UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI SCIENZE ECONOMICHE E SOCIALI

DO RURAL DEVELOPMENT POLICY MEASURES REALLY AFFECT FARMERS' BEHAVIOUR AND PERFORMANCE? A SYNTHETIC DIFFERENCE-IN-DIFFERENCES ESTIMATION.

ROBERTO ESPOSTI

QUADERNO DI RICERCA n. 494
ISSN: 2279-9575

January 2025

The views expressed in the articles are those of the authors and do not involve the responsibility of the Department.

Abstract

This paper deals with the response of farmers to targeted policy measures. The research question consists of whether and to what extent this response, upon voluntary adoption and eventually driven by private motivations, also generates the outcome of societal interest. Challenges posed by these policies are multiple since they may admit a staggered adoption with often very few, at least initially, treated units. In order to deal with these challenges, the paper adopts a Synthetic Difference-in-Differences approach. Within this causal inference logic, an appropriate theoretical framework is elaborated to model farmers' behavioural response to the policy distinguishing between the private and societal outcomes. This approach is applied to a balanced sample of Italian farms and to some selected measures of the second pillar of the EU Common Agricultural Policy over the period 2014-2022. Results point to the identification and estimation issues emerging when the entry into the treatment is staggered and treated units are few and heterogeneous. For some policy measures the estimated treatment effect is significant for the private outcome while it seems weaker and more volatile for the societal outcome.

KEYWORDS: Rural Development Policy, Common Agricultural Policy, Farmers' Decision-Making, Staggered Treatments, Synthetic Difference-in-Differences.

JEL classification: C21, Q12, Q18

Address:

Roberto ESPOSTI
Department of Economics and Social Sciences - Università Politecnica delle Marche, 60121, Ancona (Italy). Email: r.esposti@staff.univpm.it

Do rural development policy measures really affect farmers' behaviour and performance?
A Synthetic Difference-in-Differences estimation.

Roberto Esposti
Department of Economics and Social Sciences, Università Politecnica delle Marche, Ancona, Italy;
r.esposti@staff.univpm.it

Abstract

This paper deals with the response of farmers to targeted policy measures. The research question consists of whether and to what extent this response, upon voluntary adoption and eventually driven by private motivations, also generates the outcome of societal interest. Challenges posed by these policies are multiple since they may admit a staggered adoption with often very few, at least initially, treated units. In order to deal with these challenges, the paper adopts a Synthetic Difference-in-Differences approach. Within this causal inference logic, an appropriate theoretical framework is elaborated to model farmers' behavioural response to the policy distinguishing between the private and societal outcomes. This approach is applied to a balanced sample of Italian farms and to some selected measures of the second pillar of the EU Common Agricultural Policy over the period 2014-2022. Results point to the identification and estimation issues emerging when the entry into the treatment is staggered and treated units are few and heterogeneous. For some policy measures the estimated treatment effect is significant for the private outcome while it seems weaker and more volatile for the societal outcome.

Keywords: Rural Development Policy, Common Agricultural Policy, Farmers' Decision-Making, Staggered Treatments, Synthetic Difference-in-Differences.

JEL Classification: C21, Q12, Q18.

1. Introduction

This paper concerns the application of the Causal Inference (CI) logic to policy assessment. This logic seems to properly fit the fundamental questions underlying policy evaluation, just by interpreting the impact of the policy as a treatment effect (TE) and by establishing the necessary conditions to consider policy implementation as a natural experiment or quasi-experiment (Cerulli, 2015). In this respect, the recent literature has also revealed how this logic can be demanding and, often, unfeasible in many real-world cases especially when policies assume some specific characteristics. In particular, we intend those policies whose implementation eventually prevents any control on the set of treated units (i.e., those subject to the policy measure) and, consequently, on the set of control units (the non-treated units, i.e., not subject to the policy measure).

As an exemplary case, here we concentrate on a selection of measures of the so-called second pillar of the EU Common Agricultural Policy (CAP), also known as the Rural Development Policy (RDP), over the period 2014-2020 (then extended to 2022) (Esposti, 2022). Of these policies, we aim to assess the response in terms of both the farm-level private and societal performance. A major objective of this analysis thus consists in evaluating CAP support effectiveness across the different measures thus pointing to room for better policy targeting.

Performing such an evaluation exercise within a TE logic is empirically challenging. What is needed is an appropriate research design that conceptually implies recognizing the process from the policy measure to the result: the measure is offered to farmers; farmers accept to participate; farmers adjust their behaviour consequently; the result is produced. However, the logic underlying the farmer's decision-making concerns the private benefit and not the societal benefit associated with the policy measure. So, two different outcomes have to be considered: the private outcome that drives the farmer's decision-making; the societal outcome for which, in principle, the policy itself has been designed and implemented. In order to be consistent with this logic, we first need a proper reclassification of the different policy measures offered to the farmers on the basis of the societal outcome they are expected to bring about. Secondly, a suitable methodological approach must be adopted to properly identify and estimate the TE.

This approach has to take into account the main feature of this kind of policy. Since adoption is voluntary, entry into the treatment is staggered over time, often with very few early adopters. Moreover, adopters, and early ones in particular, often show some peculiar characteristics especially in the farming sector, typically presenting large heterogeneity across units. Challenges posed by these circumstances are widely acknowledged within the CI literature and investigated in many empirical studies. An approach known as Synthetic Difference-in-Differences (SDID) has been recently proposed to deal with these issues altogether (Arkhangelsky et al., 2021). The SDID approach is adopted in the present study not only for its suitability but also because its application to different real-world cases, like the CAP measures here considered, may provide interesting comparative evidence about its potentials and limits. The empirical investigation is performed on the 2014-2022 balanced panel of Italian Farm Accountancy Data Network (FADN) farms. Italian agriculture is often considered an interesting case study for the large heterogeneity of farming conditions it presents which inevitably affect farmers' policy adoption and their consequent response (Coderoni and Esposti, 2018; Baldoni et al., 2021).

The rest of the paper is structured as follows. Section 2 overviews the main challenges and the literature about the assessment of the impact of farm-level policies and, in particular, of the EU RDP measures. Section 3 illustrates the theoretical framework underlying the farmers' behaviour upon the adoption of these policy measures and its interpretation in a TE logic. Section 4 presents the adopted sample and dataset and describes the research design. Specific attention is paid to the reclassification of the EU RDP measures and the consequent definition of the treatment and outcome variables. Section 5 details the SDID estimation approach, its limits and potentials, and the set of results it is able to return. Results are then reported and discussed in Section 6 while Section 7 concludes by deriving some methodological and policy implications.

2. Challenges in EU RDP evaluation

2.1. The main open issues

Since its origin (1962), the CAP has supported the agricultural sector and farmers' income. Over time, however, this policy has undergone major reforms so most of the original market support has been gradually transformed into direct income support while measures designed to promote agricultural structures have been progressively introduced, redesigned and reinforced. This latter segment of the CAP is also known as the EU Rural Development Policy (RDP) and designated as Pillar 2 of the CAP since the Agenda 2000 reform (Sotte and Brunori, 2025). In the last two decades it has become the backbone of the strategic intervention of the EU in rural areas, accompanying and promoting farmers' adaptation to new societal needs and objectives, usually summarised by competitiveness, diversification and public (mostly environmental) goods provision (Swales, 2007; Esposti et al., 2016; Esposti, 2022).

Despite this evolution and the many reforms, the RDP within the CAP continues to face criticisms from various quarters. In particular, they highlight the lack of a sound evaluation of efficiency and effectiveness of the diverse measures and stress the importance of prioritising objectives to ensure a more focused and effective policy-making process. Criticisms also insist on the need for a greater emphasis on empirical, interdisciplinary evidence which seems necessary to evaluate the impacts of policy interventions more accurately (Esposti and Sotte, 2013; Mennig, 2024). One major source of these issues can be detected in the fundamental nature of some RDP measures that remain substantially addressed to support farmers' income without any (or very limited) behavioural implication (Esposti, 2022). It is the case, for instance, of the compensatory payments for farming in Less Favoured Areas (LFA). For the sake of simplicity, we call them *conservative measures* (or *passive measures* according to the terminology adopted by Esposti, 2022). These policy measures make it inherently very problematic to assess their effectiveness and efficiency since there is no response expected upon the adoption of the policy on which the evaluation can be grounded.

Nonetheless, many RDP measures can be regarded as non-conservative: their adoption requires a behavioural response, that is, farmers have to change their production choices in order to be entitled to receive the support implied by the measure. For instance, a compensatory payment to farmers covering additional costs and foregone income upon the adoption of more environmentally friendly practices (Coderoni et al., 2024). As a consequence of this behavioural response, a change in farms' performance can also be observed and this can be regarded as the outcome (the TE, in the CI terminology) of the policy measure (the treatment). They can thus be designated as *non-conservative measures* (or *active measures*) (Esposti, 2022). In principle, their non-conservative nature should make these RDP measures more prone to evaluation within a CI logic, namely in TE terms. In practice, this assessment remains highly challenging and often unfeasible.

The main reason is that CI requires some specific characteristics of the policy under evaluation that we can summarize here with the term granularity. With granular policy we intend a policy that combines two main characteristics. Firstly, it is designed and implemented to target the specific characteristics of a beneficiary. Secondly, and consequently, a granular policy is expected to induce a specific response with a well-defined observable, outcome. The problem with the non-conservative RDP measures is that they are not sufficiently granular. On the one hand, they are intended to support a specific response often associated with farm characteristics. On the other hand, they are implemented more like a general (or universalistic) policy. Requisites to enter Pillar 2 measures or treatments are often very limited (if any). It follows that most farmers are potential beneficiaries and they can voluntarily enter the policy at any point in time (Esposti, 2022). The fact that for some RDP measures the number of adopters is very small, especially in the first years (see Section 4), does not depend on some strong selectivity of the policy. It simply depends on the lack of convenience for the farmers or on the technical and bureaucratic complexity of the adoption (or, more often, on a combination of the two factors).

There are two main consequences of this lack of granularity, from the perspective of the CI logic. First of all, a very limited control of the treated units (thus also of the control units) by the policy maker (and the analyst). Secondly, a potential ambiguity about the outcome variables on which policy evaluation has to be grounded. The present paper focuses on these non-conservative but "ungranular" measures. The fundamental research question concerns whether we can reliably assess their impact or effect, namely, whether we have a methodological strategy to deal with these cases. There are two main methodological challenges posed by these policies: numerosity and heterogeneity of adopters; voluntariness and staggered timing of adoption.

First of all, adopters are few and, often, very heterogeneous. The heterogeneity of these units is intrinsically irreducible since it depends on some hardly observable characteristics of the units themselves, like capacity and ability, motivation and beliefs (Esposti, 2024). In CI terms, this makes it very difficult to control this heterogeneity. Secondly, the voluntary participation in the treatment raises a severe theoretical challenge in distinguishing between (and controlling for) those units' characteristics that affect the policy adoption from those that affect the policy effect. However, participation does not only imply that units may decide whether or not to adopt the policy measure, but also when to enter the treatment itself. This makes the treatment entry time-variant, a circumstance usually referred to as *staggered treatment* in recent CI literature (Sun and Abraham, 2021). This latter characteristic may be reinforced by bureaucratic and budgetary constraints in policy implementation that delay (or even prevent) the entrance of some units that would voluntarily adopt (Baldoni and Esposti, 2023).

Ultimately, the number of treated units could be, at least initially, very small and then evolve over time due to the progressive entry of new units into treatment. Consequently, also the number of untreated units (thus of controls) is itself time-variant. This time variability implies that the Average Treatment Effect (ATE) is not only an average across units but also across time (i.e., a cohort-specific ATE, different across entry times). This may make the conventional Difference-In-Differences (DID) logic inappropriate because time variability may jeopardize the validity of the Parallel Trend Assumption (PTA) over different cohorts of units (Sun and Abraham, 2021). In addition, the poor number of treated units can also jeopardize matching with controls. One possible way to deal with the sparsity of treated units consists in the Synthetic Control Method (SCM) (Abadie et al., 2010; Abadie et al., 2015; Abadie, 2021) that has been proposed as an alternative to conventional matching methods. However, SCM shows limitations in considering the staggered entry into the treatment and the consequent cohort-specific TE. This makes suitable a sort of mixture of these approaches also known as Synthetic Difference-in-Differences (SDID) (Arkhangelsky et al., 2021; Clarke et al., 2023).

This recent approach is adopted here and applied to some selected CAP RDP measures over the period 2014-2022. In particular, we want to investigate whether the SDID approach can be useful to evaluate these policies. Differences across policy measures can also provide interesting comparative evidence on the suitability, potential and limits of this approach in such real-world cases. This methodological strategy entails the definition of an appropriate research (i.e., quasi-experimental) design that has both a theoretical and an empirical justification. From both the theoretical and empirical perspective, the search for an appropriate quasi-experimental setting encounters another major issue, namely the ambiguity about the outcome variable to be considered. For several RDP measures a target variable is simply neither explicit nor univocal, i.e. there can be several legitimate outcome variables depending from which perspective the policy has to be evaluated. In any case, whatever the societal outcome underlying the policy measure of interest, its voluntary adoption, as well as the farmers' response, is driven by the farmers' own interest and this private outcome may differ (and usually does) from the societal outcome (Coderoni et al., 2025).

Besides adopting the SDID identification and estimation approach, the present paper aims to provide another original contribution in properly dealing with the possible ambiguity in the definition of the expected policy outcome. Sections 3 and 4 detail the quasi-experimental design to deal with this issue.

2.2. The RDP measures under analysis

To pursue the research objective discussed above, some 2014-2022 RDP measures are here considered as exemplary. They share the same logic: farmers voluntarily accept to participate in measures that mimic a sort of contract between the farmers themselves and the whole society. In practice, farmers receive a monetary support if and only if they behave in a way to produce predetermined and measurable performance of societal interest.

The 2014-2022 EU RDP (the Pillar 2 of the CAP) is structured in 6 priorities and 18 Focus Areas. They are pursued through the implementation of 20 measures with respective 66 sub-measures. One of these measures (Measure 18) is actually specific to accompanying the accession of Croatia (that occurred in 2014). Therefore, measures of actual interest here are 19 with 65 sub-measures. Annex 1 (Table A1) details this policy structure. For the scope of the present analysis, however, it seems helpful to reorganize these policy measures differently, i.e. in terms of the underlying expected societal outcome. Some attempts in this respect have already been made in previous studies mostly focusing on the fact that several measures actually share the same fundamental purposes, as expressed by the underlying priorities themselves (Sotte, 2009; Camaioni and Sotte, 2010; Sotte, 2012). Here, this reclassification is mostly grounded on a different perspective, that is on the basis of the expected behavioural response of farmers. First of all, we distinguish between behavioural (i.e., non-conservative or active) and non-behavioural (conservative or passive) policies. This paper focuses only on the former type of measures. They can then be structured on the basis of this expected response. As clarified more extensively in Section 4, we reclassify the behavioural measures into 6 groups (from P1 to P6) depending on the expected societal outcome.

Table 1 summarizes this classification. It also shows that three types of measures emerge in terms of numerosity of treatment adoption and entry timing: very few units and a little staggered entry (measures 1.1, 2.1, 2.2; 6.4; 16.1, 16.9); few units but largely staggered entry (measures 3.1-3.2; 8.1-6); many units and largely staggered treatment (measures 4.1-5.2; 14). In any case, though staggered, all these policy measures are here considered as absorbing: once a farm enters the treatment, it remains treated for the whole following observed period (see also Section 4). It is worth noticing that a staggered treatment implies a cohort-specific TE but it does not necessarily imply a dynamic or time-variant TE for any given cohort (Sun and Abraham, 2021). Here, we exclude this latter circumstance and only admit that the TE may vary across cohorts, that is, depending on the timing of treatment adoption.

[Table 1 here]

3. Theoretical framework: farmers' decision-making and policy indicators

Recent literature emphasizes the shift from an action-based to a result-based logic within the RDP design (Eichorn et al., 2024; Targetti et al., 2024). The former would show low effectiveness and efficiency because it disregards the large heterogeneity in farming conditions, so that the same farmers' behaviour generates a very different outcome. Focusing on the latter (the result) would improve both effectiveness and efficiency. However, this is particularly challenging because the result may be hard to observe and, more importantly, can be independent of policy measures and the farmers' actions themselves (Targetti et al., 2024). The key point here is that action and result-based policies pose the same conceptual and empirical challenges: the identification of the Treatment Effect (TE) and the appropriate research design to achieve it together with TE estimation.

What is often neglected is that assessing effectiveness and efficiency of these results-based measures is empirically demanding precisely because identification of the result as a consequence of the policy, that is, the TE is challenging. Designing a result-based contract is not enough. What is needed is an appropriate research design. The latter conceptually implies recognizing the process from the policy measure to the result: the measure is offered to a farmer; the farmer accepts to participate; the farmer adjusts his behaviour consequently; the result is produced. However, the logic underlying the farmer's decision-making concerns the private benefit and not the societal benefit associated with the policy measure. Voluntariness in policy adoption implies that the private and the societal outcome of the treatment can substantially differ and that, consequently, the adoption of the policy requires a full understanding of the underlying motivation.

This requires an appropriate conceptual and theoretical background. As policy measures are those intended to be non-conservative, that is, they are expected to induce a behavioural response by the farmers, this theoretical framework has to concentrate on farmers' decision-making and performance generating. In turn, this latter implies two different results: the private one, that drives the farmer's decision-making; the societal one, that is a sort of side-effect from the farmers' perspective but it is what really matters for the calculation of policy effectiveness and efficiency.

Consider a panel of *N* production units (farms) observed over *Z* time periods. For any i-th farm (*i*=1,…,*N*) and t-th time period (*t*=1,…,*Z*), an aggregated multi-input multi-output production technology can be represented by the feasible production set $F_{it} \subset \mathbb{R}^M$. F_{it} is farm specific and contains all sources of heterogeneity in farmers' decision-making in terms of both treatment participation and production choices (Esposti, 2024). F_{it} is shaped by all the specific features of the i-th farms, depending on both external and internal factors, that we generally indicate with the $(Q \times 1)$ vector $\boldsymbol{X}_{it}$. Given F_{it}, a $(M \times 1)$ vector of netputs $\boldsymbol{Y}_{it} = (Y_{it}, \dots, Y_{itM})'$ is feasible if $\boldsymbol{Y}_{it} \in F_{it}$. This netput vector, $\boldsymbol{Y}_{it}$, contains both farm outputs (with a positive sign) and inputs used (negative).

Each farm is offered a binary policy *T=0,1*. Non-conservative policy measures can be defined as those targeting specific farming activities to improve farms' performance concerning some societal objective. If the i-th farm adopts a given measure, receiving the respective support, it is $T_{it} = 1$; it is $T_{it} = 0$ otherwise. In a given period *t* (*t*=0,…,*Z*), if chosen by the i-th farm, the treatment (T_{it}=1) is expected to induce specific production choices ($\boldsymbol{Y}_{it}$). Therefore, for a given farm, a treatment can be univocally mapped into production choices ($T_{it} \leftrightarrow \boldsymbol{Y}_{it,k}$). It is then possible to express these choices as a function of the treatment itself, plus those abovementioned exogenous farm-specific characteristics $\boldsymbol{X}_{it}$ (or confounding variables) influencing farmers' behaviour: $\boldsymbol{Y}_{it} = f(T_{it}, \boldsymbol{X}_{it})$ where $f(\cdot)$ is a vector-valued function.

Within this theoretical framework and a binary TE logic, if both $\boldsymbol{Y}_{it}$ and T_{it} are observed $\forall i \in N$ and $\forall t \in T$, two different TEs can be identified. The first TE concerns the outcome of interest for the farmers, that is, their farm profit (or net income; see Section 4), π_{it}. Ultimately, what matters from the farmers' perspective is whether policy measure adoption and the associated policy support S_{it} (namely, the amount of monetary support received by the i-th farm at time *t* to adopt the respective policy measure), improves their profit. Let us generically express the farm-specific profit function as:

(1) $$\pi_{it} = p[f(T_{it}, \boldsymbol{X}_{it})]$$

where $p(\cdot)$ is a single-valued function. Consequently, the first TE associated with the generic policy will be π_{it}, that is:

(2) $$\pi_{it} = [(\pi_{it}|T_{it} = 1, \boldsymbol{X}_{it}) - (\pi_{it}|T_{it} = 0, \boldsymbol{X}_{it})] = p[f(T_{it}, \boldsymbol{X}_{it})]$$

where T_0represents the baseline case, that is, the non-treatment condition and $f(\cdot) = \boldsymbol{Y}_{it} = [(\boldsymbol{Y}_{it}|T_{it} = 1, \boldsymbol{X}_{it}) - (\boldsymbol{Y}_{it}|T_{it} = 0, \boldsymbol{X}_{it})]$. It is worth noticing that $\Delta\pi_{it}$ here does not include the policy support, S_{it}, thus it is also referred to as before-support farm net income.[1]

The second TE concerns the outcome of interest for the society (or the policy maker), that is an improved performance of the treated (i.e., supported) farms. This improved performance from a societal perspective can be either a negative (for instance, lower use of polluting inputs)[2] or a positive (for instance, some biodiversity indicator)[3] variation. Assume that for any i-th farm the production choice expressed by $\boldsymbol{Y}_{it}$ univocally determines its societal performance $\boldsymbol{G}_{it}$, where $\boldsymbol{G}_{it}$ is a $(P \times 1)$ vector of *P* indicators of societal interest. Therefore, production choices can be univocally mapped into these indicators ($\boldsymbol{Y}_{it} \leftrightarrow \boldsymbol{G}_{it}$). If the interest is in a single performance, G_{it} will be a scalar term. Therefore, it will be possible to express the farm-level societal performance as a function of what

[1] As clarified in section 4, this is the definition of the outcome variable adopted in the present empirical analysis and it is also the variable provided within the FADN dataset.
[2] This is the case of LSU/UAA for measures P5 (see Table 1).
[3] This is the case of FA/UAA for measures P6 (see Table 1).

determines the production choice $\boldsymbol{Y}_{it}$. It will be $G_{it} = g(T_{it}, \boldsymbol{X}_{it})$ where $g(\cdot)$ is a single-valued function. The second TE will thus be G_{it}:

(3) $$G_{it} = [(G_{it}|T_{it} = 1, \boldsymbol{X}_{it}) - (G_{it}|T_{it} = 0, \boldsymbol{X}_{it})] = g(T_{it}, \boldsymbol{X}_{it})$$

Given the definition of π_{it} as a before-support net income, we can argue that, in general terms, the farmers' voluntary choice of the treatment ($T_{it} = 1$) implies $(\pi_{it} + S_{it})$≥0. However, since we admit here that both π_{it} and G_{it} are heterogenous, namely farm-specific, we can not exclude that non-monetary motivations of farmers' policy adoption might also justify a non-positive profit outcome of the treatment $(\pi_{it} + S_{it})$ ≤0 (thus, π_{it} <0) (Esposti, 2024).

Consequently, the actual social cost of the policy measure is $(S_{it} - \Delta\pi_{it})$. On the basis of the two TEs identified above, it is thus possible to express the effectiveness of the policy under investigation, i.e., to compute an indicator expressing how much it costs to the society as a whole to obtain a unit-variation of the farm-level performance G_{it}. Here, we generally call this indicator the *Social Cost of Farm Performance* (SCFP) and it is calculated as:[4]

(4) $$SCFP_{it} = \frac{(S_{it} - \Delta\pi_{it})}{\Delta G_{it}}$$

As the outcome variables equation (4) could also be expressed in relative (i.e., size-independent) terms (see Section 4 for details), (4) becomes:

(5) $$SCFP_{it} = \frac{\left(\frac{S_{it}}{R_{it}} - \frac{\Delta_{it}}{R_{it}}\right)}{\frac{\Delta G_{it}}{R_{it}}}$$

where R_{it} indicates the farm size possibly expressed either as physical size (e.g., Utilised Agricultural Area or unit of family labour) or economic size (e.g., Gross Production Value or Value Added). If different indicators of the farm size are used to relativize the two outcome variables, say R_{it} and V_{it} respectively, (5) becomes:

(6) $$SCFP_{it} = \frac{\left(\frac{S_{it}}{R_{it}} - \frac{\Delta_{it}}{Z_{it}}\right)}{\frac{\Delta G_{it} V_{it}}{R_{it}\, Z_{it}}}$$

As detailed in Sections 4 and 5, (5) and (6) have the further advantage of reducing the variability (or dispersion) of the outcome variables due to the highly heterogeneous farm size. This may significantly improve the statistical robustness of the estimation of the TE. Expressing outcome variables in relative terms, however, does not exclude the possible impact of the farm size on the TE calculation due to scale economies. Consequently, variables expressing the size can still be part of the $\boldsymbol{X}_{it}$ set.

4. Data and research design

4.1. The observational dataset and the treatment sets

In the present study we use information from the Italian FADN dataset. Italian agriculture is often considered an interesting case study for the wide heterogeneity of its farming conditions and traditions which inevitably affect farmers' decision-making and farm performance ((Esposti, 2017a,b; Coderoni and Esposti, 2018; Coderoni et al., 2024). The Italian FADN sample consists of a representative collection of commercial farms observed over the CAP regime under consideration (2014-2020 then extended to 2022). It comes as an unbalanced panel consisting of 10573 observations in 2014, 9569 observations in 2015, 10153 in 2016, 10792 in 2017, 10769 farms in 2018, 10805 observations in

[4] Equation (4) would suggest that the treatment intensity may differ across farms and, therefore, that the treatment should be considered multi-valued rather than binary. Nonetheless, though farm-specific, S_{it} actually depends on the farm size. Therefore, if variables are expressed in relative terms as in the present case, considering the policy support as a binary treatment remains a reasonable approximation (Coderoni et al, 2024).

2019, 10764 observations in 2020, 11040 observations in 2021 and 11084 observations in 2022. Though the programming period is 2014-2022, in the year 2014 the only treated units concern payments from the previous programming period. Therefore, the actual period of policy implementation of interest (and the consequent sample) is 2015-2022.[5]

On this observational data set two further elaborations are needed. Both are required by the SDID estimation approach here adopted (see Section 5). First of all, it requires a balanced panel sample. Secondly, for any unit at least two years of observation before entry into the treatment are required. Therefore, for the remainder of the analysis two balanced panels are extracted. One covers the period 2013-2022 and concerns those policy measures for which some units enter the treatment already in 2015 (policies P1, P3, P5 and P6 of Table 1). The other balanced panel covers the period 2015-2022 and concerns those policy measures for which early adopters enter the treatment only in 2017 (policies P2 and P4). Details on the numerosity of the two balanced panels across the policy measures are provided in Table 2. Panel 2013-2022 contains 2605 units observed over 10 years, thus 26050 observations overall; for panel 2015-2022 the units are 3899 corresponding to 31192 observations overall. It could obviously be possible to align the evaluation of all policy measures on period 2015-2022 with the respective balanced sample. At the cost of losing more than one thousand observations,[6] going back to 2013 has the advantage to include a longer period under the treatment. As will be clarified in Sections 5 and 6, this brings about more robustness in the identification of the TE especially when TE itself shows wide variability across cohorts.

As anticipated in Section 2.2, of the 65 sub-measures of the 2014-2022 RDP only non-conservative ones are considered. The logic here followed to investigate non-conservative sub-measures consists of reclassifying and regrouping them according to the expected farmer's response and consequent societal outcome, that is the performance of interest for the policy maker and which eventually justifies the monetary support. Similarly to previous reclassification exercises (Sotte, 2009; Camaioni and Sotte, 2010; Sotte, 2012), here sub-measures are grouped into six general policies: P1: Information, Education, Training; P2: Food chain and product quality; P3: Structural agricultural investments; P4: Structural non-agricultural investments; P5: Animal production (or livestock) management; P6: Forest management. Table 1 details the allocation of the RDP non-conservative sub-measures across these six general policies or treatment groups (see Annex 1 for the codification and extensive definition of the sub-measures). Table 2 presents the sample numerosity by treatment group and year while Figure 1 displays the progressive entries into the treatment during the period of observation.[7]

[Table 2 here]

[Figure 1 here]

4.2. Definition of the outcome variables

The proper definition of the outcome and confounding (**X**) variables is driven by the theoretical framework illustrated in section 3. It thus follows that two distinct outcome variables have to be simultaneously considered in any policy measure assessment: the private and the societal outcome.

[5] At the time of writing, validated FADN data for 2023had yet to be released and, in any case, this year would refer, at least partially, to the new CAP regime (2023-2027) that started on 1st January 2023.

[6] Part of the loss of observations in the balanced panel going back from 2015 to 2013 depends on the fact that several farms, which would remain in the balanced sample, still report the adoption of analogous or similar measures for years 2014 and 2013, though actually referring to the previous CAP programming period (2007-2013). Since at least two before-treatment observations are needed, and in order to avoid overlap between two different CAP regimes, these units have been removed from the balanced panel.

[7] Cases of multiple treatments (i.e., units receiving simultaneously two or more of the six treatments here considered) can not be excluded, in principle. However, as clearly emerging from Table 2, the numerosity of adopters is usually quite low and this possibility can be ruled out in the two balanced panels here considered. In any case, the presence of multiple treatments would not necessarily undermine the analysis if the assumption of independence across treatments is maintained. Considering the different social outcome associated with the different treatments, this independence assumption seems to implicitly hold true here.

Table 1 also associates the private and societal outcome variables with the six treatment groups. The private outcome is always the Net Farm Income (NI), that is, variable π of Section 3, to be intended as net of RDP measure support. One practical problem in using net income as an outcome variable is that it is highly size-dependent. This dependency can hardly be controlled by including size variables in **X**, as it cannot be excluded that a size bias remains and, in any case, this would not eliminate the very large variability of this private outcome. Therefore, to make π size-independent, NI is here divided by the annual units of farm family work (FAWU). After all, per capita farm income is what the farmers really care about, and what drives their choices. Therefore, it seems the most appropriate definition of the private outcome variable within the present study.

The societal outcome variable (that is, variable G of Section 3) must be selected according to what is explicitly established by the policy itself. On this aspect an imponent effort was continuatively and incrementally made by EU institutions (European Commission 2012; 2024; 2015a,b; 2016, 2018). These documents clarify that RDP measures are expected to chase very general purposes defined by the three fundamental objectives of the "Europe 2020" strategy (smart, sustainable, inclusive growth) and the six key priorities of the 2014-2022 RDP. Some measures can be univocally associated with a single priority while others refer to several of them, as in the case of the so-called horizontal measures (Measure 20, for instance). In turn, any priority entails several (from two to five) focus areas for a total number of 19 focus areas.

The evaluation approach proposed by the European Commission for the 2014-2020 Rural Development Programs includes a comprehensive set of Common Monitoring and Evaluation Framework (CMEF) indicators. The CMEF outlines the comprehensive approach for monitoring and evaluating the performance of the RDP. The CMEF provides key information on the implementation of the RDP through the use of indicators and sub-indicators (European Commission, 2014; 2015b; 2019). These indicators are categorized into three main types. Output Indicators (OI) measure the direct products of the program's activities.[8] Result Indicators (RI) assess the immediate effects of the program's activities on the target groups.[9] Impact Indicators (II) evaluate the broader, longer-term impacts of the RDP on the economic, social, and environmental dimensions of rural areas.[10] The European Commission also included a further set, that of Target Indicators (TI), for the evaluation of the 2014-2020 Rural Development Program. Target indicators are specific, measurable goals set at the beginning of the program to define expected achievements by the end of the programming period.[11] For more details, see also Table A2 and Figure A1 in Annex 2.

These sets of indicators are obviously informative for the present investigation in order to properly identify the expected societal outcome of the different RDP measures. Eventually, this plethora of indicators put forward by the EU Commission for the RDP evaluation does not allow an univocal identification of the societal outcome variable to be considered in investigating the farm-level response to policy adoption. Not only this outcome necessarily differs across measures. More importantly, for several policies, more than one single societal outcome can be legitimately associated with the measure under consideration. In principle, RI and TI are closer to this idea as both measure direct, short-term outcomes and concentrate on the involvement in terms of land and farms. However, most RI and TI are not expressed at the farm level thus they do not capture farmers' choices and behaviour that eventually link to the ultimate societal objective.

What we need here is some mixture between alle the abovementioned sets of indicators, though at the different levels. On this basis and on the basis on our own consideration, the set of societal

---

[8] Examples include: number of operations supported; total public expenditure; number of beneficiaries.

[9] Examples include: jobs created in supported projects; percentage of agricultural land under management contracts to improve biodiversity; percentage of the rural population covered by local development strategies.

[10] Examples include: increase in agricultural productivity; reduction in greenhouse gas emissions; Improvement in water quality; reduction in the rate of population decline in rural areas.

[11] Additionally, the CMEF framework also includes Context Indicators (CI): These provide information about the general conditions and trends in the rural areas and are used for contextual analysis.

outcome variables here considered is detailed in Table 1 that also reports, in the last column, the closer reference indicators of the CMEF terminology associated to each of the selected societal outcome variables. Some of these latter are, in fact, mid-way between impact and result indicators. They are measure-specific and may be multiple leading to respective alternative estimates of the societal TE. In order to facilitate the comparison across largely heterogenous farm structure, this impact or result indicator is preferably referred to variables expressing the farm size like labour units (expressed by Annual Working Units, AWU), physical size (Utilised Agricultural Area, UAA) or economic size (Gross Production Value, GPV; Gross Value Added, GVA; Standard Output, SO).

4.3. Covariate set

The confounding variables are selected to capture the key features of the unknown farm-specific technology (Coderoni et al., 2024; Esposti, 2024). Following Brown et al. (2021) and Stetter et al. (2022), these farm attributes can be assembled into four groups: economic factors (i.e., factor endowment); socio-demographic characteristics (of the farm's holder and workforce); environmental factors (mostly geographical but also including other forms of policy support); and idiosyncratic characteristics (such as ability, knowledge, motivations, beliefs, and values of the farm's holder and workforce, as well as unobserved environmental features such as agronomic characteristics and fertility).[12]

Among economic factors, in order to express the possible presence of scale economies, a size variable is included (SIZE). Size here enters as a dimensional class, that is, an ordered categorical variable[13] and it is preferred to a continuous variable expressing the economic size (for instance, farm's GPV or SO) as it is much more stable over time. To take the multioutput nature of technology into account, the farm's production specialization (PRO) is included. Also in this case a (unordered) categorical variable is preferred as the share of specific products on the farm's GPV or SO would be highly unstable.[14] Moreover, as this PRO variable does not consider forestry production, an additional variable (FOR) expressed as the ratio of the farm's forest area on the utilised agricultural area (UAA) is included.[15] To represent the multiple production factors, a four-input technology (i.e., capital (K), labour (L), energy (E), materials (M)) is considered. Capital is expressed by the total farm's machinery power (KW), by the UAA and by the livestock units (LSU). Labour is expressed by annual working units (AWU). Energy is expressed by the total expenditure for energy (fuels included) (ENE), while materials are expressed by the total expenditure for fertilizers, pesticides and animal feed (MAT).

To complete the definition of the farm's factor endowment, two further corrections are made. First of all, input quality matters and this is particularly true for labour. To take this aspect into account, the farm holder's age is included (AGE) together with the dummy expressing the farmer's gender (GEN, where female=1). In addition, to be consistent with the size-independent outcome variable, all the non-categorical input variables are expressed as intensities, that is, divided by the farm's SO. For instance, AWU is computed as the annual working units per unit of SO.

A further set of confounding variables is intended to take the farm's geographical context into account. In particular, the farm's average altitude (ALT), latitude (LAT) and longitude (LON) are included among covariates. We include among confounders also the overall CAP Pillar 1 payments received by the farms divided by the farm's SO (PIL1). This variable reflects a combination of several farm-specific factors. In the Italian case, the 2014-2022 CAP regime implies that PIL1 partially

[12] See also Eichorn et al. (2024, Figure 1) for a slightly different articulation of the covariates affecting farmers' policy adoption and outcome.

[13] Farms are distributed across classes from 1 to 7 (variables from SIZE1 to SIZE7) where class 1 represents the smallest size and class 7 the largest.

[14] Five farm types are considered: Livestock farms (PRO1), Livestock&crops farms (PRO2), Annual-crops farms (PRO3), Perennial-crops farms (PRO4), Other farms (PRO5).

[15] This variable is excluded in the case of treatment P6 where it appears as one of the outcome variables.

depends on the farm's specialization (due to the historical base of the Pillar 1 basic payment), on geographical location (due to internal convergence) and on the farmer's characteristics (for instance, due to the additional payment for young farmers). But PIL1 also affects farmers' decision-making due to the associated cross-compliance requirements (Sotte and Brunori, 2025).

The (Px1) vector $\boldsymbol{X}$ is thus made of the following 11 variables (P=11): SIZE, PRO, FOR, KW, UAA, LSU, AWU, ENE, MAT, AGE, GEN, ALT, LAT, LON, PIL1.[16] All these covariates, enter the estimation stage as biennial averages computed on the two pre-treatment years, namely 2013-2014 and 2015-2016 for the two balanced panels considered. In the case of categorical variables for which a change is observed, the category observed in the first year is used.

5. Estimation approach

As discussed in Section 2, for several RDP measures here considered early adopters are very few and often very peculiar. Late adopters then enter the treatment at different moments. Besides farms' heterogeneity in itself, heterogeneity in the response to the treatment may depend on the different external conditions farmers are exposed to when they enter the policy. Due to the staggered treatment, therefore, we have an entry time-specific, or cohort-specific, response to the treatment. It follows that the estimated average TE is not only an average across units but also across cohorts. Here we assume that, once units enter the treatment, they remain exposed to treatment thereafter (this assumption is also referred to as an *absorbing treatment*). Thus we exclude dynamic treatments (Lechner and Miquel, 2010). We also exclude time-varying TE after the entry, that is, the TE is observed once after the entry and it does not vary over years (Sun and Abraham, 2021). Nonetheless, a staggered treatment still brings about several methodological challenges.

Within a panel data environment, the DID identification and estimation approach, possibly combined with matching techniques, has become standard in the recent empirical literature in the field. However, the DID logic may suffer under staggered treatment both because of the poor number of early treated units and because of the possible violation of the PTA on which TE identification is grounded. DID identification and estimation admit different trends between treated and control units but assume they are parallel. However, the validity of this latter assumption is questionable whenever treated units, thus also the controls, vary over time, as the PTA should hold true for any entry time, that is, for any cohort. To tackle the former issue, a feasible alternative consists of the SCM approach (Abadie et al., 2010; Abadie et al., 2015; Abadie, 2021). The SCM can handle the low and varying number of treated units as well as their peculiarity compared to controls. However, though suitable also within a panel data environment, the SCM can suffer because of the farm heterogeneity across the time dimension, meaning both different, though parallel, trends and staggered entry into the treatment.

Some combinations of these two strategies can be helpful. For this purpose, the SDID has been recently developed (Arkhangelsky et al., 2021; Clarke et al., 2023). It aims to consistently estimate the average TE of a policy measure on the treated (or ATT) even under the possible violation of the PTA between treatment and control units on average. Our argument is that this approach seems particularly suitable to deal with all the abovementioned complications of performing CI on RDP measures under investigation. One of the main contributions of the present paper, therefore, consists in assessing whether SDID estimation performs well in the case of RDP measures and, above all, whether it really performs better than SC and DID also considering the differences across RDP measures.[17]

[16] The number of covariates is actually larger since, as illustrated, some of these 11 confounders behave as categorical variables, thus entering the estimation stage as a set of dummies. See also Annex 3.

[17] There is some similarity between the SDID approach and the Causal Machine Learning (CML) estimation that has recently gained popularity (Coderoni et al., 2024; Esposti, 2024) in the sense that, though differently, both use data to train systems (the synthetic pre-treatment counterfactuals in the SDID approach) and use information very flexibly (via weights in SDID estimation).

The SDID procedure requires as input, a balanced panel, a binary treatment, well-defined outcome variables and a set of confounders. The information provided by the dataset illustrated in Section 4 is thus fully suitable for this estimation approach. Though binary, the treatment variable can be time-variant (i.e., $T_{it} = 0{,}1$) since staggered adoption is admitted (Athey and Imbens 2022). In the particular setting of the SDID estimation, not always treated units can be included in the estimation. For estimation to proceed, at least two pre-treatment periods are required to determine control units. For this reason, as anticipated in Section 4, two different time periods are considered (either 2013-2022 or 2015-2022) depending on the policy measure under consideration (see Table 1).

Estimation proceeds as described in Arkhangelsky et al. (2021) and Clarke et al. (2023). Basically, the ATT estimates are generated from a Two-Way Fixed Effect (TWFE) regression specified as follows:

$$(7) \qquad Y_{it} = a + \alpha_i + \lambda_t + \tau_{p(s)} T_{it} + \boldsymbol{X}'_{it}\boldsymbol{\beta} + \varepsilon_{it}$$

Where: Y_{it} indicates the generic outcome variable that here alternatively takes the form of the private ($\Delta\pi_{it}$) or social (ΔG_{it}) performance discussed in Section 3; a expresses a time-invariant and farm-invariant constant term; α_i is the farm-specific time-invariant term; λ_t is the shared year-specific term; $\tau_{p(s)}$ is the ATT, the estimand, for the private (*p*) and social (*s*) outcome; $\boldsymbol{\beta}$ is the (*P*x1) vector of the time-invariant and farm-invariant coefficients associated to the exogenous time-varying covariates, $\boldsymbol{X}_{it}$; ε_{it} is the conventional spherical disturbance assumed i.i.d. ~$N(0,s^2)$.[18]

The SDID estimation looks for the parameter values that minimize the sum over the whole panel of squared residuals of (7) with *N* unit-specific (ω_i) and *Z* time-specific (γ_t) weights:

$$(8) \qquad \left(\hat{\tau}_{p(s)}, \hat{a}, \hat{\alpha}_i, \hat{\lambda}_t, \hat{\boldsymbol{\beta}}\right) = \underset{\tau_{p(s)}, a, \alpha_i, \lambda_t, \boldsymbol{\beta}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{Z} \left[Y_{it} - \left(a + \alpha_i + \lambda_t + \tau_{p(s)} T_{it} + \mathbf{X}'_{it}\boldsymbol{\beta}\right)^2 \right] \omega_i \gamma_t \right\}$$

The presence of farm-specific FE implies that SDID seeks to match treated and control units on pre-treatment trends, and not necessarily on both pre-treatment trends and levels, allowing for a constant difference between treatment and control units. Consistently with the SCM logic, the selection of farm weights, ω_i, aims to ensure that comparison is made between treated units and controls approximately following parallel trends prior to the policy adoptions. The selection of time weights, γ_t, aims to give more relevance to pre-treatment periods which are more similar to post-treatment periods, in the sense of finding a constant difference between each control unit's post treatment average, and pre-treatment weighted averages across all selected controls. Here, these unit-specific and time-specific weights, $\omega_i$and γ_t, are found following the same procedure proposed by Arkhangelsky et al. (2021, pp. 4091-4092) consisting of an iterative constrained optimization process that alternates between finding the unit-specific weights and time-specific weights minimizing the pre-treatment discrepancy between the treated unit and the synthetic control.

Equation (8) underlying the SDID estimation may imply a large computational burden particularly when there are many covariates, the sample is large and treated units are numerous. Alternatively, the actual estimation procedure proposed by Arkhangelsky et al. (2021) proceeds in two stages. The first stage regresses Y_{it} on $\boldsymbol{X}_{it}$ in order to obtain a consistent estimate of covariate coefficients ($\hat{\boldsymbol{\beta}}$) and estimated residuals $\left(Y_{it} - \boldsymbol{X}'_{it}\hat{\boldsymbol{\beta}}\right)$.[19] In the second stage, the latter enter the SDID estimation procedure illustrated in (8) where term $\boldsymbol{X}'_{it}\boldsymbol{\beta}$ now disappears and Y_{it} is replaced by the estimated residuals.

---

[18] To avoid perfect collinearity, as standard in fully saturated FE models, one α_i and one λ_t are set at zero.

[19] Following Clarke et al. (2023), in this first stage all covariates are standardized in order to resize the variability of very-high-variance confounders while capturing the same underlying variation in covariates. In addition, two alternative approaches can be followed to estimate $\boldsymbol{\beta}$ parameters. One adopts an interactive optimization procedure which additionally allows for the efficient calculation of optimal weights $\omega_i$and γ_t. This approach, however, can be highly computationally demanding and time consuming and may incur the risk of converging towards local minima. Alternatively, these parameters can be estimated much faster in a single step via OLS. This second procedure is adopted here.

Though originally designed for the block treatment, that is, all treated units enter the treatment in the same year (thus a single pre- versus post-treatment date can be used to conduct estimation), this SDID estimation approach can be extended to a staggered adoption case, as in the present study. This extension is achieved by specifying a (NxZ) adoption matrix $\mathbf{W}$ whose term w_{it} indicates whether at time t the i-th unit is treated (w_{it}=1) or not (w_{it}=0). In turn, $\mathbf{W}$ can be broken down into (Z-2) adoption-date-(or cohort-)specific matrices, with Z columns and a number of rows equal to the number of units entering the treatment in that specific year (thus, units belonging to the same cohort). With this adjustment, the SDID estimation illustrated above provides cohort-specific ATT estimates, as well as cohort-specific unit-specific and time-specific weights (ω_i and γ_t).[20] It follows that also the ATT over the whole sample can be computed by calculating a weighted average of the cohort-specific ATT, where weights are assigned on the basis of the relative number of treated units and time periods in each cohort in order to give more relevance to larger and early-entry cohorts.

Arkhangelsky et al. (2021) demonstrate that the abovementioned ATT estimators are asymptotically normal. This allows constructing confidence intervals provided that a consistent estimation of the respective variance is available. Arkhangelsky et al. (2021) propose a block (or clustered) bootstrap approach. Though computationally demanding, this bootstrap procedure shows particularly good properties. However, estimated variance and confidence interval may be less reliable when the number of treated units is small. In these cases, and particularly when there is a single treated unit, an alternative placebo (or permutation-based) inference procedure can be adopted.

It seems finally interesting to stress how and to what extent this SDID estimation differs compared to the standard DID and SCM procedures. DID estimates can be obtained from (8) by simply assigning equal weights to all units and time periods, whereas the SCM estimation can be obtained from (8) by maintaining optimally chosen farm-specific weights (ω_i) and assigning equal weights to all years. Moreover, in (8) the SCM omits the farm-specific fixed effects α_i, thus implying that the synthetic control and treated units should maintain approximately an equivalent pre-treatment level, as well as trend. Therefore, the SDID estimation offers greater flexibility than the DID and SCM procedures. Compared to DID estimations, it permits a violation of the PTA in aggregate data. Compared to the SCM, it aims to optimally weigh time periods when considering counterfactual outcomes and allows for level differences between treatment and control groups.

Ultimately, both DID and SCM estimations have pros and cons. On the one hand, the DID is a special case of the SCM where all unit-specific weights are equal. On the other hand, the DID estimation controls for the fixed effects while the SCM does not. This difference may also explain why the SCM tends to perform better than the traditional DID when the number of treated units is small. The flexibility granted by the SDID estimation approach aims to magnify the pros of both approaches while minimizing the respective drawbacks by looking for the best compromise between the two given the specific case and data under study (Lu, 2021). For this reason, applying the SDID estimation to different policy measures over the same time period, as in the present study, may provide interesting insight into the circumstances that make the actual advantage of the SDID estimation particularly evident.

6. Results

Table 3 reports ATT estimates for both the private and societal outcome variables over the whole period (namely, over all cohorts) for the six policy treatments.[21] It clearly emerges that for some of

[20] It is worth stressing that the cohort-specific ATT estimates provided by the SDID approach have nothing to do with the so-called carryover effect or dynamic treatment effect. Even though treatment is staggered, for any cohort there is only one ATT, that is, the treatment effect does not change over time in the post-treatment period (Sun and Abraham, 2021).

[21] All results have been obtained with the SDID command ran in STATA 18.5 software. Annex 3 reports the estimates of the covariates' coefficients ($\boldsymbol{\beta}$in equations (7) and (8)) for a selected treatment, P4. Since these estimates are cohort-specific, Annex 4 only reports

them (P1 and P2, in particular) no estimated ATT is statistically different from zero, meaning that there is no clear evidence of an impact of these RDP measures on both the private and societal performance.[22] This result could have two different interpretations. The first reading is that, in the case of the private outcome, the adoption of these policies does not bring about any significant before-support income gain for the farm. In other words, these measures would not be adopted if an economic incentive (the monetary support associated to the measure) were not granted. In the case of the societal outcome, according to this interpretation, the lack of a significant impact would indicate that, beside the alleged and declared policy measures, these measures simply turn out to be an income support to farmers', so they consist in a purely income redistributive transfer from the rest of the society to the farmers without any appreciable counterpart or justification for the society itself.

Though this interpretation is not new within the literature on the real underlying motivations of the RDP measures (Esposti, 2022), it should be acknowledged that these results for P1 and P2 can be more prudently interpreted in terms of an inconclusive evidence on the actual impact of these measures on farmers' behaviour and performance. Inconclusiveness that can be arguably attributable to the lack of more robust information due to the limited number of post-treatment years,[23] the limited number of treated units especially in the early years and, as will be discussed below, the highly year-by-year volatility of the outcome variables in use. In fact, while this latter interpretation seems to be reasonable for P2 where both the private and societal ATT show small magnitude, it seems more questionable for P1. In this case, both outcome variables show a response of remarkable magnitude (more than 35 thousand € per unit of FAWU in the case of the private outcome), but the high variance associated to these estimates makes this evidence inconclusive.

For policies P3, P4, P5 and P6, all estimated private ATT are significantly different from zero and positive. They are all quite consistent in magnitude as estimates range between about 8.5 thousand € per unit of FAWU (for P5) to about 38 thousand € per unit of FAWU (for P6). On the one hand, this seems consistent with the rationale of farmers' decision-making as illustrated in Section 3: a negative impact on the private outcome, if larger than the underlying policy support, would imply a negative net income as a consequence of the policy adoption. On the other hand, however, the evident policy implication of these results is that for all these measures farmers obtain a monetary advantage that goes beyond the policy support itself. In other words, and in principle, the societal outcome could have been obtained even without the policy support as farmers would still have found an advantage to produce that response.

In the case of these four policies, also results (ATT) for the societal outcome seem quite robust though slightly less statistically significant. In all cases, the societal outcome moves in the expected direction (positive for P3, P4 and P6 and negative for P5). For P4 and P5, however, one of the candidate social outcomes does not show a statistically significant difference from 0 ATT. In any case, it can be concluded that these measures demonstrate to be effective in inducing the response of societal interest, but their efficiency is questionable since this response would have been obtained, in principle, even with a lower (possibly zero) public expenditure.

As discussed, however, besides these ATT averaged over the whole period and cohorts, SDID estimation also returns cohort-specific ATT providing interesting evidence on the volatility of farmers' response to these measures across the different adoption times. Table 4 provides evidence about this time heterogeneity by reporting the cohort-specific ATT estimates. Table 4 confirms the

coefficient estimates for P4 as it only presents two cohorts (2017 and 2020) but still shows a statistically significant overall ATT for both the private and societal outcome variables. Therefore, Table A4 reports four coefficient vectors, two for each cohort and two for each outcome variable. Estimates for all other treatments are available upon request.

[22] The ATT estimate standard errors is computed following the bootstrap procedure described in Arkhangelesky et al. (2021). As clarified in Clarke et al. (2023), for treatments in which there are cohorts with only one treated unit (as P4 and P6 in the present case) the bootstrap procedure is replaced by the placebo inference.

[23] It is worth noticing, however, that his lack of statistical significance is found for both period 2013-2022 (P1) and period 2015-2022 (P2).

cohort-dependency of the ATT estimates, both private and societal. These latter, in particular, are not only volatile, thus instable, in magnitude but also in statistical significance and, sometimes, even in sign. This volatility may eventually explain the weak (or lack of) statistical significance of the ATT computed over the whole period at least for some policy measures.

More in detail, in the case of P1 only the first cohort (2015) shows a statistical significance for private and social ATT, while significance is found only in the last two cohorts for policy P2. P4 reports statistical significance for both available cohorts and for both the private and societal outcome variables. The other policies (P3, P5, P6), for both outcomes, alternate years showing significant ATT with others in which significance is lost though the sign remains the same (with the only exception of P5).

While Table 4 demonstrates the heterogeneity of the cohort-specific ATT estimates, Figures 2-7 display one of the major causes of this heterogeneity, namely, the large volatility of outcome variables over time.[24] Heterogeneity across cohorts, therefore, is not only in terms of different ATT but also in the volatility of outcome variables that inevitably reflects in the cohort-specific ATT estimation. This evidently holds true for the private outcome which is always the same across all measures. But it also seems true for the different societal outcomes across measures.

These graphs show the trends of the outcome variables for both the treated units and the synthetic controls before and after the entry and this makes it possible to visualize the impact of the treatment but also the possible violation of the PTA. The weights used to average pretreatment time periods are also reported at the bottom of the graphs (shaded areas) to emphasize how the SDID estimation flexibly makes use of the pre-treatment observations.[25] The estimated effect can be detected by the deviation from a parallel trend between treated and control unis after the treatment. It is worth noticing that, as stressed in Arkhangelsky et al. (2021, Figure 1), the ATT applies to the year of the treatment but since then it deviates from the trend for any year after the treatment so the deviation from parallelism applies over the whole post-treatment period.

Therefore, for instance, a positive ATT implies that the treated units either get closer or even overcome the synthetic control if it was underperforming in the pre-treatment period (see Figure 2b) or divaricates from the synthetic control if it was overperforming in the pre-treatment period (see Figure 4b). In fact, all possible cases are found. There are cases for which the treatment makes the treated units closer to the controls (Figure 2b), and other cases in which they divaricate (Figure 6a); they also signals both positive (all figures referring to the private outcome variable) and negative (Figure 6b) effects consistently with what reported in Table 4. Another interesting consideration about the volatility of outcome variables before and after the treatment is that one often disregarded (and often unintended) effect of a policy may be to stabilize (or destabilize) the outcome variable beside affecting its level.

The key message here is that unlike Arkhangelsky et al. (2021) and Clarke et al. (2023), the time dimension does not bring about a clear and regular dynamics but an highly volatile and unstable movement of outcome variables. This might not surprise when farm-level data are in use. However, it implies that the DID logic, being it based on the PTA, is by itself less helpful. Moreover, this may explain why in recent studies concerning farm-level data the time dimension has been collapsed into a cross-section dataset in order to get rid of this volatility (Baldoni and Esposti, 2024; Coderoni et al., 2024). Alternatively, as done in the present study, the SDID estimation can more flexibility take care of the time dimension without assuming unlikely parallel trends. Table 4 and Figures 2-7 highlight

[24] Only exemplary cohorts are reported. They are selected by looking at the cohort-specific ATT (significance and magnitude) but, above all, concentrating on cohorts in the middle of the period to better visualize some years before and some years after, to appreciate volatility pre and post-treatment but also allowing for an appreciable number of treated units. All other years are available upon request.
[25] Unit-specific weights associated to any SDID estimation are available upon request.

the advantages of the SDID compared to DID and SCM when unstable outcome data and these kind of policies are under consideration.

Table 5 and Figures 8-11 compare SDID estimates with the respective DID and SCM counterparts for a selected group of measures. Table 5 reports both private and societal ATT over the whole period while, for the sake of space limitations, figures concern only the private outcome variable for the policy measures showing a statistically significant SDID ATT (P3, P4 P5 and P6; see Table 4). We select only one cohort per policy treatment with a statistically significant ATT but different across measures in order to represent all the periods considered.[26] The objective here is only to show how the SDID looks for a compromise between DID and SCM estimations and this compromise depends on the specific characteristics of the treatment under consideration, in particular the number of treated units.

The volatility of the outcome variables discussed above implies that the validity of DID estimates may be questionable since PTA barely holds true. In principle, the advantage of the SCM on DID estimation concerns cases with very few (early) adopters (such as P1 and P4, in the present study). Table 5 confirms the SDID advantage over both SCM and DID estimations but it seems more evident compared to SCM. For both cases, SDID shows higher efficiency: no SCM ATT estimate is statistically different from 0, while few (P5 and P6, in particular) are significant under the DID estimation. The flexible use of the panel informative set made by SDID estimation evidently brings about this gain in efficiency particularly in admitting different trends between treated units and controls (like SCM and unlike DID). To stress this, Figures 8-11 visually compare SDID, DID and SC as in Arkhangelsky et al. (2021). DID and SCM alone might not work properly, while flexibly combining them in the more data-driven SDID estimation allows for a more efficient use of the available information within the panel.

Finally, Table 6 reports evidence on policy efficiency using the SCFP indicator illustrated in Section 3 (equation (6)). It is a sort of back-of-the-envelope calculation based on the average ATT estimates in Table 3 and on the average expenditure per unit of FAWI. SO or UAA as reported in the first three columns of Table 6. Calculation is performed on selected policies for which the private and societal ATT are statistically different from 0 over the whole period, in order to show how different this policy indicator can be. The SCFP indicator, across the different policies, points to an apparent low efficiency since a unit of societal outcome costs up to several thousand Euros. Evidently, these results are only indicative as they suggest room for efficiency improvement and, more importantly, confirmation in further studies.

7. Concluding remarks

The wide recent literature on the assessment of RDP measures adopting a CI logic has often disregarded several complications that these policies, and the farming context, bring about. These complications may be summarized by the term heterogeneity. But it does not only refer to farmers' characteristics but also, and more importantly, to farmers' response to the treatment, thus to the TE itself. This latter heterogenous response, in turn, may also depend on the different entry times and, consequently, on the different pre vs. after-treatment trends of outcome variables, namely on their volatility.

Assessing policy impact, effectiveness and efficiency, is very challenging under these circumstances. The SDID estimation approach here adopted seems appropriate to take some characteristics of these policies into account. It is flexible in using information to better account for the abovementioned heterogeneities. At the same, SDID estimates also clearly reveal the consequent instability and lack of robustness of the cohort specific TE and of the trends of outcome variables.

[26] All other figures are available upon request.

Results confirm the validity of the SDID approach compared to alternative DID and SCM estimations. However, some results show weak statistical robustness and are not of immediate policy interpretation. Empirical evidence thus points to improvements in the methodological approach and in the adopted dataset, particularly in terms of longer pre and post-treatment observation periods and in a better definition of the outcome variables of societal interest.

REFERENCES

Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature* 59(2): 391–425. https://www.aeaweb.org/articles?id=10.1257/jel.20191450.

Abadie, A., Diamond, A., Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association* 105(490): 493–505.

Abadie, A., Diamond, A., Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science* 59(2): 495-510.

Arkhangelsky, D., Athey, S., Hirshberg, D.A., Imbens, G.W., Wager, S. (2021). Synthetic Difference-in-Differences. *American Economic Review*, 111 (12), 4088-4118.

Athey, S., Imbens, G.W. (2022). Design-based analysis in Difference-In-Differences settings with staggered adoption. *Journal of Econometrics* 226(1): 62-79.

Baldoni, E., Coderoni, S., Esposti, R. (2021). Immigrant workforce and agriculture productivity: evidence from Italian farm-level data, *European Review of Agricultural Economics* 48 (4): 805–834.

Baldoni, E., Esposti, R. (2023). Estimating the impact of policies under spatial interference. The case of CAP support to organic farming. Quaderno di Ricerca n. 475, Dipartimento di Scienze Economiche e Sociali, Università Politecnica delle Marche.

Brown, C. Kovács, E., Herzon, I., Villamayor-Tomas, S., Albizua, A., Galanaki, A., Grammatikopoulou, I., McCracken, D., Alkan Olsson, J., Zinngrebe, Y. (2021). Simplistic understandings of farmer motivations could undermine the environmental potential of the common agricultural policy. *Land Use Policy* 101, 105136.

Camaioni B., Esposti, R., Pagliacci F., Sotte F. (2016). How Does Space Affect the Allocation of the EU Rural Development Policy Expenditure? A Spatial Econometric Assessment. *European Review of Agricultural Economic*s, 43(3), 433–473.

Camaioni, B., Sotte, F. (2010). Un primo bilancio della politica di sviluppo rurale in Europa. *Agriregionieuropa* 6(20): 41-45.

Cerulli, G. (2015). *Econometric Evaluation of Socio-Economic Programs: Theory and Applications*. Berlin: Springer.

Clarke, D., Pailañir, D., Athey, S., Imbens, G. (2023). Synthetic Difference-in-Differences Estimation. IZA-Institute of Labour Economics, Discussion Paper No. 15907, Bonn.

Coderoni, S., Esposti, R. (2018). CAP payments and agricultural GHG emissions in Italy. A farm-level assessment. Science of the Total Environment 627: 427-437.

Coderoni, S., Esposti, R., Varacca, A. (2024). How differently do farms respond to agri-environmental policies? A probabilistic machine-learning approach. *Land Economics* 100(2): 370-397. DOI: 10.3368/le.100.2.060622-0043R1.

Coderoni, S., Esposti, R., Varacca, A. (2025). Evaluating the Dual Impact of Organic Farming Support: Private Response and Societal Outcome for GHG Emission Abatement. A Causal Machine Learning assessment. Unpublished manuscript, Department of Economics and Social Sciences, Università Politecnica delle Marche.

Eichorn, T., Schaller, L., Hamunen, K., Runge, T. (2024). Exploring acro-environmental factors influencing adoption of result-based and collective agri-environmental measures: a PESTLE approach based on stakeholder statements. *Bio-based and Applied Economics* 13(1), 47-91.

Esposti, R. (2017a) The Empirics of Decoupling: Alternative Estimation Approaches of the Farm-level Production Response. *European Review of Agricultural Economics*, 44(3), 499-537.

Esposti, R. (2017b). The heterogeneous farm-level impact of the 2005 CAP-first pillar reform: a multivalued treatment effect estimation. *Agricultural Economics*, 48(3), 373-386.

Esposti, R. (2022). The Coevolution of Policy Support and Farmers Behaviour and Performance. An investigation on Italian agriculture over the 2008-2019 period. *Bio-Based and Applied Economics* 11(3): 231-264.

Esposti, R. (2024). Non-Monetary Motivations of Agri-Environmental Policy Adoption. A Causal Forest Approach. *Journal of Environmental Management* 352, 119992. DOI: 10.1016/j.jenvman.2023.119992.

Esposti, R., Sotte, F. (2013). Evaluating the Effectiveness of Agricultural and Rural Policies: An Introduction. *European Review of Agricultural Economics*, 40(4), 535-539.

European Commission (2012). Strategic Programming and Monitoring and Evaluation for RDPs 2014-2020. Joint ENRD Coordination Committee and Evaluation Expert Committee workshops, Newsletter 14th and 15th March, Brussels.

European Commission (2014). Getting the most from your RDP: Guidelines for the ex-ante Evaluation of 2014-2020 RDPs. European Evaluation Network for Rural Development, Brussels.

European Commission (2015a): Establishing and implementing the Evaluation Plan of 2014-2020 RDPs. Guidelines. Directorate-General for Agriculture and Rural Development – Unit E.4, Brussels.

European Commission (2015b). Common Evaluation Questions for Rural Development Programmes 2014-2020. Directorate-General for Agriculture and Rural Development – Unit E.4, Working Paper, Brussels.

European Commission (2016). Guidelines: Assessment of RDP results: how to prepare for reporting on evaluation in 2017. Directorate-General for Agriculture and Rural Development – Unit E.4, Brussels.

European Commission (2018). Guidelines. Assessing RDP achievements and impacts in 2019. Directorate-General for Agriculture and Rural Development – Unit C.4 (2018): Brussels.

European Commission (2019). Interactive CAP Indicator Dashboards: transparency and growth for a more sustainable tomorrow's. Directorate-General for Agriculture and Rural Development, European Evaluation Helpdesk for Rural Development, Rural Evaluation NEWS, Number 12, Brussels.

Lechner, M., Miquel, R. (2010). Identification of Effects of Dynamic Treatments by Sequential Conditional Independence Assumptions. *Empirical Economics* 39(1): 111-137.

Lu, J. (2021). Synthetic Control Method, Stationarity and Pointwise Statistical Inference. http://www.uone-tech.cn/ChinaStataConference2021/pdf/lujiaxuan.pdf.

Mennig, P. (2024). A never-ending story of reforms: on the wicked nature of the Common Agricultural Policy. *npj Sustain. Agric.* 2, 20. https://doi.org/10.1038/s44264-024-00027-z.

Sotte F. (a cura di). (2009). La politica di sviluppo rurale 2007-2013. Un primo bilancio per l'Italia, Quaderni del Gruppo 2013- Edizioni Tellus, Roma.

Sotte, F., (2012). La Politica di Sviluppo Rurale. In: De Filippis, F. (a cura di). La nuova PAC 2014-2020. Un'analisi delle proposte della Commissione. Edizioni Tellus, Roma: 89-105.

Sotte, F., Brunori, G. (2025). *European Agricultural Policy: History and Analysis*. Springer Nature Switzerland, Basel.

Stetter, C., Mennig, P., Sauer, J. (2022). Using Machine Learning to Identify Heterogeneous Impacts of Agri-Environment Schemes in the EU: A Case Study. *European Review of Agricultural Economics* 49(4): 723-739.

Sun, L., Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics 225(2): 175-199.

Swales, V. (2007). The Role of Pillar 2 in Delivering Environmental Outcomes. A paper prepared by for the LUPG/BfN conference 'Future Policies for Rural Europe 2013 and beyond', Brussels, 19 – 20 September.

Targetti. S., Niedermayr, A., Hafner, K., Schaller, L. (2024). New pathways for improved delivery of public goods from agriculture and forestry. *Bio-based and Applied Economics* 13(1), 3-11.

Zhang W., Cotton C.A. (2023). Causal inference for recurrent events via aggregated marginal odds ratio. *Statistics in Medicine* 42(18): 3208-3235. doi: 10.1002/sim.9802.

Table 1 – List of EU RDP 2014-2022 measures included and excluded in the study with association with respective outcome variables.

Priority & Focus Area	Treatment	RDP Measures (active behavioural response) (Annex 1)	Theme/Reclassification	Period	Outcome variable		Reference indicators (Annex 2)
					Private	Societal	
Pr1 & 1A-1C	P1	1.1, 2.1, 2.2, 16.1	Information, Education, Training	2013-2022	NI/FAWU	GVA/AWU, GVA/UAA	I2, I3
Pr3 & 3A	P2	3.1, 3.2, 16.2, 16.4	Food chain, Quality	2015-2022	NI/FAWU	GVA/SO	R2, R4, T6, I1, I2
Pr2 & 2A	P3	4.1-4.4, 5.1, 5.2	Structural agricultural investments	2013-2022	NI/FAWU	Inv/AWU Inv/SO	C28, R1, T4
Pr2 & 5A-5C, 6A	P4	6.4, 16.9	Structural non-agricultural investments	2015-2022	NI/FAWU	OGA/SO, OGA/AWU, Other Inv/SO	C28, R1, R14, R15, T4
Pr3 & 5E	P5	14	Animal/livestock production management	2013-2022	NI/FAWU	LSU/AWU LSU/UAA VE/SO	C21, C45, O8, R16, T17, I7
Pr4-5 & 4A, 5E	P6	15.1-15.2, 8.1-8.6	Forests	2013-2022	NI/FAWU	FA/UAA	C38, R6, R18, R9, R11, R20, T8, T11, T13, T19, I9

LEGEND: AWU = Annual Working Units; FA = Forest Area; FAWU= Family Annual Working Units. GVA = Gross Value Added; Inv = Investments; LSU = Livestock Units; NI=before-support Net farm Income; OGA = Other Gainful Activities; SO = Gross Production Value; UAA = Utilised Agricultural Area; VE = Veterinary and sanitary Expenditure.

Table 2 – Number of treated farms by policy in the two alternative panels during the period (2015-2022)

Treatment	2015	2016	2017	2018	2019	2020	2021	2022	Never treated units[a]	Total panel
Unbalanced Panel										
P1	3	3	3	4	6	7	8	9	9566-11075	
P2	0	2	10	32	52	63	69	74	9569-11010	
P3	6	38	66	134	218	307	366	402	9563-10726	
P4	0	0	2	4	7	14	17	19	9569-11065	
P5	28	268	382	420	482	455	494	511	9541-10573	
P6	1	33	34	49	70	74	77	81	9568-11003	
Balanced Panel										
P2	0	0	3	10	17	32	38	43	3856	3899
P4	0	0	1	1	1	3	3	3	3896	
Balanced Panel (2013-2022)										
P1	2	2	2	2	2	3	3	3	2602	2605
P3	2	13	21	41	55	65	80	98	2507	
P5	3	79	113	130	142	148	153	160	2445	
P6	1	14	17	25	32	31	36	40	2565	

[a] In the unbalanced panel the yearly minimum and maximum number of untreated units during the period are reported

Table 3 - Private and societal ATT by treatment (estimated standard errors in parenthesis)

	Private ATT ($\Delta\pi$) (.000€)	Societal ATT (ΔG)[a] (1)	Societal ATT (ΔG)[a] (2)	Societal ATT (ΔG)[a] (3)
P1	32.80 (31.1)	192.2 (658.3) [GVA/AWU]	519.6 (521.9) [GVA/UAA]	-
P2	-0.244 (5.27)	0.152 (0.859) [GVA/SO]	-	-
P3	11.06 (5.19)**	812.5 (304.6)** [INV/AWU]	0.005 (0.002)** [INV/SO]	-
P4	37.63 (12.3)**	2.757 (1.359)** [OGA/SO]	5581 (4314) [OGA/AWU]	-
P5	8.443 (3.861)**	-0.084 (0.146) [LSU/UAA]	-1.468 (0.763)* [LSU/AWU]	-0.006 (0.003)** [VE/SO]
P6	38.30 (12.93)**	0.045 (0.024)* [FA/UAA]	-	-

*,**: Statistically significant at 5% and 10% level, respectively.
[a]Societal outcome variable in square brackets.

Table 4 – Cohort-specific (i.e., by year of entry in the treatment) private and societal ATT (standard errors in parenthesis).

	P1[a]	P2	P3[b]	P4[c]	P5[d]	P6
Private ATT ($\Delta\pi$) (.000€)						
2015	35.3 (17.3)**	-	36.4 (33.5)	-	7.625 (8.601)	25.130 (12.060)**
2016	-	-	34.08 (16.8)**	-	22.03 (11.65)**	15.39 (9.024)*
2017	-	18.1 (25.7)	8.06 (15.3)	2.606 (1.293)**	13.69 (13.81)	13.16 (13.61)
2018	-	-5.81 (12.9)	3.53 (18.0)	-	14.19 (13.187)	36.11 (14.68)**
2019	-	-1.18 (6.76)	2.07 (17.1)	-	7.261 (7.283)	32.39 (6.940)**
2020	19.5 (30.3)	3.21 (5.54)	11.5 (5.51)**	74.91 (34.42)**	8.057 (7.629)	35.536 (19.43)*
2021	-	7.59 (3.96)*	17.4 (9.25)*	-	14.09 (5.581)**	32.95 (26.97)
2022	-	4.82 (2.33)**	7.18 (15.8)	-	13.07 (6.535)**	18.42 (7.144)**
Total ATT	32.80 (31.1)	-0.244 (5.27)	11.06 (5.19)**	37.63 (12.3)**	8.443 (3.861)**	38.30 (12.93)**
Societal ATT (ΔG)						
2015	7110 (1973)**	-	182.4 (294.8)	-	-0.007 (0.004)	0.014 (0.012)
2016	-	-	1871 (1131)	-	0.000 (0.001)	0.071 (0.029)**
2017	-	-0.182 (0.396)	2378 (1069)**	5.598 (0.416)**	0.001 (0.003)	-0.014 ((0.021)
2018	-	-0.276 (3.41)	104.3 (131.2)	-	-0.002 (0.001)**	0.021 (0.026)
2019	-	0.048 (0.055)	198.1 (259.0)	-	-0.003 (0.001)**	0.078 (0.039)**
2020	-6707 (7969)	0.019 (0.088)	525.9 (286.3)*	0.823 (0.443)*	-0.002 (0.001)**	0.013 (0.018)
2021	-	0.225 (0.045)**	1526 (755.3)**	-	-0.003 (0.001)**	0.002 (0.015)
2022	-	0.141 (0.040)**	149.2 (328.8)	-	-0.003 (0.002)*	0.035 (0.019)*
Total ATT	519.6 (521.9)	0.152 (0.859)	812.5 (304.6)**	2.757 (1.359)**	-0.006 (0.003)**	0.045 (0.024)*

*,**: Statistically significant at 5% and 10% level, respectively.
[a]Societal outcome: GVA/UAA; [b]Societal outcome: INV/AWU; [c]Societal outcome: OGA/SO; [d]Societal outcome: VE/SO

Table 5 – Comparison of SDID, DID and SCM ATT estimates by policy (standard errors in parenthesis).

	P1[a]	P2	P3[b]	P4[c]	P5[d]	P6
Private ATT ($\Delta\pi$) (.000€)						
SDID	32.80 (31.1)	-0.244 (5.27)	11.06 (5.19)**	37.63 (12.3)**	8.443 (3.861)**	38.30 (12.93)**
DID	19.91 (29.5)	-0.679 (5.22)	8.49 (9.36)	15.22 (19.6)	10.23 (2.35)**	38.47 (20.9)*
SCM	27.70 (64.1)	-1.052 (3.11)	5.07 (8.04)	16.48 (13.6)	5.30 (5.21)	46.04 (37.13)
Societal ATT (ΔG)						
SDID	519.6 (521.9)	0.152 (0.859)	812.5 (304.6)**	2.757 (1.359)**	-0.006 (0.003)**	0.045 (0.024)*
DID	381.4 (463.0)	-1.44 (2.55))	849.8 (936.0)	2.710 (2.113)	-0.005 (0.002)**	0.161 (0.138)
SCM	390.3 (808.6)	0.960 (2.86)	307 1(504.5)	1.789 (1.889)	-0.0065 (0.077)	0.053 (0.057)

*,**: Statistically significant at 5% and 10% level, respectively.
[a]Societal outcome: GVA/UAA; [b]Societal outcome: INV/AWU; [c]Societal outcome: OGA/SO; [d]Societal outcome: LSU/AWU

Table 6 – Social Cost of Farm Performance (SCFP) for selected societal ATT by policy.

	Avg. S_i on FAWU[a]	Avg. S_i on SO[a]	Avg. S_i on UAA[a]	Societal ATT $(\Delta G)^b$ (1)	Societal ATT $(\Delta G)^b$ (2)	Societal ATT $(\Delta G)^b$ (3)
P3	38668.8	0.384	2135.8	47.58 [INV]	76.77 [INV]	-
P4	46945.8	0.466	2593.0	1.697 [OGA]	8.407[c] [OGA]	-
P5	10596.2	0.105	585.3	-6963[c] [-LSU]	-7214 [-LSU]	-17.52 [-VE]
P6	4628.6	0.046	255.7	5649 [FA]	-	-

[a]The average refers only to treated units.
[b]Societal outcome variable in square brackets.
[c]Not statistically significant ATT

Figure 1 – Evolution of the treated farms across the period under consideration in: a) units; b) as percentage on the balanced panel.

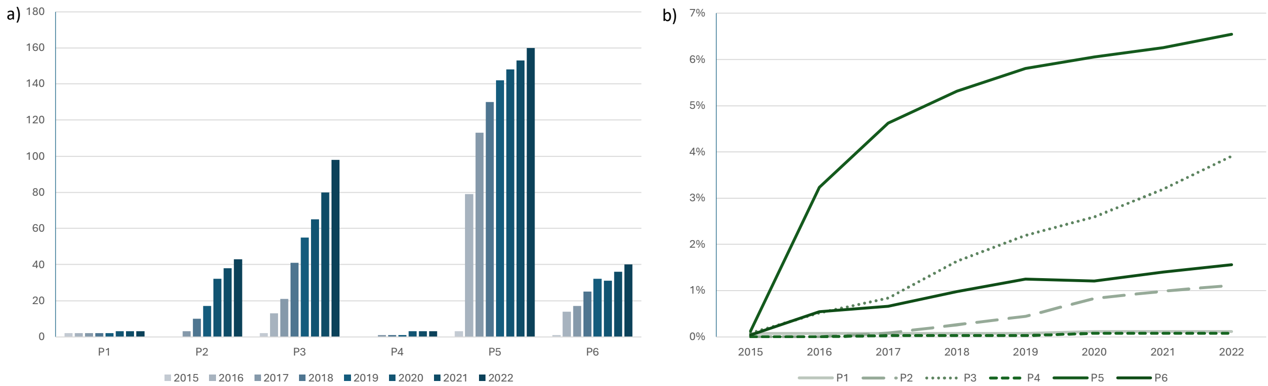


Figure 2 – Outcome trends for units entering treatment P1 in year (cohort) 2015: a) private outcome (NFI/FAWU); b) societal outcome (GVA/UAA). The shaded area indicates the time-specific weights.

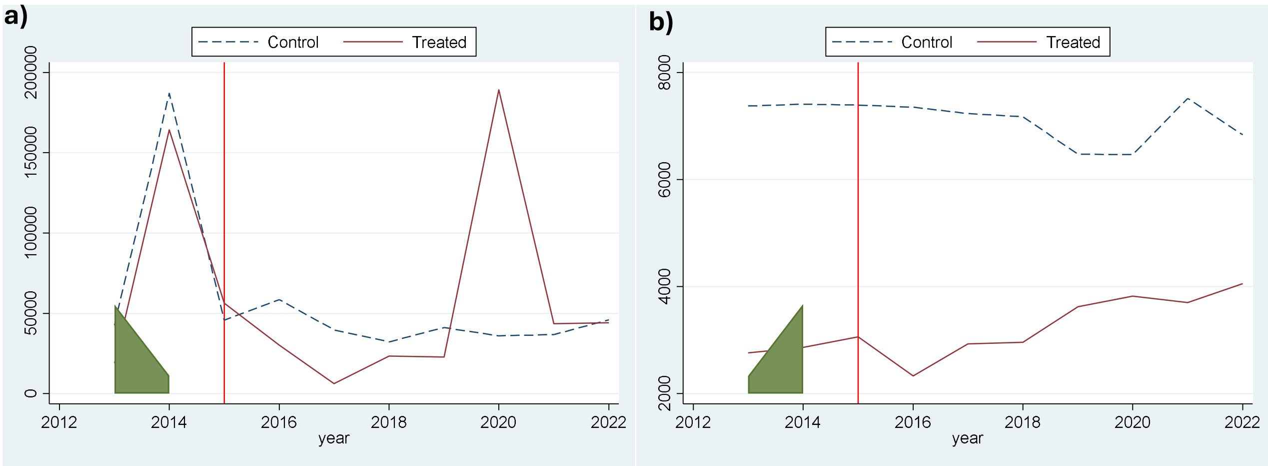


Figure 3 - Outcome trends for units entering treatment P2 in year (cohort) 2021: a) private outcome (NFI/FAWU); b) societal outcome (GVA/SO). The shaded area indicates the time-specific weights.

Figure 4 - Outcome trends for units entering treatment P3 in year (cohort) 2021: a) private outcome (NFI/FAWU); b) societal outcome (INV/UAA). The shaded area indicates the time-specific weights.

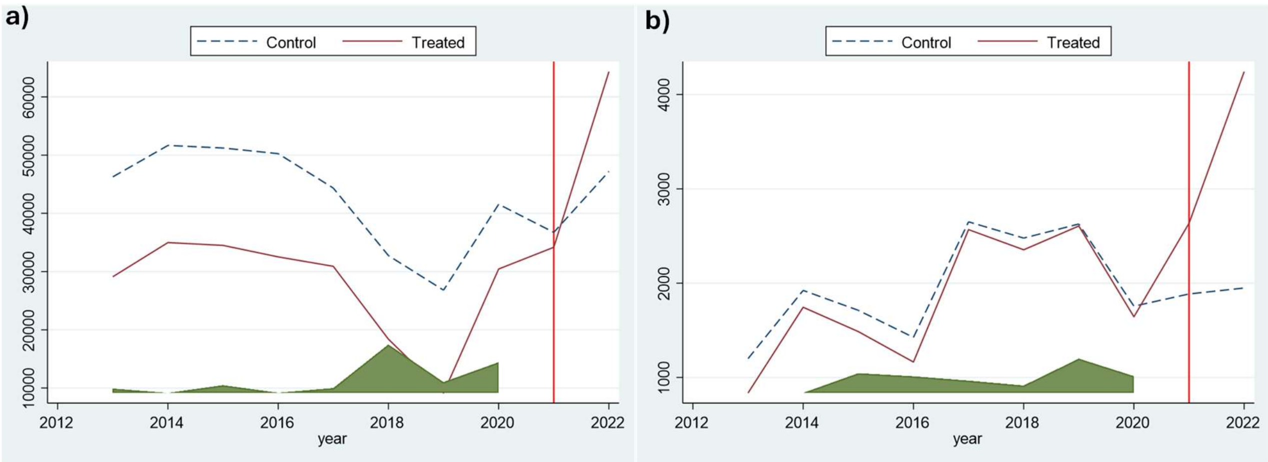


Figure 5 – Outcome trends for units entering treatment P4 in year (cohort) 2017: a) private outcome (NFI/FAWU); b) societal outcome (OGA/SO). The shaded area indicates the time-specific weights.

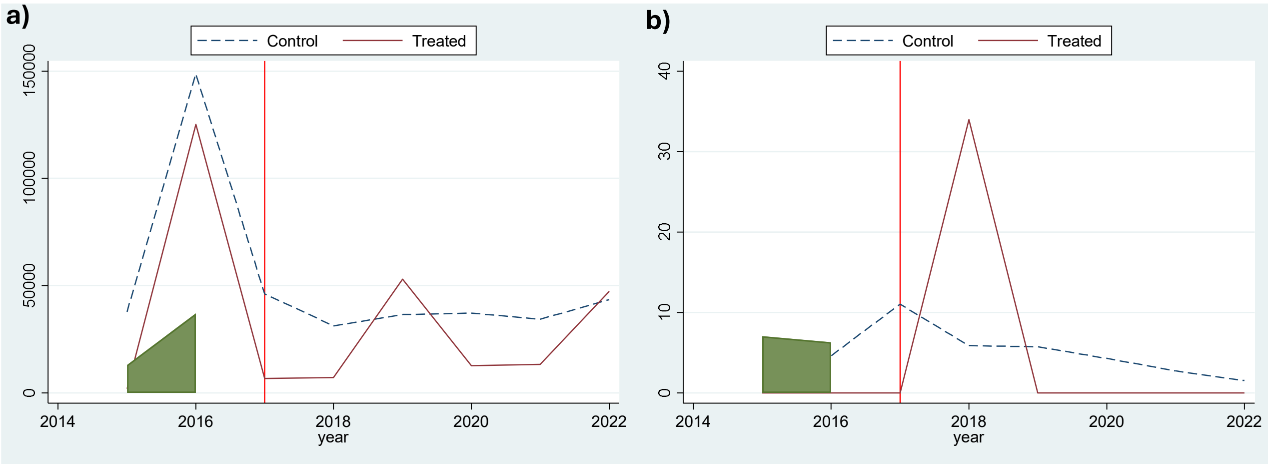


Figure 6 - Outcome trends for units entering treatment P5 in year (cohort) 2016: a) private outcome (NFI/FAWU); b) societal outcome (LSU/AWU). The shaded area indicates the time-specific weights.

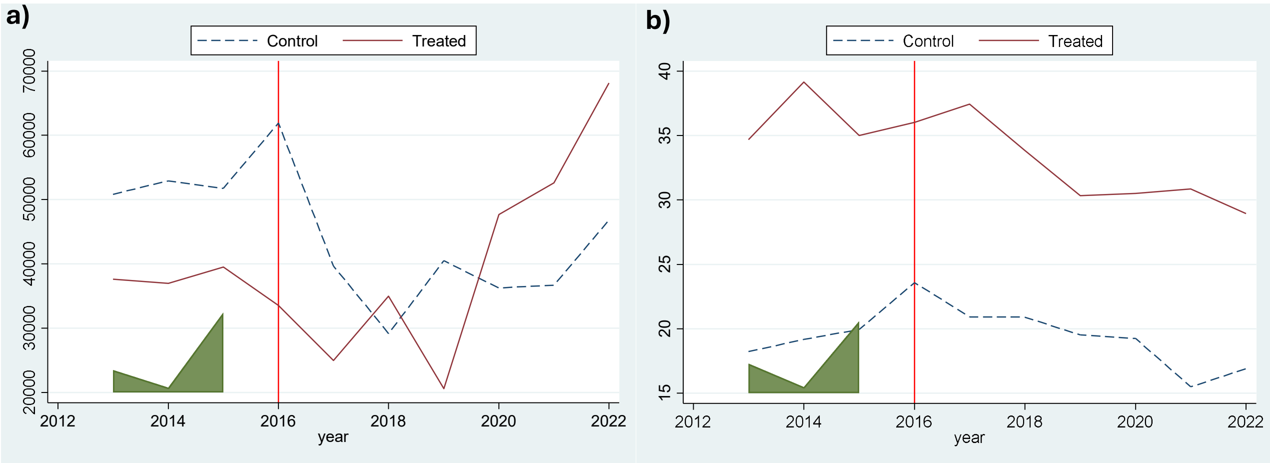


Figure 7 - Outcome trends for units entering treatment P6 in year (cohort) 2019: a) private outcome (NFI/FAWU); b) societal outcome (FA/UAA). The shaded area indicates the time-specific weights.

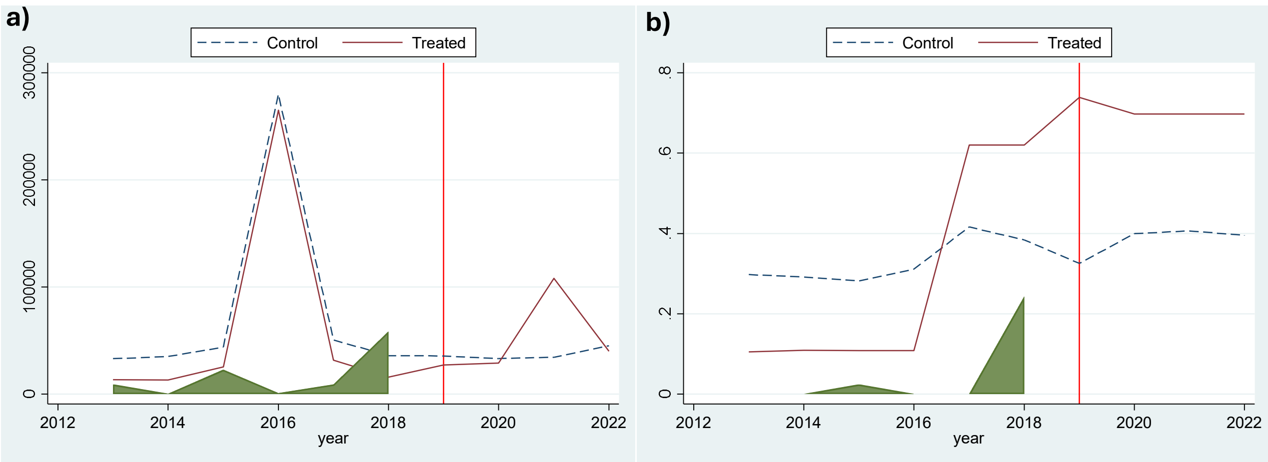


Figure 8 – Comparison of SDID, DID and SCM estimates for P3, entry year (cohort) 2021.

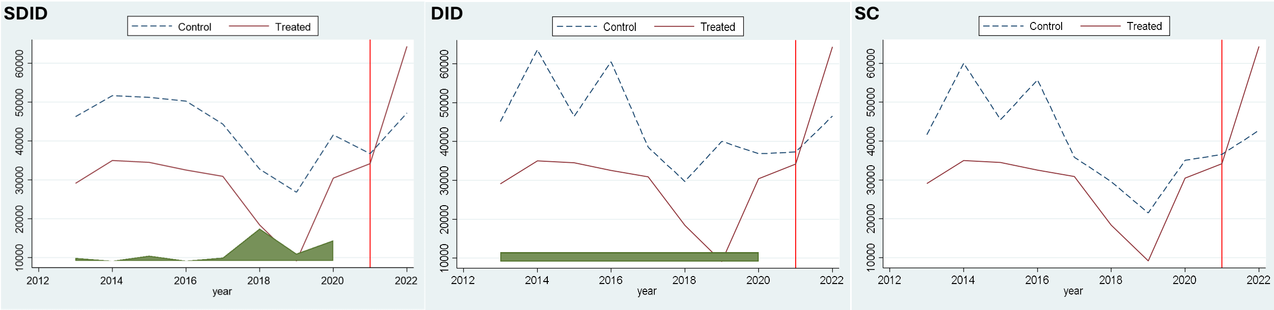


Figure 9 – Comparison of SDID, DID and SCM estimates for P4, entry year (cohort) 2017.

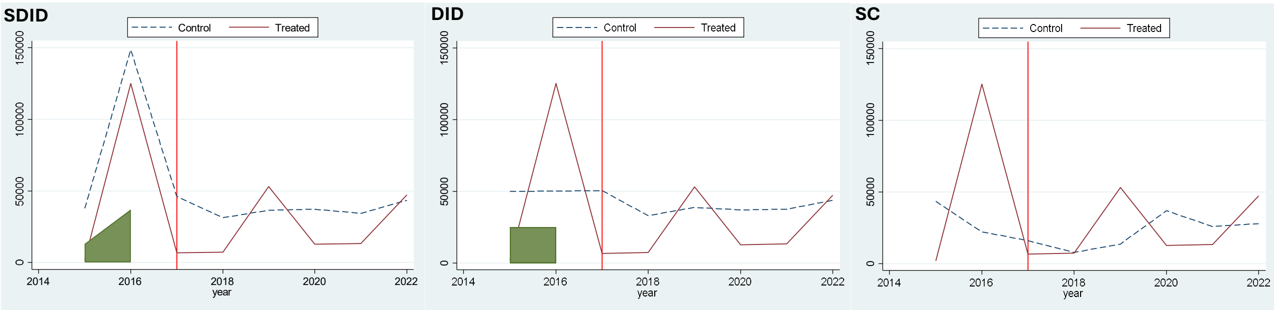


Figure 10 – Comparison of SDID, DID and SCM estimates for P5, entry year (cohort) 2016.

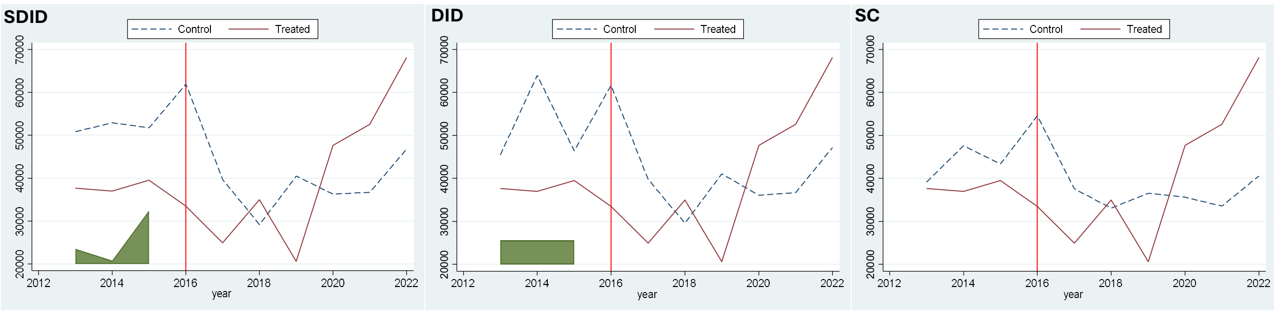
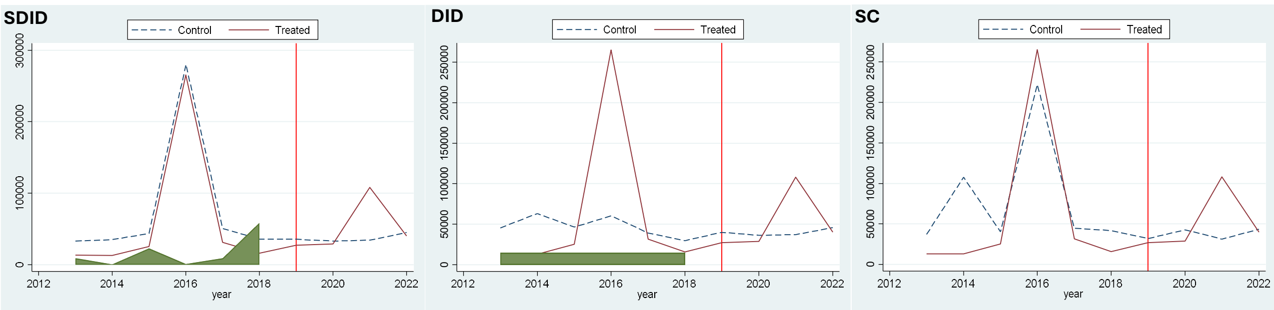


Figure 11 – Comparison of SDID, DID and SCM estimates for P6, entry year (cohort) 2019.



ANNEX 1

Table A1 – List of measures of the EU RDP 2014-2020/2022 and correspondence with priorities and focus areas

Code	Title
Priority	
Pr1	Fostering knowledge transfer and innovation in agriculture, forestry and rural areas
Pr2	Enhancing farm viability and competitiveness of all types of agriculture in all regions and promoting innovative farm technologies and the sustainable management of forests
Pr3	Promoting food chain organisation, including processing and marketing of agricultural products, animal welfare and risk management in agriculture
Pr4	Restoring, preserving and enhancing ecosystems related to agriculture and forestry
Pr5	Promoting resource efficiency and supporting the shift towards a low carbon and climate resilient economy in agriculture, food and forestry sectors
Pr6	Promoting social inclusion, poverty reduction and economic development in rural areas
Focus area	
1A	Fostering innovation, cooperation, and the development of the knowledge base in rural areas
1B	Strengthening the links between agriculture, food production and forestry and research and innovation, including for the purpose of improved environmental management and performance
1C	Fostering lifelong learning and vocational training in the agricultural and forestry sectors
2A	Improving the economic performance of all farms and facilitating farm restructuring and modernisation, notably with a view to increasing market participation and orientation as well as agricultural diversification
2B	Facilitating the entry of adequately skilled farmers into the agricultural sector and, in particular, generational renewal
3A	Improving competitiveness of primary producers by better integrating them into the agri-food chain through quality schemes, adding value to agricultural products, promotion in local markets and short supply circuits […] inter-branch organisations
3B	Supporting farm risk prevention and management
4A	Restoring, preserving and enhancing biodiversity, including in Natura 2000 areas, and in areas facing natural or other specific constraints and high nature value farming, as well as the state of European landscapes
4B	Improving water management, including fertiliser and pesticide management
4C	Preventing soil erosion and improving soil management
5A	Increasing efficiency in water use by agriculture
5B	Increasing efficiency in energy use in agriculture and food processing
5C	Facilitating the supply and use of renewable sources of energy, of by products, wastes, residues and other non food raw material for the purposes of the bio-economy
5D	Reducing green house gas and ammonia emissions from agriculture
5E	Fostering carbon conservation and sequestration in agriculture and forestry
6A	Facilitating diversification, creation and development of small enterprises, as well as job creation
6B	Fostering local development in rural areas
6C	Enhancing the accessibility, use and quality of information and communication technologies (ICT) in rural areas
Measures	
M01	M01 - Knowledge transfer and information actions (art 14)
M02	M02 - Advisory services, farm management and farm relief services (art 15)
M03	M03 - Quality schemes for agricultural products and foodstuffs (art 16)
M04	M04 - Investments in physical assets (art 17)
M05	M05 - Restoring agricultural production potential damaged by natural disasters and catastrophic events and introduction of appropriate prevention actions (art 18)
M06	M06 - Farm and business development (art 19)
M07	M07 - Basic services and village renewal in rural areas (art 20)
M08	M08 - Investments in forest area development and improvement of the viability of forests (art 21-26)
M09	M09 - Setting-up of producer groups and organisations (art 27)
M10	M10 - Agri-environment-climate (art 28)
M11	M11 - Organic farming (art 29)
M12	M12 - Natura 2000 and Water Framework Directive payments (art 30)
M13	M13 - Payments to areas facing natural or other specific constraints (art 31)
M14	M14 - Animal welfare (art 33)
M15	M15 - Forest environmental and climate services and forest conservation (art 34)
M16	M16 - Co-operation (art 35)
M17	M17 - Risk management (art 36)
M18	M18 - Financing of complementary national direct payments for Croatia (art 40)
M19	M19 - Support for LEADER local development (CLLD – community-led local development) (art 35 Regulation (EU) No 1303/2013)
M20	M20 - Technical assistance Member States (art 51-54)
M113	M113 - Early retirement
M131	M131 - Meeting standards based on Community legislation
M341	M341 - Skills acquisition, animation and implementation
Sub-measures	
M01.1	1.1 - support for vocational training and skills acquisition actions
M01.2	1.2 - support for demonstration activities and information actions
M01.3	1.3 - support for short-term farm and forest management exchange as well as farm and forest visits
M02.1	2.1 - support to help benefiting from the use of advisory services
M02.2	2.2 - support for the setting up of farm management, farm relief and farm advisory services as well as forestry advisory services
M02.3	2.3 - support for training of advisors
M03.1	3.1 - support for new participation in quality schemes
M03.2	3.2 - Support for information and promotion activities implemented by groups of producers in the internal market
M04.1	4.1 - support for investments in agricultural holdings

M04.2	4.2 - support for investments in processing/marketing and/or development of agricultural products
M04.3	4.3 - support for investments in infrastructure related to development, modernisation or adaptation of agriculture and forestry
M04.4	4.4 - support for non-productive investments linked to the achievement of agri-environment-climate objectives
M05.1	5.1 - support for investments in preventive actions aimed at reducing the consequences of probable natural disasters, adverse climatic events and catastrophic events
M05.2	5.2 - support for investments for the restoration of agricultural land and production potential damaged by natural disasters, adverse climatic events and catastrophic events
M06.1	6.1 - business start up aid for young farmers
M06.2	6.2 - business start up aid for non-agricultural activities in rural areas
M06.3	6.3 - business start up aid for the development of small farms
M06.4	6.4 - support for investments in creation and development of non-agricultural activities
M06.5	6.5 - payments for farmers eligible for the small farmers scheme who permanently transfer their holding to another farmer
M07.1	7.1 - support for drawing up and updating of plans for the development of municipalities and villages and of protection and management plans relating to N2000/HNV areas
M07.2	7.2 - support for investments in the creation/improvement of all types of small scale infrastructure, including investments in renewable energy and energy saving
M07.3	7.3 - support for broadband infrastructure and provision of access to broadband and public e-government
M07.4	7.4 - support for investments in the setting-up, improvement or expansion of local basic services for the rural population
M07.5	7.5 - support for investments for public use in recreational infrastructure, tourist information and small scale tourism infrastructure
M07.6	7.6 - support for studies/investments associated with the maintenance, restoration and upgrading of the cultural and natural heritage of villages, landscapes and HNV sites
M07.7	7.7 - support for investments targeting the relocation of activities and conversion of buildings , to improve the quality of life or increasing their environmental performance
M07.8	7.8 - others
M08.1	8.1 - support for afforestation/creation of woodland
M08.2	8.2 - support for establishment and maintenance of agro-forestry systems
M08.3	8.3 - support for prevention of damage to forests from forest fires and natural disasters and catastrophic events
M08.4	8.4 - support for restoration of damage to forests from forest fires and natural disasters and catastrophic events
M08.5	8.5 - support for investments improving the resilience and environmental value of forest ecosystems
M08.6	8.6 - support for investments in forestry technologies and in processing, mobilising and marketing of forest products
M09.1	9.1 - setting up of producer groups and organisations in the agriculture and forestry sectors
M10.1	10.1 - payment for agri-environment-climate commitments
M10.2	10.2 - support for conservation and sustainable use and development of genetic resources in agriculture
M11.1	11.1 - payment to convert to organic farming practices and methods
M11.2	11.2 - payment to maintain organic farming practices and methods
M12.1	12.1 - compensation payment for Natura 2000 agricultural areas
M12.2	12.2 - compensation payment for Natura 2000 forest areas
M12.3	12.3 - compensation payment for agricultural areas included in river basin management plans
M13.1	13.1 - compensation payment in mountain areas
M13.2	13.2 - compensation payment for other areas facing significant natural constraints
M13.3	13.3 - compensation payment to other areas affected by specific constraints
M14.1	14.1 - payment for animal welfare
M15.1	15.1 - payment for forest -environmental and climate commitments
M15.2	15.2 - support for the conservation and promotion of forest genetic resources
M16.0	16.0 - others
M16.1	16.1 - support for the establishment and operation of operational groups of the EIP for agricultural productivity and sustainability
M16.2	16.2 - support for pilot projects, and for the development of new products, practices, processes and technologies
M16.3	16.3 - co-operation among smalls operators in organising joint work processes and sharing facilities and resources, and for developing/marketing tourism
M16.4	16.4 - support for co-operation among supply chain actors for the establishment and development of short supply chains and local markets, and for promotion activities
M16.5	16.5 - support for joint action undertaken to mitigate or adapt to climate change, and for joint approaches to environmental projects and practices
M16.6	16.6 - support for cooperation among supply chain actors for sustainable provision of biomass for use in food and energy production and industrial processes
M16.7	16.7 - support for non-CLLD local development strategies
M16.8	16.8 - support for drawing up of forest management plans or equivalent instruments
M16.9	16.9 - support for diversification of farming activities into activities concerning health care, social integration, community-supported agriculture and education about environment/food
M17.1	17.1 - Crop, animal and plant insurance premium
M17.2	17.2 - Mutual funds for adverse climatic events, animal and plant diseases, pest infestations and environmental incidents
M17.3	17.3 - Income stabilisation tool
M18.1	18 - Financing of complementary national direct payments for Croatia
M19.1	19.1 - Preparatory support
M19.2	19.2 - Support for implementation of operations under the community-led local development strategy
M19.3	19.3 - Preparation and implementation of cooperation activities of the local action
M19.4	19.4 - Support for running costs and animation
M20.1	20.1 - support for technical assistance (other than NRN)
M20.2	20.2 - support for establishing and operating the NRN

Source: European Commission

ANNEX 2

Table A2 – Different sets of CMEF indicators

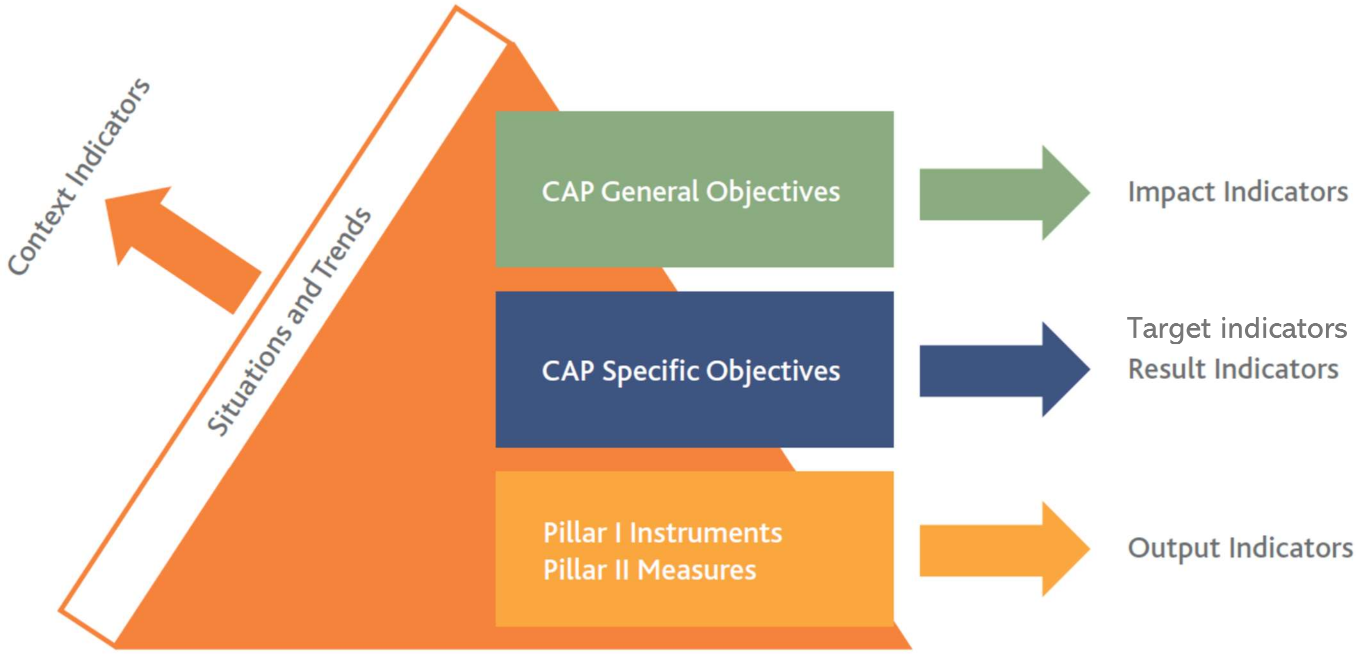
Code	Description
Context Indicators (CI)	
C1	Population
C2	Age Structure
C3	Territory
C4	Population Density
C5	Employment Rate
C6	Self-employment rate
C7	Unemployment rate
C8	GDP per capita
C9	Poverty rate
C10	Structure of the economy (GVA)
C11	Structure of Employment
C12	Labour productivity by economic sector
C13	Employment by economic activity
C14	Labour productivity in agriculture
C15	Labour productivity in forestry
C16	Labour productivity in the food industry
C17	Agricultural holdings (farms)
C18	Agricultural Area
C19	Agricultural area under organic Farming
C20	Irrigated Land
C21	Livestock units
C22	Farm labour force
C23	Age structure of farm managers
C24	Agricultural training of farm managers
C25	Agricultural factor income
C26	Agricultural Entrepreneurial Income
C27	Total factor productivity in agriculture
C28	Gross fixed capital formation in agriculture
C29	Forest and other wooded land (FOWL) (000)
C30	Tourism infrastructure
C31	Land Cover
C32	Areas with Natural Constraints
C33	Farming intensity
C34	Natura 2000 areas
C35	Farmland Birds index (FBI)
C36	Conservation status of agricultural habitats (grassland)
C37	HNV Farming
C38	Protected Forest
C39	Water Abstraction in Agriculture
C40	Water Quality
C41	Soil organic matter in arable land
C42	Soil Erosion by water
C43	Production of renewable Energy from agriculture and forestry
C44	Energy use in agriculture, forestry and food industry
C45	GHG emissions from agriculture
Output Indicators (TI)	
O1	Total public expenditure
O2	Total investment
O3	Number of actions/operations supported
O4	Number of holdings/beneficiaries supported
O5	Total area (ha)

O6	Physical area supported (ha)
O7	Number of contracts supported
O8	Number of Livestock Units supported (LU)
O9	Number of holdings participating in supported schemes
O10	Number of farmer benefiting from pay-outs
O11	Number of training days given
O12	Number of participants in trainings
O13	Number of beneficiaries advised
O14	Number of advisor trained
O15	Population benefiting of improved services/infrastructures
O16	Number of EIP groups and operations supported and number/type of partners in EIP groups
O17	Number of cooperation operations supported (other than EIP)
O18	Population covered by LAG
O19	Number of LAGs selected
O20	Number of LEADER projects supported
O21	Number of cooperation project supported
O22	Number and type of project promoters
O23	Unique number of LAG involved in cooperation project
O24	Number of thematic and analytical exchanges set up with the support of NRN
O25	Number of NRN communication tools
O26	Number of ENRD activities in which the NRN has participated
Result Indicators (TI)	
R1	Percentage of agricultural holdings with RDP support for investments in restructuring or modernisation (focus area 2A)
R2	Change in Agricultural output on supported farms/AWU (Annual Work Unit) (focus area 2A)*
R3	Percentage of agricultural holdings with RDP supported business development plan/investments for young farmers (focus area 2B)
R4	Percentage of agricultural holdings receiving support for participating in quality schemes, local markets and short supply circuits, and producer groups/organisations (focus area 3A)
R5	Percentage of farms participating in risk management schemes (focus area 3B)
R6	Percentage of forest or other wooded areas under management contracts supporting biodiversity (focus area 4A)
R7	Percentage of agricultural land under management contracts supporting biodiversity and/or landscapes (focus area 4A)
R8	Percentage of agricultural land under management contracts to improve water management (focus area 4B)
R9	Percentage of forestry land under management contracts to improve water management (focus area 4B)
R10	Percentage of agricultural land under management contracts to improve soil management and/or prevent soil erosion (focus area 4C)
R11	Percentage of forestry land under management contracts to improve soil management and/or prevent soil erosion (focus area 4C)
R12	Percentage of irrigated land switching to more efficient irrigation systems (focus area 5A)
R13	Increase in efficiency of water use in agriculture in RDP supported projects (focus area 5A)*
R14	Increase in efficiency of energy use in agriculture and food-processing in RDP supported projects (focus area 5B)*
R15	Renewable energy produced from supported projects (focus area 5C)*
R16	Percentage of LU (Live-stock Unit) concerned by investments in live-stock management in view of reducing GHG (Green House Gas) and/or ammonia emissions (focus area 5D)
R17	Percentage of agricultural land under management contracts targeting reduction of GHG and/or ammonia emissions (focus area 5D)
R18	Reduced emissions of methane and nitrous oxide (focus area 5D)*
R19	Reduced ammonia emissions (focus area 5D)*
R20	Percentage of agricultural and forest land under management contracts contributing to carbon sequestration or conservation (focus area 5E)
R21	Jobs created in supported projects (focus area 6A)
R22	Percentage of rural population covered by local development strategies (focus area 6B)
R23	Percentage of rural population benefiting from improved services / infrastructures (focus area 6B)
R24	Jobs created in supported projects (Leader) (focus area 6B)
R25	Percentage of rural population benefiting from new or improved services / infrastructures (Information and Communication Technology - ICT) (focus area 6C)
Target Indicators (TI)	
T1	Percentage of expenditure under Articles 14, 15 and 35 of Regulation (EU) No 1305/2013 in relation to the total expenditure for the RDP (focus area 1A)
T2	Total number of cooperation operations supported under the cooperation measure (Article 35 of Regulation (EU) No 1305/2013) (groups, networks/clusters, pilot projects…) (focus area 1B)
T3	Total number of participants trained under Article 14 of Regulation (EU) No 1305/2013 (focus area 1C)
T4	Percentage of agricultural holdings with RDP support for investments in restructuring or modernisation (focus area 2A)
T5	Percentage of agricultural holdings with RDP supported business development plan/investments for young farmers (focus area 2B)
T6	Percentage of agricultural holdings receiving support for participating in quality schemes, local markets and short supply circuits, and producer groups/organisations (focus area 3A)
T7	Percentage of farms participating in risk management schemes (focus area 3B)
T8	Percentage of forest/other wooded area under management contracts supporting biodiversity (focus area 4A)
T9	Percentage of agricultural land under management contracts supporting biodiversity and/or landscapes (focus area 4A)
T10	Percentage of agricultural land under management contracts to improve water management (focus area 4B)
T11	Percentage of forestry land under management contracts to improve water management (focus area 4B)

T12	Percentage of agricultural land under management contracts to improve soil management and/or prevent soil erosion (focus area 4C)
T16	Total investment in renewable energy production (focus area 5C)
T13	Percentage of forestry land under management contracts to improve soil management and/or prevent soil erosion (focus area 4C)
T14	Percentage of irrigated land switching to more efficient irrigation system (focus area 5A)
T15	Total investment for energy efficiency (focus area 5B)
T17	Percentage of LU concerned by investments in live-stock management in view of reducing GHG and/or ammonia emissions (focus area 5D)
T18	Percentage of agricultural land under management contracts targeting reduction of GHG and/or ammonia emissions (focus area 5D)
T19	Percentage of agricultural and forest land under management contracts contributing to carbon sequestration and conservation (focus area 5E)
T20	Jobs created in supported projects (focus area 6A)
T21	Percentage of rural population covered by local development strategies (focus area 6B)
T22	Percentage of rural population benefiting from improved services/infrastructures (focus area 6B)
T23	Jobs created in supported projects (Leader) (focus area 6B)
T24	Percentage of rural population benefiting from new or improved services/infrastructures (ICT) (focus area 6C)
Impact Indicators (II)	
I01	Agricultural entrepreneurial income
I02	Agricultural factor income
I03	Total factor productivity in agriculture
I04	EU commodity price variability
I05	Consumer price evolution of food products
I06	Agricultural trade balance
I07	Emissions from agriculture
I08	Farmland bird index
I09	High nature value (HNV) farming
I10	Water abstraction in agriculture
I11	Water quality
I12	Soil organic carbon in arable land
I13	Soil erosion by water
I14	Rural employment rate
I15	Degree of rural poverty
I16	Rural GDP per capita

Source: European Commission

Figure A1 – Articulation of the different sets of CMEF indicators with reference to the CAP objectives



Source: Adaptation from European Commission (2019).

ANNEX 3

Table A3 – SDID Estimates of confounders' coefficients (**β** in equations (7) and (8)) for both outcome variables under policy P4 (and the two respective cohorts).

Covariates (**X**)	Cohort 2017		Cohort 2020	
	Private ATT ($\Delta\pi$) ($.000€$)	Societal ATT (ΔG)[a]	Private ATT ($\Delta\pi$) ($.000€$)	Societal ATT (ΔG)[a]
SIZE1	-3733.4	1299.7	-4650.9	-7755.2
SIZE2	-3962.5	-2313.8	-7455.2	-333.4
SIZE3	2024.7	695.8	-14734.5	-41211.4
SIZE4	4846.3	1607.6	-12533.3	-5332.8
SIZE5	1829.1	-1792.6	-38987.1	-406.1
SIZE6	24587.8	36834.1	8718.8	5058.4
PRO1	743.1	581.2	-14008.7	5422.1
PRO2	-787.7	-229.3	5281.4	4785.1
PRO3	-1086.8	-211.1	2544.7	7447.6
PRO4	1431.2	61.2	6582.5	5959.5
FOR	286.2	-414.7	11574.6	-787.1
KW	556.4	10219.1	-72794.9	-4703.5
UAA	2699.7	1132.6	-1783.0	-1027.7
LSU	7776.3	17650.5	584991.6	13746.8
AWU	12845.0	14644.3	21047.0	4261.0
ENE	1200.3	1911.3	276.4	131.8
MAT	1364.9	1569.4	-3323.8	18535.6
AGE	-351.3	-660.1	-1479.2	6466.9
GEN	-4558.8	109.3	-9744.7	-4369.7
ALT	-715.8	23.69	2104.5	-58.48
LAT	-3.645	4.355	4.120	3.058
LON	-0.186	-0.174	0.487	-0.289
PIL1	4762.41	1904.3	1402.2	8849.1

[a]Societal outcome: OGA/AWU