl and Weiss, 2009) and farm employment (Arata and Sckokai, 2016; Petrick and Zier, 2011). Notably, Arata and Sckokai (2016) in their analysis evaluated also the effect of AEMs on environmental outcomes, such as the number of crops grown on the farm and the proportion of grassland in the farmland. The effect of AEMs on the proportion of grassland in the farm was also considered by Pufahl and Weiss (2009). In performing such analyses, some of these studies took advantage of the Farm Accountancy Data Network dataset (Arata and Sckokai, 2016; Mary, 2013), while others resorted to administrative datasets (Mennig and Sauer; 2019; Chabé-Ferret and Subervie; 2013; Pufahl and Weiss, 2009), as is done in the present analysis.

Considering studies that are closer to our analysis, Chabé-Ferret and Subervie (2013) found that generally AEMs provide additional effects to varying degrees, even if they are sometimes attenuated by the presence of windfall effects. Particularly in the organic farming measure, additional effects

largely exceed windfall effects, which are negligible. Similarly, additional effects prevail on windfall effects for the AEM incentivizing the cultivation of cover crops, while an opposite situation was found for the measure promoting the maintenance of buffer strips. Significant results were detected also for schemes seeking to increase crops diversification (+0.65/+0.85 crops cultivated in the farms participating in the two schemes analysed). Arata and Sckokai (2016) estimated the significant effects of AEMs in increasing the number of cultivated crops among English and Italian farms (respectively +0.28 and +0.32), and in enhancing the share of grassland in French farms (+2.75%). Finally, Pufahl and Weiss (2009) found that participation in AEMs increased by 8.8% the share of grassland among treated farms in Germany.

Our main results appear broadly in line with these previous findings, and particularly with the ones of Chabé-Ferret and Subervie (2013), though with some caveats. On average, we show that the effect of the considered AEMs go in a direction consistent with the policy targeting. When estimating additional and windfall effects, for two out of three AEMs (i.e. Grassland Maintenance and Organic Farming), the magnitudes of additional effects are often small, suggesting the existence of important windfall effects and less than satisfactory cost-effective measures.

3. Data and variables

AEMs represented the main policy measure of the RDP 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 the 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, generally providing per hectare payments to farmers, who maintain or convert to some farmland uses or farming methods for a period of at least 5 years. 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 the 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. Such a measure encouraged farmers to adopt rotation of arable crops. Its introduction was motivated to overcome environmental problems arising from the widespread diffusion of an intensive maize monoculture in Lombardy (in 2005 maize covered 50.5% of Lombardy's arable crops farmland). The second AEM concerns *Grassland maintenance*, paying farmers to preserve farmland devoted to such land use. This scheme, long-established in Lombardy RDPs, seeks to counter the persistent decrease of meadows, especially in the flat area of the Region. Such decline was mainly attributable to a progressive substitution of grass with maize (grain and silage) for livestock feed, with notable negative consequences especially from the biodiversity and landscape perspectives. The third covers *Organic farming*, delivering differentiated payments both for maintenance of and conversion to this farming method.[3]

Note that the last two schemes have their own equivalents in the 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

---

[3] With reference to the Lombardy RDP, the AEMs are identified as: *214_a – crops diversification - 2014_c – grassland maintenance –* and *214_e – organic farming*.
[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.

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 the 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. It is worth noting that the three AEMs examined have been chosen because they are the only ones, whose outcome variables can be quantified for the entire universe of Lombardy farms (participants and non-participants) from administrative datasets.

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 implementation of the AEMs in Lombardy began in 2008, with the first farms adopting this policy, and the 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 that 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. In addition, as shown in Table 2, most of 2012 participants enrolled in the 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

---

[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.

participated in an agri-environmental scheme and a value of 0 when the farm was potentially 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 crops cultivated, the main part of the outcome variables is focused on farmland use.[6] 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.

---

[6] We are aware that farmland use and farming practices represent only the outcomes of the AEMs analysed and not their (actual) environmental effect, as correctly pointed by Oñate et al. (2000). 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.. Hence, as the objective of this paper is to estimate the effect of participation in AEMs for a large sample of farms, variables on farmland use represent a valuable proxy for capturing the (potential) environmental effects of the selected schemes.

Starting from these data we were able to calculate a set of outcome variables for each farm potentially eligible for AEMs.[7]

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 varieties on the farm. In particular, for each farm, we computed the area covered by the main arable crop, the share of the main arable crop of the total arable crop area of the farm, and the number of arable crops present on the farm. In order to assess how much the farmland was diversified, we also quantified for each farm the degree of heterogeneity of arable crops, computed as the Gini Index of Heterogeneity of arable crops. Furthermore, we calculated the share of leguminous crops and non-soil-depleting crops of the 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. There were 22 different leguminous crops and 135 non-soil-depleting ones. This latter count includes, besides leguminous crops, other cereals (maize and sorghum), vegetable crops and other crops useful for their beneficial effects on soil structure. Also, set aside land was included in this last broad 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 –

---

[7] 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).

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 farms' total UAA converted to organic farming and that under conversion[8].

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 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 and 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

---

[8] It is worth noting that in the reference area only a part of organic farmland was covered by the corresponding AEM. Starting with the premise that participation in the measure is voluntary, there are, at least, three possible interpretations explaining such behaviour. Firstly, as farmers under organic farming rules already undergo controls by a private certifying body, some of them do not accept to be also subject to controls by the RDP managing authority (that is a necessary condition for participation in the AEM). The second reason regards farmers with a little managed farmland, who would not be particularly attracted by apply to obtain a modest overall payment (especially when they grow high added value crops), given the per-hectare nature of these subsidies. The third explanation could be linked to the presence of farmers who do not want to be conditioned by the AEM obligation to maintain organic farming commitments for at least 5-years.

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 in the UAA

- livestock density

- farmer's age

- previous participation in AEMs during the 2000-2006 programming period.[9]

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 Tabellini, 2008), others hold the opposite view (Rubin and Thomas, 1996). 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 the CEM is chosen *ex-ante*, rather than discovered *ex-post,* as in PSM, the 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

---

[9] 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.

affect the policy adoption. The complete list of control variables and thresholds used to define strata in the CEM procedure are reported in the Additional Materials section (Table A.5).

3.4 Sample description

Descriptive statistics of the selected outcome variables for the three schemes considered are reported in Appendix 2 (see Tables A.2, A.3 and A.4). 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 the other schemes, these numbers are 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).[10]

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 number of arable crops and the heterogeneity of the arable crops were already higher among future participants, while the share of the main arable crop was already lower. 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 in 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.

---

[10] The main parts of other AEMs have eligibility criteria similar to those for the *Organic farming* scheme.

4. Methodology

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. In particular, this occurs as these matching methodologies operate a simple random sampling by stratifying the sample *ex-post* on the initial covariate space, or based on the propensity score or on the distance metric space (Iacus et al., 2019). In practice, these methodologies approximate matching within each stratum as if it were an exact matching.

Against this background, Iacus et al. (2012; 2019) 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 on. Specifically, they include an assumption about the *ex-ante* stratification of the data. 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 to data analysts.

In short, our empirical analysis is based on the use of 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 simple random sampling axiom given a set of covariates on which treated and untreated units are matched, the properties of CEM allow this methodology to be consistent and coherent with its theoretical underlying axiom developed by Iacus et al. (2019) 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, 2009; 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.

4.1 CEM and identification

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.

In what follows we present concisely the formal logic of CEM. Note that, a non-technical and straightforward description of how CEM works in practice is presented in Appendix 1 of the paper. 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 a 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.

This typical observational problem is addressed by the CEM algorithm through the following main steps.[11] 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 meaningful strata (i.e. in equal intervals of the same size or in intervals of different dimension from each other), 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 (the higher the number of strata), the lower will be the number of matches provided by the CEM, as well as the lower will be the imbalance.

---

[11] Note, since the PSM methodology has been widely described in previous literature, we decided to focus only on the description of the CEM.

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:

$$E(Y_{it}(1) - Y_{it'}(0)|T = 1, X) - (Y_{it}(0) - Y_{it'}(0)|T = 0, X).$$

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.[12] 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.[13]

---

[12] Note that the parallel line assumption is common to all the DiD application. Moreover, it is worth mentioning that this assumption considers the outcome variable only. Therefore, it does not affect the way treated and untreated unit are matched.

[13] 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).

5. Results

5.1 CEM versus PSM: a comparative analysis

In order to test the properties of CEM, we conducted a direct comparison with PSM. PSM, so far, has been one of the most widely methodologies used by other researchers to evaluate the effect of AEMs and other agricultural policy outcomes. 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 the CEM (based on the same specification as used in the empirical analysis) with the nearest neighbour PSM.[14, 15]

Our exercise compares the level of imbalance obtained through the use of CEM and PSM, together with the estimated ATT. Following Iacus et al. (2019), we run our demonstrative analysis through a randomization of 500 CEM and 500 PSM. We then compare the results obtained in terms of their level of imbalance and ATT.[16] For the CEM, this randomization 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.[17] The randomization for the PSM is 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 1 plots the results of this test, where on the vertical axis we show the ATT obtained with our 500 CEM and PSM randomizations, while on the horizontal axis we have their respective measures of imbalance (as measured by the L1 distance).[18] The imbalance represents a measure of

---

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

[15] For the purpose of this analysis, we used a nearest neighbour propensity score matching with one neighbour, with replacement.

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

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

[18] 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).

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 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. 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 randomizations plotted in Figure 1, some interesting 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). The CEM estimations show a level of imbalance that is either much higher (between 0.5 and 0.8) or lower (between 0.1 and 0.3) than those obtained with PSM, which are concentrated around values of 0.4-0.5. From this perspective, although some of the estimations obtained through CEM show level of imbalance particularly low, this exercise does not allow to reach any conclusive evidence on which of the two approaches (CEM and PSM) is able to address the imbalance better.

The most striking result emerging from this analysis, however, concerns the ATT estimations. The variability of the results obtained with our randomized CEM (included in the black circle) is significantly lower than that obtained with the PSM (included in the red circle). This finding suggests that, in this specific case, the high variability of the results obtained with the PSM may lead the ATT estimation 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, suggesting almost a uniquely negative ATT, no matter how the model is specified, with only very few exceptions showing positive effect. It is worth noting that PSM estimations with the higher level of imbalance (higher than 0.5) present the most heterogeneous results. The results of these estimations are indeed either positive or negative, and present a magnitude much higher than those obtained, on average, from PSM estimations with lower imbalance.

To summarise the findings emerging from our comparative test we can argue that the CEM has shown some relevant characteristics that should be worth considering. While from our exercise does not clearly emerge a systematic capability of CEM of obtaining lower level of imbalance than PSM, although some results look promising, it is worth noting that 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 the CEM.

Despite our tests highlighting some important characteristics of the CEM, we are not arguing that the use of the 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, first, when considering the use of matching methodologies for empirical analysis, it is worth considering the use of more than one estimator, as different methodologies may lead to significantly different results. Second, it is of fundamental importance to find the more suitable specification of the estimation model that allows having the lowest level of imbalance possible. This is important, in general, to minimize the bias in the estimations. From the exercise proposed in this section, however, it emerges that, for the PSM estimations more than for the CEM ones, the reliability of the estimated ATT is strictly influenced by this parameter.

5.2 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 the CEM. Two sets of estimates are considered: one with the full set of farms and one where all treated and untreated farms participating in any AEMs in the previous programming period were dropped from the database. The choice of running this second set of estimations is driven by the fact that the value of the pre-treatment outcomes of farms participating in AEMs during the 2000-2006 programming period may have been indeed affected from their

participation. From this perspective, the estimated ATT of the DiD CEM may be biased. However, it is worth recalling that the first set of estimations includes the set of control variables from the farms' previous participations in AEMs. Although controlling for this issue is of relevant importance for the estimations, the inclusion of these variables does not exclude the farms' participations in the 2005 AEMs' contributions to the construction of the counterfactual of treated farms. Thus, to get rid of any potential bias arising from a farm's previous participation, we ran a second set of estimations where we dropped all farms participating in previous AEMs. However, the exclusion of these farms leads to a massive loss of data. Therefore, to exclude potential sample selection bias, we decided to show the results obtained through both sets of estimations.

Looking at the CEM results obtained using the full set of farms, the share of the matched, non-treated farms is around 40-50% for all the AEMs considered, 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 not 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 farms participating in previous AEMs the number of matched farms, in particular in the treated sample, reduces dramatically. As shown by the figures reported in Table 4, both sets of estimates present the number of matched non-treated farms, which is considerably higher than the number of treated farms, and so ensures a good representativeness for the non-treated (counterfactual) sample.

Tables 5, 6, 7 and 8 present results of estimating the ATT of the different AEMs considered on a set of outcome variables by using DiD CEM. Figures in the tables show the average variation for each outcome variable for the period 2005-2012, for both samples of treated and non-treated farms matched with the CEM. This information is 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 of the sample of treated matched farms.

The estimated (positive or negative) ATT may arise from the difference in variations between treated and controls 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 ATTs 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 not in bold).

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 the CEM, including a full set of control variables, as discussed in Section 3.3. To check for the potential existence of a bias arising from farms having participated in AEMs during the previous programming period, we present in the last three columns the results obtained excluding from the data all the treated and untreated farms with previous participation in AEMs. However, because qualitatively the estimated ATTs are robust in the two sets of data, for the sake of simplicity in what follows we focus the discussion on the results obtained with the first set of estimates (columns 1).

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% 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 cultivated arable crops (ATT = 0.68), in line with Chabé-Ferret and Subervie (2013) and Arata and Sckokai (2016), 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, for the period 2005-2012. Thus, the measure attenuates a 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 in the outcome for the non-treated group. The proportion of leguminous 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%).

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 of UAA (ATT = 5.8%); both significant at the 1% level. Such result are in line with those found by Arata and Sckokai (2016) and Pufahl and Weiss (2009). Quantitatively, these are relevant effects, even if they derive mainly from the counterbalance of a dominant trend of decreasing values in permanent grassland rather than a net increasing effect in treated farms. Concerning potential cross-over effects, participation in such a scheme seems to lead farmers to significantly reduce 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 AEM schemes 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 the heterogeneity of arable crops 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, participation in *Other AEMs* leads farms to a higher increase in their organic farming area, both expressed as number of hectares and under conversion.

One of the main advantage of DiD estimations is the possibility to use the outcome levels in the pre-treatment period as covariates (i.e. year 2005). As a robustness check, we indeed exploit this possibility in our analysis. This test allows further refining the match between treated and controlling farms, by introducing as an additional criterion their similarities in the level of the considered outcome variable. The results of this test, presented in the Additional Materials section (Table A.10), largely confirm those previously presented in this section. It is interesting to note that, on average, the estimated ATTs in this set of estimations are quantitatively slightly higher than those obtained when the outcome variable is not included in the set of control variables. The case of the effect of the *Crops diversification* measure on the percentage of leguminous crops is particularly relevant in this respect. The estimated ATTs in the full sample and in the reduced sample are about four times, and twice as high in magnitude than those showed in Table 5.[19]

---

[19] Understanding the reason for the general slightly higher magnitude of the ATT estimated when including the outcome variable in the set of control variables is beyond the scope of this analysis. Our robustness check just aims to prove the

5.3 Robustness check: ATT based on DiD PSM

To check the consistency of our results to the use of difference matching estimators, we run all the estimations from Table 5 to Table 8 using a DiD PSM instead of the CEM. Among the various PSM estimators our choice fell on the use of a nearest neighbour matching with one neighbour, with replacement.[20] The results obtained using the PSM are presented in Additional Materials (tables from A.6 to A.9). In most cases, the estimated ATTs are qualitatively similar to the ones obtained with the CEM. However, some significant differences with respect to the ATT estimated with the CEM are worthy of consideration. Focusing on the outcome variables that we considered as the main targets of the different schemes, two results change significantly using PSM. First, the effect of the *Crops diversification* measure on the *main arable crop* area (see Table A.6). Using the PSM on the full sample, the estimated effect is negative and significant and higher in magnitude than that obtained with the CEM. The second main difference is related to the ATT of *Crops diversification* on the percentage of non-soil-depleting crops (see Table A.6). Here, the coefficient for this variable in the full sample of farms is no longer significant.

Moreover, we detect also some differences between the CEM and PSM when considering the magnitude of the estimated ATT. Three main findings are worthy of attention in this respect: the first two cases are related to the effect of the Crops diversification measure on the percentage of leguminous crops in the full sample of farms and on the percentage of non-soil-depleting crops in the reduced sample, respectively (Table A.6). The third case concerns the effect of the Organic farming measure on the UAA under organic farming estimated on the full sample.

Summarizing, the results obtained with the PSM do not alter the main message coming from the analysis carried out with the CEM on the effect played by AEMs in affecting farmers'

---

consistency of the results when including the level of the outcome variable in the pre-treatment period as an additional criterion to match treated and control farms.

[20] Our choice to adopt a nearest neighbour PSM is motivated by the fact that, given that the CEM is based on exact matching, we argue that this PSM estimator allows the most accurate comparison with the CEM, as it matches any treated farm with the closest untreated one, based on the estimated propensity score.

environmentally greener practices. However, our findings highlight one more time the importance of considering different matching estimators in this kind of analysis.

6. 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 estimating the actual effect, i.e. the ATT, it is important to evaluate the result taking into consideration the cost of (policy) implementation. In what follows, starting from the estimations presented above, we have performed an assessment of the windfall effects and preliminary cost-benefit analysis. In such a preliminary cost-benefit analysis we do not pretend to compare the cost of the policies with their environmental effects (a difficult task in a large sample like ours, as pointed out in section 3.2), but we aim to quantify the money spent to obtain an additional agri-environmental outcome by farmers. This allows for a better understanding of the extent to which the moneys spent to implement the policy have produced the desired outcomes. However, we are aware that, theoretically, even a measure with a small additional effect, and a high cost per unit of ATT, could be still cost-effective. However, we underline that increasing the level of cost per unit of ATT reduce the AEMs' overall potential social benefits.

From a conceptual 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' effects on *PS*, we do not have information on the impact on farm profits 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.[21] 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.

---

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

Our preliminary 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 a counterfactual situation, under the hypothesis 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.

As many of 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. Therefore, following Chabé-Ferret and Subervie (2013), for the sake of simplicity we applied the preliminary cost-benefit analysis only on the AEMs that display three characteristics: (1) they are targeted to just one goal, (2) they aim to increase the farmland allocated to a specific use or managed with a specific farming method, (3) they are related to farmland use and can be quantified by hectares of land. As only the *Grassland maintenance* and *Organic farming* AEMs meet these criteria, a cost-benefit analysis will be performed for these two measures, making exclusive reference to the outcome variables directly related to the number of hectares under their commitments.

Table 9 presents the result of this analysis. In row (a) we report for treated matched farms the average farmland subject to the practice after the treatment (in 2012). Such value corresponds to the outcome variables directly linked to the goal of the measure, indicating the number of hectares devoted to grassland and those to organic farming. Next, we compute windfall effects (row c) as the difference between the post-treatment observed level of the outcome variables in the treated matched farms (row a) and the estimated ATTs (row b). In a second step, we considered for each scheme the average annual farm payment received by treated matched farms (row d), their average area under agri-environmental commitments in 2012 (row a) and the estimated ATT (row b). Using these values,

we compute the average payment per hectare under the constraints (row e) and the average payment per unit of additional effect (row f).

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 3,020 per farm), generated by the voluntary uptake of the measure on 13.25 ha, in reality has contributed to obtaining an additional effect of only 2.32 ha, with a payment for additional units of ATT of EUR 1,302/ha, against a payment of EUR 228/ha under the contract.

Similarly, considering the case of the *Organic farming* scheme the interpretation of windfall and additional effects is 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 769/ha against an average payment per area under contract of EUR 260/ha. Interestingly, payments for conversion to organic farming in practice determine only the additional effects, in line with Chabé-Ferret and Subervie (2013). Such a result, if compared to that of *Grassland maintenance*, may suggests that AEMs directed to improving farming methods or practices could be more efficient than those paying for preserving a specific farmland use.

In summary, from this stylized cost-benefit analysis the main conclusion that can be drawn is that the AEMs investigated 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., 2019). Indeed, the impact of AEMs on job creation, though often positive, is not enough to reverse our cost-benefit analysis conclusion.

7. Conclusions

This paper provides an *ex-post* evaluation of the direct and cross-over effects of three AEMs implemented in the 2007-2013 RDP programme 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. The empirical analysis was based on the use of an innovative matching procedure called Coarsened Exact Matching (CEM). This methodology allows the causal effect of AEMs to be assessed for the selected outcome indicators. 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 interesting properties of the 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 the CEM with the PSM through a randomization process, the former allows for having, on average, more stable results.

As regards the effect of AEMs on green farming practices, our results provide evidence that the AEMs considered were apparently effective in improving the environmental performances of 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 (additional) effects is 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, our stylized cost-benefit analysis highlights that the cost of the implementation of this policy appears to be extraordinarily high.

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 the 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 the CEM cannot be considered an option for a policy-evaluation analysis.

References

Abadie, A., Imbens, G.W., 2006. Large sample properties of matching estimators for average treatment effects. Econometrica, 74(1), 235-267. https://doi.org/10.1111/j.1468-0262.2006.00655.x

Ansell, D., Freudenberger, D., Munro, N., Gibbons, P., 2016. The cost-effectiveness of agri-environment schemes for biodiversity conservation: A quantitative review. Agric Ecosyst Environ, 225, 184-191. https://doi.org/10.1016/j.agee.2016.04.008

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

Balana, B.B., Vinten, A., Slee, B., 2011. A review on cost-effectiveness analysis of agri-environmental measures related to the EU WFD: Key issues, methods, and applications. Ecol Econ, 70(6), 1021-1031. https://doi.org/10.1016/j.ecolecon.2010.12.020

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

Carey, P.D., Short, C., Morris, C., Hunt, J., Priscott, A., Davis, M., Finch, C., Curry, N., Little, W., Winter, M., Parkin, A., Firbank, L.G., 2003. The multi-disciplinary evaluation of a national agri-environment scheme. J Environ Manage, 69(1), 71-91. https://doi.org/10.1016/S0301-4797(03)00120-8

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

European Commission, 2019. Agriculture and rural development. Agri-environment measures.
https://ec.europa.eu/agriculture/envir/measures_en. Accessed 15 April 2019

Evans, N.J., Morris, C., 1997. Towards a geography of agri-environmental policies in England and Wales.
Geoforum, 28(2), 189-204. https://doi.org/10.1016/S0016-7185(97)00003-1

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

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. 2019. Jobs and Agricultural Policy: Impact of the Common
Agricultural Policy on EU Agricultural Employment. *Food Policy*. (Forthcoming).

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., G. King, and G. Porro. 2008. Matching for causal inference without balance checking. http://gking.harvard.edu/files/cem.pdf.

Iacus, S.M, King, G. and Porro, G. (2019). A Theory of Statistical Inference for Matching Methods in Causal Research. Political Analysis 27: pp. 46–68. DOI: 10.1017/pan.2018.29

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

Kleijn, D., Sutherland, W.J., 2003. How effective are European agri-environment schemes in conserving and promoting biodiversity?. J Appl Ecol, 40(6), 947-969. https://doi.org/10.1111/j.1365-2664.2003.00868.x

Lansink, A.O., Pietola, K., Bäckman, S., 2002. Efficiency and productivity of conventional and organic farms in Finland 1994–1997. Eur Rev Agric Econ, 29(1), 51-65. https://doi.org/10.1093/erae/29.1.51

Laukkanen, M., Nauges, C., 2014. Evaluating greening farm policies: a structural model for assessing agri-environmental subsidies. Land Econ, 90(3), 458-481. https://doi: 10.3368/le.90.3.458

Lindenmayer, D.B., Zammit, C., Attwood, S.J., Burns, E., Shepherd, C.L., Kay, G., Wood, J., 2012. A novel and cost-effective monitoring approach for outcomes in an Australian biodiversity conservation incentive program. PLoS One, 7(12) https://doi.org/10.1371/journal.pone.0050872

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

Marconi, V., Raggi, M., Viaggi, D., 2015. Assessing the impact of RDP agri-environment measures on the use of nitrogen-based mineral fertilizers through spatial econometrics: The case study of Emilia-Romagna (Italy). Ecol Indic, 59, 27-40. https://doi.org/10.1016/j.ecolind.2015.02.037

Marriott, C.A., Bolton, G.R., Fisher, J.M., Hood, K., 2005. Short-term changes in soil nutrients and vegetation biomass and nutrient content following the introduction of extensive management in upland sown swards in Scotland, UK. Agric Ecosyst Environ, 106(4), 331-344. https://doi.org/10.1016/j.agee.2004.09.004

Mary, S. (2013). Assessing the impacts of pillar 1 and 2 subsidies on TFP in French crop farms. J Agric Econ, 64(1), 133-144. https://doi.org/10.1111/j.1477-9552.2012.00365.x

Matzdorf, B., Lorenz, J., 2010. How cost-effective are result-oriented agri-environmental measures? -An empirical analysis in Germany. Land use policy, 27(2), 535-544. https://doi.org/10.1016/j.landusepol.2009.07.011

Mennig, P., Sauer, J., 2019. The impact of agri-environment schemes on farm productivity: a DID-matching approach, Eur. Rev. Agric. Econ, jbz006. https://doi.org/10.1093/erae/jbz006

OECD, 2005. Evaluating Agri-environmental Policies: Design, Practice and Results, OECD Publishing, Paris, https://doi.org/10.1787/9789264010116-en.

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

Oñate, J.J., Andersen, E., Peco, B., Primdahl, J., 2000. Agri-environmental schemes and the European agricultural landscapes: the role of indicators as valuing tools for evaluation. Landsc Ecol, 15(3), 271-28. https://doi.org/10.1023/A:1008155229725

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

Parrott, A., Burningham, H., 2008. Opportunities of, and constraints to, the use of intertidal agri-environment schemes for sustainable coastal defence: A case study of the Blackwater Estuary, southeast England. Ocean & Coastal Management, 51(4), 352-367.
https://doi.org/10.1016/j.ocecoaman.2007.08.003

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

Petrick, M., Zier, P., 2011. Regional employment impacts of Common Agricultural Policy measures in Eastern Germany: a difference-in-differences approach. Agric Econ, 42(2), 183-193.
https://doi.org/10.1111/j.1574-0862.2010.00509.x

Poole, A.E., Bradley, D., Salazar, R., Macdonald, D.W., 2013. Optimizing agri-environment schemes to improve river health and conservation value. Agric Ecosyst Environ, 181, 157-168.
https://doi.org/10.1016/j.agee.2013.09.015

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

Primdahl, J., Vesterager, J.P., Finn, J.A., Vlahos, G., Kristensen, L., Vejre, H., 2010. Current use of impact models for agri-environment schemes and potential for improvements of policy design and assessment. J Environ Manage, 91(6), 1245-1254. https://doi.org/10.1016/j.jenvman.2009.12.012

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

Rubin, D., Thomas, N. 1996. Matching Using Estimated Propensity Scores: Relating Theory to Practice. *Biometrics* 52 (1): 249–64.

Salhofer, K., Streicher, G., 2005. Self-selection as a problem in evaluating agri-environmental programs. In: K.M. Ortner (ed.), Assessing Rural Development Policies of the Common Agricultural Policies. Selection of papers from the 87th EAAE-Seminar. Kiel: Vauk, 203–213.

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

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

Walker, K.J., Critchley, C.N.R., Sherwood, A.J., Large, R., Nuttall, P., Hulmes, S., Rosed, R., Mountford, J.O., 2007. The conservation of arable plants on cereal field margins: an assessment of new agri-environment scheme options in England, UK. Biol Conserv, 136(2), 260-270. https://doi.org/10.1016/j.biocon.2006.11.026

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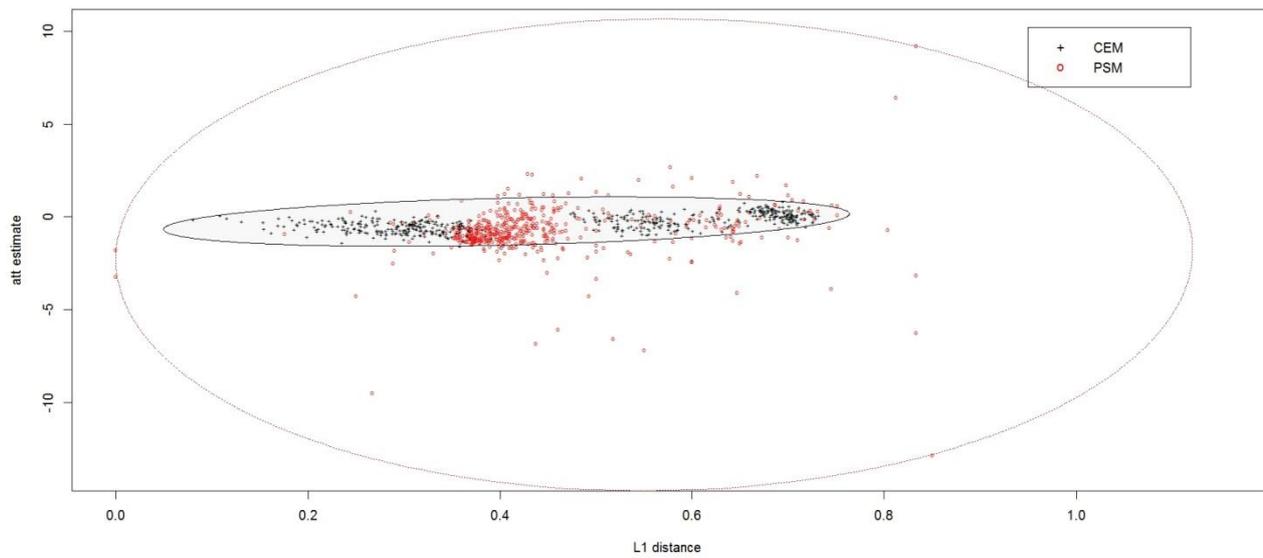
Note: The figure plots ATT and the level of imbalance (L1 distance) for 500 simulations of CEM and PSM. CEM results are identified with black crosses, while red circles show the results for PSM. CEM results are included in the black ellipse, while PSM results are shown within the red ellipse.

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 leguminous 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
	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
Crops diversification	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 (ha)	The area covered by grassland in the farm	Number of hectares
Grassland maintenance	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 Farms			Exluding farms with previous participation		
		Non Treated	Treated	Total	Non Treated	Treated	Total
Crops diversification	All	24,960	1,113	26,073	20,802	275	21,077
	Matched	10,557	796	11,353	9,785	252	10,037
	Unmatched	14,403	317	14,720	11,017	23	11,040
	Share matched	**42.3%**	**71.5%**	**43.5%**	**47.0%**	**91.6%**	**47.6%**
Grassland maintenance	All	27,810	1,668	29,478	24,143	127	24,270
	Matched	11,016	1,208	12,224	9,720	112	9,832
	Unmatched	16,794	460	17,254	14,423	15	14,438
	Share matched	**39.6%**	**72.4%**	**41.5%**	**40.3%**	**88.2%**	**40.5%**
Organic farming	All	33,517	272	33,789	27,088	47	27,135
	Matched	4,861	84	4,945	4,590	34	4,624
	Unmatched	28,656	188	28,844	22,498	13	22,511
	Share matched	**14.5%**	**30.9%**	**14.6%**	**16.9%**	**72.3%**	**17.0%**
Other AEMs	All	30,648	3,141	33,789	25,792	1,343	27,135
	Matched	14,547	2,381	16,928	13,098	1,211	14,309
	Unmatched	16,101	760	16,861	12,694	132	12,826
	Share matched	**47.5%**	**75.8%**	**50.1%**	**50.8%**	**90.2%**	**52.7%**

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

Table 5 – Estimated additional effects for crops diversification measure

Crops diversification measure	CEM with full set of farms			CEM exluding farms with 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)**	1.097	1.725	**-0.628** **(0.947)**
Main arable crop(%)	-0.964	10.384	**-11.35*** ** **(0.839)**	-5.584	6.033	**-11.62*** ** **(1.330)**
Number of arable crops	-0.158	-0.835	**0.677*** ** **(0.0571)**	0.262	-0.504	**0.766*** ** **(0.0898)**
Arable crops heterogeneity	0.690	-12.205	**12.90*** ** **(0.929)**	5.283	-7.589	**12.87*** ** **(1.514)**
Leguminous crops(%)	-1.077	-3.033	**1.956*** **(1.124)**	6.698	0.464	**6.234*** ** **(1.705)**
Non soil-depleting crops(%)	-2.414	1.075	**-3.490*** ** **(1.013)**	-5.897	-1.593	**-4.304** ** **(1.870)**
Grassland(ha)	0.307	-0.252	0.558*** (0.103)	0.493	0.066	0.427** (0.175)
Grassland(%)	1.078	-0.358	1.436*** (0.278)	1.455	0.392	1.063** (0.507)
Organic farming(ha)	0.019	0.421	-0.402* (0.243)	-0.047	0.013	-0.0597 (0.260)
Organic farming - under conversion(ha)	0.023	-0.016	0.0396 (0.118)	0.066	0.009	0.0564* (0.0303)

*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 6 – Estimated additional effects for grassland maintenance measure

Grassland maintenance measure	CEM with full set of farms			CEM exluding farms with 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)**	4.434	0.142	**4.292***** **(0.256)**
Grassland(%)	1.863	-3.941	**5.804***** **(0.540)**	19.976	1.355	**18.62***** **(1.405)**
Main arable crop(ha)	1.837	1.384	0.453 (0.367)	1.308	1.268	0.0405 (1.271)
Main arable crop(%)	-2.127	-1.452	-0.676 (0.832)	-7.630	2.888	-10.52*** (2.578)
Number of arable crops	-0.131	-0.251	0.120*** (0.0368)	0.063	-0.329	0.392*** (0.128)
Arable crops heterogeneity	-0.482	-1.538	1.056 (0.650)	0.637	-5.547	6.184*** (2.079)
Leguminous crops(%)	1.673	1.927	-0.254 (0.589)	1.196	0.673	0.523 (2.215)
Non soil-depleting crops(%)	-1.214	-2.745	1.531 (0.931)	-6.321	-1.664	-4.658 (3.171)
Organic farming(ha)	-0.140	0.077	-0.217*** (0.0802)	-1.477	0.061	-1.538*** (0.263)
Organic farming - under conversion(ha)	0.000	0.011	-0.0106 (0.0474)	0.000	0.016	-0.0161 (0.0632)

*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 7 – Estimated additional effects for organic farming measure

Organic farming measure	CEM with full set of farms			CEM exluding farms with 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)**	3.430	0.192	**3.238***** **(0.774)**
Organic farming - under conversion(ha)	3.198	-1.245	**4.443***** **(0.631)**	2.392	0.007	**2.385***** **(0.193)**
Main arable crop(ha)	2.943	1.170	1.772 (1.246)	2.349	-0.189	2.538 (2.141)
Main arable crop(%)	0.499	-2.306	2.805 (4.458)	2.245	1.359	0.887 (6.927)
Number of arable crops	-0.167	-0.255	0.0887 (0.153)	0.441	-0.068	0.509* (0.262)
Arable crops heterogeneity	-6.654	-4.587	-2.067 (2.672)	-1.162	-1.548	0.386 (4.262)
Leguminous crops(%)	-1.058	-2.363	1.305 (4.266)	-0.426	-1.214	0.788 (5.193)
Non soil-depleting crops(%)	-4.588	-7.101	2.514 (5.085)	2.521	-1.137	3.657 (8.099)
Grassland(ha)	-0.029	-0.041	0.0119 (0.236)	1.131	0.039	1.091*** (0.332)
Grassland(%)	0.905	1.433	-0.528 (1.454)	0.734	0.901	-0.168 (2.483)

*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 full set of farms			CEM exluding farms with 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)	0.813	1.024	-0.211 (0.414)	0.274	1.250	-0.976 (0.627)
Main arable crop(%)	4.798	4.175	0.622 (0.930)	2.881	6.167	-3.286*** (1.077)
Number of arable crops	-0.137	-0.393	0.256*** (0.0286)	-0.140	-0.343	0.203*** (0.0401)
Arable crops heterogeneity	-2.643	-5.852	3.209*** (0.450)	-2.939	-5.773	2.834*** (0.626)
Leguminous crops(%)	-0.831	0.577	-1.408*** (0.546)	1.184	2.232	-1.048 (0.669)
Non soil-depleting crops(%)	3.167	0.754	2.413** (1.006)	1.348	3.214	-1.866 (1.212)
Grassland(ha)	0.022	-0.034	0.0561 (0.0709)	0.160	0.131	0.0284 (0.0937)
Grassland(%)	-0.046	0.754	-0.800*** (0.226)	0.178	0.727	-0.549* (0.296)
Organic farming(ha)	0.545	0.070	0.474*** (0.105)	0.558	0.111	0.447*** (0.143)
Organic farming - under conversion(ha)	0.146	-0.009	0.155*** (0.0390)	0.114	0.010	0.103** (0.0417)

*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 9 – Assessment of windfall effects and stylized cost-benefit analysis for treated matched farms after the treatment (2012)

Measure		Grassland maintenance	Organic farming (maintenance)	Organic farming (conversion)
Average farmland subject to the practice (ha)	**(a)**	13.25	17.09	4.06
ATT (ha)	**(b)**	2.32	5.78	4.44
Windfall effect (ha)	**(c=a-b)**	10.93	11.31 -	0.38
Average cost per farm subject to the practice (€)	**(d)**	3,020	4,444	1,254
Cost per hectare subject to the practice (€/ha)	**(e=d/a)**	228	260	309
Cost per unit of ATT (€/ha)	**(f=d/b)**	1,302	769	282

Appendix

A non-technical explanation of how CEM works in practice

Section 4.1 of this paper presents the main properties of the coarsened exact matching (CEM). As this methodology is rather new and has never been applied in the field of agricultural economics, we believe that a non-technical explanation of how CEM works in practice may be useful. In what follows we present a numerical example that presents in a simple way the mechanism through which the CEM allows estimation of the causal effect of the participation in a treatment.[22]

Suppose we are interested in analysing the difference in revenue between farms in the Lombardy Region that participate in any AEMs and those that do not participate in any AEMs. Consider now a hypothetical scenario like the one in the illustrative example in Table A.1, where we have eight farms (Farm ID in column A), three of which are subject to the treatment, as shown in column B (ID 1, 5 and 6). Suppose that we are going to use only two variables to match treated and non-treated farms, such as farm size expresses in terms of number of arable hectares (column E) and the age of the leading farmer (column F). We then decide to divide each of the considered variables into two bins. The variable Farm size will include in the first one farms with less than 10 ha, while the second one includes farms with more than 10 ha. The CEM allows only treated and non-treated farms having a size lower than 10 ha to be included in the first bin, and only farms having more than 10 ha to be included in the second one. Considering *Farmer's Age*, farmers less than 50 years will be included in the first bin, while the second one will include farmers more than 50 years old. Also in this case the CEM allows only treated and non-treated farms having farmers less than 50 years old to be included in the first bin and only those with farmers more than 50 years old in the second.

---

[22] Note that the figures do not correspond to actual data used in the analysis, but have been created by the authors for the purpose of this exercise.

Table A.1: A numerical example of how CEM works in practice

A	B	C	E	F	G	H	I	L	M
Farm ID	Treatment	CEM Strata	Farm's Size (ha)	Farmer's Age	Revenue	Average Revenue Treated Farms	Average Revenue non-Treated Farms	Treatment Effect	ATT
1	1		8	45	40				
2	0	1	9	44	30	40	20	20	
3	0		7	45	20				
4	0		8	47	10				**ATT=(20+15+15)/3=16.6**
5	1		11	51	100				
6	1	2	12	52	80	90	75	15	
7	0		13	53	70				
8	0		12	54	80				

To estimate the causal effect of participating in the AEMs, the CEM first divides farms into the corresponding strata of the considered set of variables, which has been defined as detailed above. In our example, farms will be allocated in two strata. [23] Then, it computes the average outcome (column G) for treated and non-treated farms in each stratum separately (columns H and I). Successively, it computes the treatment effect for each stratum in all the considered variables, as the difference between the average outcome of treated and non-treated farms in each stratum (column L). Finally, the ATT of the entire sample (column M) is computed considering the average causal effect (computed in column L) for the treated units (in this case, for ID=1, the treatment effect is 20, while the mean treatment effect is 15 for both ID=5,6, that belongs to the same stratum). The result in column M suggests, on average, a positive effect on revenue of the farm's participation in the AEMs.[24]

This straightforward example highlights some interesting properties of the CEM.

- The level of imbalance is chosen *ex-ante* rather than being discovered *ex-post*. For instance, if we consider the farm size variable in Table A.1, we know that treated and non-treated farms can differ at most by 2 ha, this being the difference between the minimum value in the first strata, 7,

---

[23] For the sake of simplicity, in this example we assume that our hypothetical farms have characteristics such that they would fall only in two strata. This is because the proposed variable coarsening, would have created two additional strata: the first one including farms with more than 10 ha and leading farmer less than 50 years old, and the second one with farms with less than 10 ha and leading farmer more than 50 years old. In this example we assume that no farms in the sample have such characteristics.

[24] It is worth mentioning that CEM also produces weights that can be then used in weighted least squares regression (as we have done in the empirical analysis of this paper) to estimate the causal effect of the participation in a treatment. The procedure using weights leads to exactly the same results that we presented in this section, but is less straightforward to understand. As the objective of this section is to show how CEM works in a simple way, we prefer to focus on this easy-to-understand, but involved procedure.

and 9 the maximum value, while in the second stratum the variable ranges from 11 to 13. Considering *Farmer's Age*, the maximum level of imbalance is set at 3 years. Other matching estimators, like PSM, do not have this property. Therefore, the level of imbalance is discovered *ex-post* after the computation of the propensity score. It is not uncommon in this case to have a very good level of imbalance for some covariates, while for others there might be a wide variation between treated and non-treated units.

- The number of strata in which a variable can be divided, and, therefore, the level of imbalance, can be decided either by the user or automatically by the CEM algorithm looking at the variable's distribution. This is an important feature of CEM, as the user may know some important characteristics of a given variable that may suggest its stratification into meaningful bins. The definition of the strata is the basis of the treatment effect calculation – tighter bins (lower coarsening) will lead to the creation of a higher number of strata. A high number of strata will then lead to more similar observations within the same stratum, which will eventually result in a lower imbalance.[25] From this perspective, the CEM departs from other matching estimators, where the user does not have any control (*a priori*) on the level of imbalance. If we take for instance the example in Table A.1, we could even decide to further reduce the level of imbalance in the two considered variables. To do that, we should create one or even more additional bins in our variables, until the desired level of imbalance is reached. For instance, if we want to set the maximum level of imbalance in the *Farm's Size* variable to 1, we could divide that variable in the following three bins (instead of two): 7-8 ha; 9-10 ha; 11-12 ha.

- Related to the previous point, the adjustment of the imbalance of one variable in CEM does not affect the imbalance of the other variables. In other matching estimator methods, the procedures for the adjustment of the level of imbalance in one variable generates unpredictable effects on the level of the imbalance of the others. While in CEM the user can intentionally reduce or increase

---

[25] Note that, by progressively increasing the number of strata, a higher number of observations will be discarded. This is because when reducing the boundaries of the strata, the number of units sharing the same characteristics of the treated units decreases consequently.

the imbalance in one variable without affecting the imbalance in the others, in other matching estimators it is not possible to predict *ex-ante* what will be the effect of a procedure to reduce the imbalance in one variable on the other. In the example made in the previous point, the reduction of the imbalance in the *Farm's Size* variable (and therefore the creation of an additional stratum), does not affect the level of imbalance in the *Farmer's Age* variable.

Additional materials

Table A.2: 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.3: 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.4: 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.

Table A.5: Control variables strata thresholds

Farm characteristic	Variable description	Strata thresholds
Farm size	Utilized Agricultural Area (ha)	1 - 2 - 5 - 10 - 20 - 30 - 50 - 100
Farm location	Mountain	0-1
	Hill	0-1
	Plain	0-1
Type of farming - agriculture specialization	Cereals	0-1
	Other Crops	0-1
	Horticulture and flowers	0-1
	Vines	0-1
	Fruit	0-1
	Dairy cattle	0-1
	Beef cattle	0-1
	Dairy-beef cattle	0-1
	other herbivorous	0-1
	granivorous	0-1
	mixed cropping	0-1
	mixed livestock	0-1
	mainly grazing livestock	0-1
	Mixed livestock	0-1
	mainly granivores	0-1
	mainly granivores	0-1
Share of the most important crops on the UAA	Share arable crops	0 - 25 - 50 - 75
	Share permanent crops	0 - 25 - 50 - 75
	Share vegetables	0 - 25 - 50 - 75
	Share grassland	0 - 25 - 50 - 75
livestock density	Livestock units per hectar	0-1-2- 3- 4
farmer's age	Farmers' age	43-54-64
	Young (age<=40)	0-1
	Old (age>=65)	0-1
Previous participation in AEMs during the 2000-2006 programming period	Participation in low-input farming AEM	0-1
	Participation in organic farming AEM	0-1
	Participation in grassland maintenance AEM	0-1
	Participation in landscape improvement AEM	0-1

Table A.6: Estimated additional effects for crops diversification measure – Propensity Score Matching

Crops diversification measure	PSM with full set of farms			PSM exluding farms with 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)	1.078	3.262	**-2.183*** **(1.13)**	0.855	1.304	**-0.449** **(1.490)**
Main arable crop(%)	-0.112	11.623	**-11.734*** **(1.269)**	-5.242	7.554	**-12.795*** **(1.931)**
Number of arable crops	-0.149	-0.847	**0.698*** **(.087)**	0.225	-0.629	**0.854*** **(0.139)**
Arable crops heterogeneity	-0.220	-13.681	**13.460*** **(1.422)**	4.878	-8.990	**13.867*** **(2.133)**
Leguminous crops(%)	-1.077	-3.033	**5.378*** **(1.940)**	6.834	-0.552	**7.38*** **(2.760)**
Non soil-depleting crops(%)	-0.860	-6.239	**-2.519** (1.711)	-6.152	1.418	**-7.569**** **(2.718)**
Grassland(ha)	0.175	-0.241	0.415*** (0.155)	0.477	-0.130	0.607** (0.309)
Grassland(%)	0.930	0.029	0.900* (0.51)	1.350	-0.031	1.3813** (0.712)
Organic farming(ha)	-0.336	0.370	-0.705** -0.322	-0.043	0.036	-0.079 (0.078)
Organic farming - under conversion(ha)	-0.051	0.017	-0.068 -0.085	0.060	0.000	0.060* (0.0343)

*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 PSM implemented in STATA13. Standard errors are in parentheses. ***, **, or * denote statistical significance at 1%, 5% and 10% level, respectively.

Table A.7: Estimated additional effects for grassland maintenance measure – Propensity Score Matching

Grassland maintenance measure	PSM with full set of farms			PSM exluding farms with 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)	0.464	-2.285	**2.748*** (0.347)**	3.998	0.063	**4.292*** (0.562)**
Grassland(%)	1.771	-4.922	**6.692*** (1.08)**	19.051	1.645	**17.406*** (3.1315)**
Main arable crop(ha)	1.734	0.981	0.752 (0.847)	1.500	1.625	-0.125 (1.544)
Main arable crop(%)	-1.088	-1.145	0.057 (1.648)	-7.296	0.857	-8.152** (3.849)
Number of arable crops	-0.128	-0.105	-0.023 (0.067)	0.071	-0.315	0.385** (0.1678)
Arable crops heterogeneity	-1.030	0.204	-1.233 (1.172)	0.291	-3.584	3.875 (2.948)
Leguminous crops(%)	1.648	3.415	-1.767 (1.1439)	1.503	4.340	-2.837 (2.834)
Non soil-depleting crops(%)	-0.589	-0.334	-0.255 (1.862)	-4.903	0.199	-5.102 (4.359)
Organic farming(ha)	-0.512	0.117	-0.628*** (0.299)	-1.302448	0	-1.30245 1.302)
Organic farming - under conversion(ha)	0.038	0.085	-0.048 (0.120)	-	-	- -

*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 PSM implemented in STATA13. Standard errors are in parentheses. ***, **, or * denote statistical significance at 1%, 5% and 10% level, respectively.

Table A.8: Estimated additional effects for organic farming measure – Propensity Score Matching

Organic farming measure	PSM with full set of farms			PSM exluding farms with 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)	1.978	-8.757	**10.734*** **(2.574)**	3.592	0.038	**3.554*** **(1.257)**
Organic farming - under conversion(ha)	1.714	-1.787	**3.501*** **(0.931)**	2.315	0.000	**2.314*** **(1.301)**
Main arable crop(ha)	2.943	1.170	2.518* (1.495)	1.912	2.205	-0.293 (2.807)
Main arable crop(%)	7.142	-1.653	8.794* (4.734)	8.049	0.752	7.298 (8.742)
Number of arable crops	-0.059	0.250	-0.309 (0.313)	0.319	0.128	0.191 (0.391)
Arable crops heterogeneity	-4.780	-3.077	-1.703 (3.368)	-0.868	1.330	-2.198 (5.563)
Leguminous crops(%)	-0.836	-5.105	4.269 (3.82)	2.753	1.256	1.497 (5.979)
Non soil-depleting crops(%)	3.455	-1.853	5.308 (5.179)	9.615	4.323	5.292 (9.012)
Grassland(ha)	0.319	-0.144	0.463 (0.413)	0.724	-0.175	0.899 (0.806)
Grassland(%)	0.816	1.508	-0.692 (1.747)	-1.253	-1.936	0.683 (2.842)

*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 PSM implemented in STATA13. Standard errors are in parentheses. ***, **, or * denote statistical significance at 1%, 5% and 10% level, respectively.

Table A.9: Estimated additional effects for other AEMs – Propensity Score Matching

Other AEMs	PSM with full set of farms			PSM exluding farms with 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)	0.857	0.280	0.577 (0.961)	0.529	0.047	0.482 (1.413)
Main arable crop(%)	4.817	7.827	-3.010 (1.832)	2.488	5.946	-3.458 (2.885)
Number of arable crops	-0.103	-0.438	0.335*** (0.070)	-0.115	-0.412	0.296*** (0.109)
Arable crops heterogeneity	-2.430	-7.541	5.111*** (1.097)	-2.658	-7.226	4.567** (1.796)
Leguminous crops(%)	-1.758	-0.569	-1.189 (1.355)	0.890	5.102	-4.211** (1.922)
Non soil-depleting crops(%)	3.305	3.935	-0.630 (1.952)	0.668	4.649	-3.981 (3.106)
Grassland(ha)	-0.017	-0.390	0.373 (0.238)	0.149	0.001	0.148 (0.196)
Grassland(%)	-0.171	0.710	-0.881 (0.627)	0.066	0.901	-0.836 (0.936)
Organic farming(ha)	0.131	0.304	-0.173 (0.316)	0.503	0.424	0.079 (0.459)
Organic farming - under conversion(ha)	0.122	-0.099	0.221** (0.110)	0.102	0.007	0.096 (0.072)

*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 PSM implemented in STATA13. Standard errors are in parentheses. ***, **, or * denote statistical significance at 1%, 5% and 10% level, respectively.

Table A.10: Estimated additional effects considering outcome variables for the year 2005 in the set of control variables

	CEM with full set of farms				CEM exluding farms with previous participation			
	Crop diversification	*Grassland maintenance*	*Organic farming*	*Other AEMs*	*Crop diversification*	*Grassland maintenance*	*Organic farming*	*Other AEMs*
	ATT	ATT	ATT	ATT	ATT	ATT	ATT	ATT
Main arable crop(ha)	**-0.535**	0.0447	1.944*	-0.926**	**0.147**	-0.0320	2.183	-2.060***
	(0.472)	(0.315)	(1.177)	(0.371)	**(1.035)**	(1.274)	(2.016)	(0.625)
Main arable crop(%)	**-13.75*****	0.989	-6.654	-0.660	**-14.53*****	-8.006***	-3.501	-0.872
	(1.196)	(1.136)	(5.252)	(1.136)	**(1.539)**	(2.528)	(6.262)	(1.317)
Number of arable crops	**0.906*****	0.0281	0.236	0.120***	**1.036*****	0.376***	0.330	0.0311
	(0.0663)	(0.0376)	(0.212)	(0.0283)	**(0.100)**	(0.128)	(0.278)	(0.0365)
Arable crops heterogeneity	**14.98*****	-1.306	-0.170	0.602	**15.84*****	4.510*	-0.884	-1.141*
	(1.267)	(0.812)	(2.884)	(0.471)	**(1.690)**	(2.307)	(4.346)	(0.643)
Leguminous crops(%)	**9.376*****	-2.030***	2.111	-0.672	**12.91*****	-1.865	4.703	-0.308
	(1.415)	(0.509)	(3.732)	(0.482)	**(1.859)**	(1.811)	(4.081)	(0.639)
Non soil-depleting crops(%)	**-4.829*****	-0.516	-0.364	-1.169	**-6.419*****	-5.159	0.617	-3.405**
	(1.162)	(1.035)	(5.141)	(1.151)	**(1.882)**	(3.540)	(6.407)	(1.352)
Grassland(ha)	0.389***	**2.484*****	0.747	-0.00395	0.269*	**5.528*****	1.416	0.0241
	(0.0770)	**(0.177)**	(0.541)	(0.0595)	(0.138)	**(0.334)**	(1.276)	(0.0835)
Grassland(%)	0.809***	**7.582*****	-0.748	-0.603***	0.618	**23.92*****	-0.776	-0.384
	(0.190)	**(0.740)**	(1.338)	(0.206)	(0.413)	**(1.820)**	(2.336)	(0.254)
Organic farming(ha)	-0.392**	-0.0348	**5.313*****	0.572***	-0.0831	-0.0810	**4.401*****	0.844***
	(0.175)	(0.0302)	**(0.771)**	(0.0839)	(0.158)	(0.147)	**(0.188)**	(0.122)
Organic farming - under conversion(ha)	-0.0920	-0.0488	**4.983*****	0.123***	0.0558*	-0.0165	**2.819*****	0.0843**
	(0.0672)	(0.0306)	**(0.341)**	(0.0281)	(0.0301)	(0.0632)	**(0.189)**	(0.0375)

*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.