EUROPEAN EVALUATION HELPDESK FOR RURAL DEVELOPMENT

European Commission

WORKING DOCUMENT

BEST USE OF FADN FOR THE ASSESSMENT OF RDP EFFECTS ON FOSTERING THE COMPETITIVENESS IN AGRICULTURE

WORKING PACKAGE 3

THEMATIC WORKING GROUP NO 8

'EX POST EVALUATION OF RDPS 2014-2020: LEARNING FROM PRACTICE'

JANUARY 2021

Disclaimer: *This document has been prepared by evaluation experts based on good practice available from the current programming period. This document is non-binding and only intended to facilitate the work of evaluators and managing authorities in the context of preparing for the ex post evaluation of the RDPs 2014-2020.*

Recommended citation:

EUROPEAN COMMISSION – Directorate-General for Agriculture and Rural Development – Unit C.4 (2021): Working Document 'Best use of FADN for the assessment of RDP effects on fostering the competitiveness in agriculture – Working Package 3 'Assessment of RDP effects on fostering the competitiveness of agriculture' - Thematic Working Group no 8 'Ex post evaluation of RDPs 2014-2020: Learning from practice'. Brussels.

The Evaluation Helpdesk is responsible for the evaluation function within the European Network for Rural Development (ENRD) by providing guidance on the evaluation of RDPs and policies falling under the remit and guidance of DG AGRI's Unit C.4 'Monitoring and Evaluation' of the European Commission (EC). In order to improve the evaluation of EU rural development policy the Evaluation Helpdesk supports all evaluation stakeholders, in particular DG AGRI, national authorities, RDP managing authorities and evaluators, through the development and dissemination of appropriate methodologies and tools; the collection and exchange of good practices; capacity building, and communicating with network members on evaluation related topics.

Additional information about the activities of European Evaluation Helpdesk for Rural Development is available on the Internet through the Europa server (http://enrd.ec.europa.eu).

CONTENT

LIST OF ACRONYMS

ATE	Average Treatment Effect
ATNT	Average Treatment Effects on the Non-Treated
ATT	Average Treatment Effect on the Treated
AWU	Annual Work Unit
CAP	Common Agricultural Policy
CEQ	Common Evaluation Question
CGE	Computable General Equilibrium model
CRI	Complementary Result Indicator
DID	Difference in Differences
EC	European Commission
ECA	European Court of Auditors
ESU	Economic Size
EU	European Union
EUR	Euro
FA	Focus Area
FADN	Farm Accountancy Data Network
FSS	Farm Structure Survey
GPSM	Generalized Propensity Score Matching
GVA	Gross Value Added
LAU	Local Administrative Unit
LSU	Livestock Unit
MS	Member State
NUTS	Nomenclature of Territorial Units for Statistics
PATT	Population Average Treatment on the Treated
PSM	Propensity Score Matching
RD	Rural Development
RDP	Rural Development Programme
SATT	Sample Average Treatment on the Treated
SO	Standard Outputs
TF	Type of Farming
TWG	Thematic Working Group
UAA	Utilized Agricultural Area

INTRODUCTION

This working document is one of the outcomes of the Working Package 3 **'Assessment of RDP effects on fostering the competitiveness in agriculture'** which analyses selected evaluation related issues when using the Farm Accountancy Data Network (FADN) data for the assessment of farm competitiveness in agriculture.

FADN has become a primary source of information for the assessment of farms' economic performance throughout the Member States. Ensuring the representativeness of the FADN data is one of the essential challenges stated by several Member States in their Annual Implementation Reports (AIRs), yearly capacity building events and good practice workshops.

This document discusses these issues and proposes practical solutions. This document provides examples from various Rural Development Programmes (RDPs) and describes what should be considered when using FADN data for the assessment of RDP effects on fostering competitiveness in agriculture and when answering the related common evaluation questions. This working document complements the existing guidelines, in particular:

- Assessment of RDP results: How to prepare for reporting on evaluation in 2017. Annex 11. Fiches for answering Common Evaluation Questions 4 and 6 for RDPs 2014-2020
- Assessing RDP achievements and impacts in 2019:
    - Part II 'Approaches for assessing RDP impacts in 2019. Chapter 2.2 'Sector related impact indicators'.
    - Part III 'Fiche for answering the Common Evaluation Question 27'.
    - Part IV 'Technical annex'. Chapter 4.1 'CAP common impact indicators related to the agricultural sector: I.01, I.02, I.03'.

This is **a non-binding document**, which aims to facilitate the exchange and learning from current practices in view to improve the quality of evaluations when preparing for the **ex post evaluation of RDPs 2014-2020**.

The drafting of this document has been carried out by evaluation experts in the context of the Evaluation Helpdesk's Thematic Working Group (TWG) on the 'Ex post evaluation of RDPs 2014-2020: Learning from practice'.

This document has been developed by Jerzy Michalek, Marili Parissaki, Valdis Kudiņš and the Evaluation Helpdesk members Julia Gallardo Gomez, Ana Prieto, Myles Stiffler, Hannes Wimmer.

Background

The assessment of the impact of the RDP is driven primarily by data availability, the type of the RDP measure analysed, the geographical coverage and the econometric approach used.

For instance, evaluations usually use survey (micro) data to estimate the RDP's effects. Individual micro-data needs to accomplish five main requirements when applied in order to estimate the RDP's impacts:

- homogenous data across assessed Member States and regions;
- coverage of policy relevant output indicators;
- disaggregated information about RDP measures and focus areas;
- coverage of the RDP implementation period;
- availability of sufficient data and balanced number of beneficiaries and non-beneficiaries.

These requirements can be theoretically satisfied in the case of the farming sector, as the EU wide data collection system FADN has been established. There are, however, some challenges associated with the representativeness and availability of data:

- non-representativeness of FADN in **comparison to the overall structure of the agricultural sector** in a given country or region (incl. no data on non-commercial and small farms);
- non-representativeness of FADN **regarding the structure of the supported farm through a given RDP.**

These challenges are further discussed and addressed in this working document by providing answers to the following guiding questions of the thematic working group:

- What common evaluation questions are linked to the assessment of RDP effects on fostering farm competitiveness in agriculture and what does it imply for data used in evaluation?
- Why is farm-level data essential to answering Common Evaluation Questions 4 and 6, which are linked to farm competitiveness?
- What requirements are needed from a sample of data at farm-level to be used for answering the Common Evaluation Question 4?
- What are the basic sources of farm-level data, which can enable evaluators to answer Common Evaluation Question 4?
- Are the variables available in the FADN sufficient to estimate the RDP's effects and to answer Common Evaluation Questions 4 and 6?
- Is the information and data available in the FADN on small farms sufficient to answer Common Evaluation Questions 4 and 6 and if not, how can one collect it?
- Given that the FADN is the first choice as a source of data for the calculation of important policy parameters (e.g., Average Treatment on the Treated (ATT)) needed for answering Common Evaluation Questions 4 and 6, how can the FADN be best utilised in order to mitigate any representation bias?
- Given that the FADN is the first choice as a data source for the calculation of important policy parameters (e.g., ATT) needed for answering Common Evaluation Questions 4 and 6, how can the FADN be best utilised in order to answer Common Evaluation Question 27?

Definitions used in this document

Programme effects can be calculated using the following effects:

- **Average Treatment Effect (ATE)** – measuring the effect of a programme on the population of the programme through assessing its beneficiaries and non-beneficiaries (randomly selected).
- **Average Treatment on the Treated (ATT)** – measuring the effect of a programme through assessing only the programme's beneficiaries.
- **Average Treatment on the Non-treated (ATNT)** – measuring the effect of a programme through assessing only the programme's non-beneficiaries.

In cases when the estimation of ATT is based on a sample of data, the ATT is understood as the **Sample Average Treatment on the Treated (SATT)**.

In cases when the ATT is estimated using the data of the whole population of beneficiaries, the ATT is called the **Population Average Treatment on the Treated (PATT)**.

When the estimation of the ATT is based on a representative sample of data the **SATT is considered equal to the PATT.**

1. COMMON EVALUATION QUESTIONS TO ASSESS RDP EFFECTS ON FOSTERING FARM COMPETITIVENESS IN AGRICULTURE

GUIDING QUESTION: What common evaluation questions are linked to the assessment of RDP effects on fostering farm competitiveness in agriculture and what does it imply for data used in evaluation?

In principle there are three common evaluation questions (CEQs) which are **directly** linked to the assessment of RDP effects on the **competitiveness** of farms in the agricultural/farming sector:

- CEQ 4 (Focus Area (FA) 2A): 'To what extent have RDP interventions contributed to improving the economic performance, restructuring and modernisation of supported farms in particular through increasing their market participation and agricultural diversification?'
- CEQ 6 (Focus Area 3A): 'To what extent have RDP interventions contributed to improving the competitiveness of supported primary producers by better integrating them into the agri-food chain through quality schemes, adding value to the agricultural products, promoting local markets and short supply circuits, producer groups and inter-branch organization?'
- CEQ 27 (CAP objectives): 'To what extent has the RDP contributed to the CAP objective of fostering the competitiveness of agriculture?'

Each of these CEQs have a **different scope** and are focused on **different target groups/sectors**[1]:

- CEQ 4 requires only the estimation of RDP effects on the population of farms which received direct support from a given RDP under FA 2A.
- Eligible target groups under FA 3A are various agri-food chain participants (e.g., agricultural producers, food processors, local market promoters), as well as their organisations and associations involved in improving agricultural products and food quality, promoting local markets and short supply circuits, producer groups and inter-branch organisations. However, CEQ 6 which relates to competitiveness only requires the estimation of RDP effects for primary producers (i.e., agricultural holdings who received a direct support from a given RDP under FA 3A(i)).
- CEQ 27 requires the estimation of all RDP effects on the competitiveness of the whole agricultural sector(ii) (i.e., explicitly targeting both RDP direct beneficiaries (supported under all focus areas 1A-6C) as well as non-beneficiaries that may have been indirectly affected by the RDP support). To answer CEQ 27 evaluators *inter alia* have to analyse interactions between RDP beneficiaries and non-beneficiaries, including various general equilibrium effects (e.g., spill-over effects, displacement effects).

Further considerations

(i) As common result indicators (linked to FA 3A) proposed by the EC are not sufficient to answer CEQ 6, competitiveness of farms supported under FA 3A has to be measured using additional indicators and other similar indicators as those suggested for farms supported under FA 2A.

(ii) Although RDP can affect many sectors in economy, e.g., the agricultural sector, the food processing sector, forestry sector, other sectors located in rural areas, etc. under CEQ 27 evaluator is asked to focus on the effect of RDP on the change of competitiveness of the agricultural sector only.

To answer these evaluation questions the evaluator will have to use a variety of **different** methodologies, representative data samples and extrapolation methods:

[1] There is an overlap of certain target groups (i.e., agricultural holdings) supported under FA 2A and FA 3A.

- Answering CEQ 4 and CEQ 6 requires using a sample of data which should be **representative** for the population of all **direct programme beneficiaries** of a given RDP only (i.e., agricultural producers supported under FA 2A and FA 3A) in order to estimate the effects.
- In contrast, answering of CEQ 27 requires carrying out the analysis in two stages:
    - Answering of CEQ 4 and CEQ 6 concerning the effects on agricultural holdings directly supported by the RDP under FA 2A and FA 3A.
    - Using the information from (i) to estimate the **aggregated effects** of a concerned RDP (or all RDPs implemented in the given country) on the competitiveness of the **whole agricultural sector.** This means the evaluator must estimate the effect of the RDP on **all the farmers (or types of farms)** in a given programming area (or country).

It is important to note that **CEQ 27 cannot be answered in a robust manner without first answering CEQ 4 and CEQ 6** (for details see: Section 3 'Specific considerations for using FADN'). Moreover, data used to answer **CEQ 27** will have **different properties** in comparison to the data used for answering **CEQ 4 and CEQ 6**:

- While data samples used to answer CEQ 4 and CEQ 6 should be **representative for farms supported** by a given RDP (under FA 2A and FA 3A), it does **not** have to be representative for all farms or the **whole agricultural sector.**
- The data sample used to answer CEQ 27 should, however, be **representative for the whole agricultural sector** (either in a given programming area or the whole country).

2. USE OF FARM-LEVEL DATA

GUIDING QUESTION: Why is farm-level data essential to answering Common Evaluation Questions 4 and 6, which are linked to farm competitiveness?

Common Evaluation Question 4 (Focus Area 2A)

> *'To what extent have RDP interventions contributed to improving the economic performance, restructuring and modernisation of supported farms in particular through increasing their market participation and agricultural diversification?'*

CEQ 4 asks evaluators to assess the effects of a given RDP on supported farms only. The estimated effects include all effects from all measures (including their primary and secondary contributions) implemented under the FA 2A.

For impact evaluations, the 'effect of the RDP' in a narrow interpretation can be understood as the 'net-effect'. In this understanding, the 'gross effect' (derived from an application of a 'before-after approach') would NOT be appropriate for answering CEQ 4 because it would also include the effects of other/non-programme related or exogenous factors (e.g., changes in prices, effects of other EU or national policies/programmes).

Treatment Effect

The **Average Treatment Effect on the Treated (ATT)** is the relevant treatment effect, which will enable the answering of the CEQ 4 (i.e., evidence-based estimation of the net effect of a given RDP on supported farms). While ATT focuses on the effect of the RDP on the programme's participants, it also describes the direct gain accrued to the economy from the existence of the programme compared a situation in which the programme did not exist[2]. The information from the ATT combined with information on the programme's indirect gains (e.g., general equilibrium effects) and programme's costs will contribute to answering various policy questions regarding the total net gain from the programme and facilitate the answering of CEQ 27[(iii)].

Data requirements

In order to answer CEQ 4 using the ATT the evaluator needs **farm-level data** (micro level data) describing the economic development of supported and non-supported farms (non-supported farms should be as similar as possible to supported farms in terms of size, characteristic, etc.). Finding an alternative to this approach is often NOT easily done.

Further considerations

[(iii)] While estimation of ATT is necessary to answer CEQ 4, CEQ 6 and others, some related evaluation questions can facilitate a better understanding of obtained results especially when answering CEQ 27 or questions linked to RDP effectiveness and efficiency. Among these questions the most important are:

- What could have been the effect of a given RDP on non-supported farms (e.g., those of the same type X as supported farms) had they participated in this programme? The relevant treatment effect is the Average Treatment Effects on the Non-Treated (ATNT).
- What was an average effect of the given RDP on the whole population of supported and non-supported farms? The relevant treatment effect is the Average Treatment Effect (ATE). ATE is the effect of assigning participation randomly to every unit (ignoring programme general equilibrium effects) and describes an expected gain from participating in the RDP for a randomly selected farm/individual from the joined sub-groups of programme participants and non-participants in a given programme area. This treatment effect averages the effect of the programme over all units in the population, including both programme participants and non-participants.

However, specific issues may arise when using ATE to policy analysis because the ATE includes the effect on units/farms/individuals for which the programme was never intended/designed. It may include the impact on units that may even be programme ineligible.

[2] Heckman and Robb, 1985; Heckman, 1997; Smith, 2000; Smith and Todd, 2003

Key points to keep in mind for alternative approaches

If **data is not available for non-supported farms**, some evaluators may try to answer CEQ 4 by comparing/contrasting data on supported farms with contextual or general trend data (e.g., provided for EU Member States and selected regions by Eurostat). Such an approach has limited applicability because the contextual or trend data includes both supported and non-supported farms. By comparing data of supported farms with contextual data, one is essentially comparing the performance of supported farms with aggregate data which already includes these supported farms in it, in essence comparing them to a large extent to themselves.

When **all farms have been supported by a given RDP**, but with different levels of intensity an evaluator is advised to apply a counterfactual based approach using the appropriate methods (e.g., Generalized Propensity Score Matching) and use farm-level data inter alia describing the structure and performance of the supported farms both before implementation of a given RDP and after it.

The only exception to the above rule which advocates answering CEQ 4 using farm-level data is a situation when there are considerable RDP **general equilibrium effects** between supported and non-supported farms, and there is **abundant cross-regional data** which covers important characteristics of highly disaggregated supported and non-supported regions. A description of such an approach is provided in Section 3 'Specific considerations for using FADN', Case 6 (below).

Common Evaluation Question 6 (Focus Area 3A):

'To what extent have RDP interventions contributed to improving the competitiveness of supported primary producers by better integrating them into the agri-food chain through quality schemes, adding value to the agricultural products, promoting local markets and short supply circuits, producer groups and inter-branch organization?'

Answering CEQ 6 similarly **requires farm-level data**. This data should be collected for agricultural producers supported under FA 3A (e.g., farms participating in quality schemes, or agricultural producers becoming a member in a producer organisation supported under FA 3A), as well as data from a similar group of producers who have not been supported through RDP funds (i.e., those producers not supported under FA 3A).

The same justification for using **farm-level data** as mentioned above for CEQ 4 (i.e., estimation of net impacts using ATT) is also **valid for answering CEQ 6**.

GUIDING QUESTION: What requirements are needed from a sample of data at farm-level to be used for answering the Common Evaluation Question 4?

Farm-level data on which an analysis can be based should cover all farms supported by a given RDP and a sufficient number of non-beneficiary farms. However, if these two populations are very large, or due to technical and/or economic constraints data is not available the evaluator will have to rely on sampling (i.e., the process of selecting a representative group from the larger population being studied).

What does representativeness of a sample mean in general?

If accurate conclusions about a population can be drawn from a sample, then it can be considered representative of the wider population. A sample is representative, if all the elements in a population have the same chance of being part of the sample. The best definition of a representative sample is a sample that has the same general characteristics as the target population.

What are the 'general characteristics' of the target population (i.e., population of the RDP supported farms) and why are they important?

General characteristics can have various interpretations depending on the context. For RDP supported farms, evaluators may be interested in the following criteria:

- the specialisation type of supported farm (e.g., crops, livestock, mixed),
- size of the supported farms (total utilised agricultural area (UAA) in hectares and economic size in EUR),
- number of persons employed,
- value of fixed assets,
- farm income,
- risk tolerance of the farm owner (e.g., his/her attitude towards risk (risk aversion))
- etc.

Some of those characteristics are observable (e.g., area of farm in ha, farm income, employment) while others are not (e.g., farm owner's attitude towards risk). While there are some methodological approaches which can attempt to tackle the 'unobservable', we will focus in this document on observable characteristics only.

Key points to keep in mind on representativeness and general characteristics

- ⇨ Since the individual effect of the RDP on a (randomly selected) farm is usually heterogeneous (it differs across farms) then one should expect that the individual effect of a given RDP will depend on the farm's general characteristics (as described above).
- ⇨ In order to reduce path dependence from previous RDP support (e.g. support provided in years 2007-2013 for both groups of farms, i.e. beneficiaries and non-beneficiaries) it is important to include in the list of farm characteristics (control variables) also those variables which show the amount of RDP support received by an individual farm from previous RDPs (e.g. those implemented in years 2007-2013). By using these control variables while matching beneficiaries with non-beneficiaries the potential effect of any previous RDP support can be neutralized.
- ⇨ When answering CEQ 4 it is important to make sure the data sample is representative for the population of the RDP supported farms under FA 2A. This does not necessarily mean it will be representative for the overall population of farms in a given region or country.
- ⇨ When answering CEQ 6 it is important to make sure the data sample is representative for the population of the RDP supported farms under FA3A. This does not necessarily mean it will be representative for the overall population of farms in a given region or country.

If the farm-level data sample is representative for all programme beneficiaries, one can expect:

- That the ATT effect calculated on the basis of this sample (Sample Average Treatment on the Treated (SATT)[iv] will be representative for the whole population of the supported farms in the given programme area.
- That the calculated SATT effect will depend on individual farm characteristics prevailing in the sample data (e.g., it will differ for small and large farms, as well as, for various types of farms (crop, livestock, mixed, etc.)).

Further considerations

[iv] As in most cases the estimation of ATT is based on a **sample** data, the ATT is normally understood as the Sample Average Treatment on the Treated (SATT). In rare cases when ATT is estimated using the whole population data, the ATT is called Population Average Treatment on the Treated (PATT). Only when estimation of the ATT is based on a **representative** sample data the SATT equals to PATT.

GUIDING QUESTION: What are the basic sources of farm-level data, which can enable evaluators to answer Common Evaluation Question 4?

In principle, there are three basic sources of farm-level (micro level) data which can be utilised when answering CEQ 4: **FADN, Farm bookkeeping data** and **specific surveys**.

Each of these sources have various advantages and disadvantages. Evaluators should carefully consider all sources of data before giving a preference to a specific one.

The FADN

The FADN is a survey carried out by the Member States of the European Union. It was established in 1965[3]. The FADN is the only source of micro-economic data that is harmonised (the bookkeeping principles are the same) among all EU Member States and is representative of the commercial agricultural holdings in the European Union.

Agricultural holdings are selected for the survey based on sampling plans established at each regional level of the EU. Currently, the FADN consists of approximately 3000 variables which are collected for roughly 80000 agricultural holdings in the EU 27 and are collected from accountancy data each year. The holdings included in the FADN represent a population of about 6400000 farms from 27 Member States, which cover approximately 90% of the total utilised agricultural area (UAA) and account for about 90% of the total agricultural production of the European Union.

While the FADN can provide useful harmonised information about the economic performance of commercial agricultural holdings per agricultural sector and size class, it has some specificities that might affect the calculation of RDP effects on the competitiveness of farms/sector when being used for evaluations:

- FADN is representative only for commercial farms (excludes non-commercial or small farms).
- Beneficiaries of CAP payments are in general underrepresented (e.g., the share of not represented beneficiaries of CAP direct payments varies from 12% in Bulgaria to 79% in Slovakia) see Table 1 (below)[4].

[3] Council Regulation (EEC) No 79/1965
[4] European Court of Auditors (ECA). (2016). Is the Commission's system for performance measurement in relation to farmers' incomes well designed and based on sound data? (Special Report No. 01). Luxembourg: European Court of Auditors.

- Certain sectors (particularly smaller ones) can be either under or overrepresented depending on the region or Member State (e.g., FADN under-represents sugar beet production by 18%, 99% and 235% compared with the total populations in Greece, Slovenia and Sweden, respectively. By contrast, in Italy, Slovakia and Spain, it overrepresents by 22%, 25% and 119% respectively)[5]. In Germany, very small and very large holdings are under-represented in the FADN sample.
- In the FADN, minimum thresholds of some important variables (e.g., standard output (SO) and therefore the coverage of the total farming population and beneficiaries of EU support) vary between Member States (e.g., minimum SO in Germany is 25000 EUR and in Bulgaria it is 2000 EUR). Some direct comparisons between Member States (e.g., for all holdings above 5000 EUR) are not possible (e.g., German holdings with SO between 5000-25000 EUR are not included in the FADN).
- The FADN sample design is achieved through a rotating panel technique that requires a certain portion of the sample to be periodically updated. This may lead to a discontinuity in the observations over a longer period.
- When the FADN sample is representative at the Member State level it does not necessarily mean it is at the regional level and can still be biased at this level.
- The FADN survey is conducted to meet the requirements of accounting records and therefore the information collected is not directly targeted for the needs of programme evaluation.

Although the FADN has become a primary source of information for the assessment of farms' economic performance among individual Member States it is important to note that the **FADN has not been designed to be representative of CAP beneficiaries nor specifically used as a monitoring tool for RDP measures**. In fact, CAP beneficiaries not covered in the FADN are mostly part-time or subsistence farmers, who fall below the respective survey thresholds and who receive only a small part of the total CAP budget[6]. Differences in the coverage of farming populations and CAP beneficiaries by FADN are depicted in Table 1.

National farm bookkeeping datasets

In some Member States (e.g., Germany, Austria, Slovakia, Poland) there is an abundance of farm bookkeeping data available for a large number of individual farm holdings. This dataset is commonly used by the national FADN Liaison Agencies to standardise the results and feed the FADN system. In most cases the number of farms included in the national survey is much larger than farms fed into the FADN[(v)] system. Farm bookkeeping data is collected yearly by professional national farm accounting organisations and/or associations, however, it is not harmonised across EU Member States. The latest update in 2020 present data for the year 2019.

Further considerations

[(v)] Before the creation of FADN, several Member States were already conducting agricultural surveys based on farm accounts. Some of these surveys were based on a selective sample of farms - as opposed to the entire population of farms. To select a sample of farms, these Member States had established their own selection plans. Most Liaison Agencies of the Member States continue to conduct national surveys and have thus retained their own selection plans (EC, 2010). For example, in Slovakia the number of farms included in the farm bookkeeping database (approximately 2500 farms in a sample) is more than 4 times bigger in comparison to FADN (approximately 600 farms in the sample).

[5] Louhichi, K., Espinosa, M., Ciaian, P., Perni, A., Vosough Ahmadi, B., Colen, L., & Gomez y Paloma, S. (2018). The EU-wide individual farm model for Common Agricultural Policy analysis (IFMCAP v.1: Economic Impacts of CAP Greening. EUR 28829 EN). Luxembourg: European Commission, Joint Research Centre, Publications Office of the European Union.

[6] European Court of Auditors (ECA). (2016). Is the Commission's system for performance measurement in relation to farmers' incomes well designed and based on sound data? (Special Report No. 01). Luxembourg: European Court of Auditors.

Farm Structure Surveys

The basic Farm Structure Survey (FSS) is carried out by all EU Member States. Farm Structure Surveys are conducted throughout the EU utilising a common methodology at regular harmonised intervals. This means that data from each FSS can be compared across Member States and time at regional levels (down to NUTS 3 level). Every 3 or 4 years the FSS is carried out as a sample survey and once in ten years as a census(vi). Like the FADN the FSS is also subject to minimum thresholds which determine if a given farm is included or not. Following the entry into force of Regulation (EC) No 1166/2008, the minimum requirements for survey coverage from the 2009/2010 FSS onwards have been modified. Countries which used a survey threshold above one hectare of UAA could fix this threshold at a level that excludes only the smallest agricultural holdings which together contribute 2% or less to the total UAA excluding common land, and 2% or less to the total number of farm livestock units (LSU).

Further considerations

(vi) For example, in Germany the Census of Agriculture 2020 has already started. In total 265000 agricultural holdings are asked to provide information on subjects such as farm succession, farm rent, animal husbandry and organic areas. The survey is conducted about every ten years by the statistical offices of the Federation and the Länder and provides comprehensive data for interested users, in particular from researchers, policy makers, the farming community and the agricultural industry. The results are presented at federal and Länder level, some also at district and municipality level. The Census of Agriculture is part of the Agricultural Census of the European Union.

In addition, countries have to include in the survey all population holdings which comply with at least one of the following set of physical thresholds[7]:

- Utilised agricultural area (arable land, kitchen gardens, permanent grassland, permanent crops): 5 hectares.
- Permanent outdoor crops (fruit, berry, citrus and olive plantations, vineyards and nurseries): 1 hectare.
- Other intensive production:
  - Fresh vegetables, melons and strawberries, which are outdoors or under low (not accessible) protective cover: 0.5 hectares.
  - Tobacco: 0.5 hectares.
  - Hops: 0.5 hectares.
  - Cotton: 0.5 hectares.
- Crops under glass or other (accessible) protective cover:
  - Fresh vegetables, melons and strawberries: 0.1 hectares.
  - Flowers and ornamental plants (excluding nurseries): 0.1 hectares.
- Bovine animals (all): 10 heads.
- Pigs (all): 50 heads.
- Breeding sows: 10 heads.
- Sheep (all): 20 heads.
- Goats (all): 20 heads.
- Poultry (all): 1000 heads.

However, only some necessary evidence for the assessment of RDP effects on the farming sector can be found in FSSs. The information collected in the FSS including covered land use, farming system, livestock numbers, management and farm labour input (including the age, gender and relationship to the holder of the agricultural holding), machinery and equipment, secondary activities, etc. Important information on RDP supported farms is still missing and the last available FSS is from 2016. Therefore, evaluators should look to the FADN for information on farm income and its generation instead.

[7] Defined in Annex II of Regulation (EC) No 1166/2008

Table 1. Coverage of farming populations covered by FADN and CAP beneficiaries covered by FADN

Member State	Number of holdings in Farm Structure Survey (FSS) 2010	National FADN threshold (in euro SO)	Percentage of FSS holdings represented in FADN	Percentage of CAP beneficiaries[8] NOT represented in FADN
Belgium	42 850	25 000	72 %	30 %
Bulgaria	370 500	2 000	31 %	12 %
Czech Republic	22 870	8 000	65 %	39 %
Denmark	42 120	15 000	68 %	51 %
Germany	299 150	25 000	65 %	47 %
Ireland	139 900	8 000	74 %	36 %
Greece	723 010	4 000	44 %	No information
Spain	989 810	4 000	53 %	61 %
Estonia	19 620	4 000	41 %	55 %
France	516 110	25 000[9] 15 000[10]	57 %	28 %
Italy	1 620 900	8 000	49 %	38 %
Cyprus	38 860	4 000	26 %	No information
Latvia	83 400	4 000	26 %	67 %
Lithuania	199 930	4 000	27 %	69 %
Luxembourg	2 210	25 000	73 %	20 %
Hungary	576 840	4 000	18 %	45 %
Malta	12 540	4 000	24 %	37 %
Netherlands	72 320	25 000	71 %	25 %
Austria	150 160	8 000	62 %	26 %
Poland	1 506 620	4 000	48 %	50 %
Portugal	305 260	4 000	36 %	53 %
Romania	3 859 030	2 000	27 %	No information
Slovenia	74 640	4 000	54 %	33 %
Slovakia	24 460	25 000	18 %	79 %
Finland	63 880	8 000	61 %	38 %
Sweden	71 100	15 000	39 %	57 %
United Kingdom	186 650	25 000[11] 15 000[12]	50 %	50 %
EU-27	12 014 740			

Source: European Court of Auditors (ECA). (2016). Is the Commission's system for performance measurement in relation to farmers' incomes well designed and based on sound data? (Special Report No. 01). Luxembourg: European Court of Auditors.

Key points to keep in mind concerning sources of farm-level data

- ⇨ Should evaluators have access to both FADN and farm bookkeeping data they should rather use farm bookkeeping data because these databases typically include more observations (i.e., the farm bookkeeping data sample is much larger than FADN) and are available with only a 1-year delay, whereas, the FADN data has a 2 year delay.
- ⇨ Given the specificities of the FADN data, it may occur that while calculating answers for CEQ 4 that in some Member States there could be an overestimation and in others an underestimation of the RDP policy impacts depending on the bias in the representation of specific types of farms in the total population of farms supported by the given RDP (under FA 2A). However, this bias will in most cases be independent of whether a FADN data sample is representative or not for the whole agricultural sector.

[8] Column 5 in Table 1 refers to CAP beneficiaries of direct payments.
[9] Continental France
[10] Guadeloupe, Martinique and La Réunion
[11] Except Northern Ireland
[12] Northern Ireland

3. SPECIFIC CONSIDERATIONS FOR USING FADN

GUIDING QUESTION: Are the variables available in the FADN sufficient to estimate the RDP's effects and to answer Common Evaluation Questions 4 and 6?

The FADN database is the most important data source for analysing the impact of agricultural policy instruments on the economic situation of farms. The methodology applied in the FADN aims to provide representative data at three dimensions: FADN region, economic size (ESU) and type of farming (TF).

The FADN database includes a large number of variables which are relevant for the estimation of RDP (net) effects on supported farms. The information is collected for each sample farm and described in a specific questionnaire called the 'Farm Return' which contains approximately 1000 variables which refer to:

- Physical and structural data (e.g., location, crop areas, livestock numbers, labour force).
- Economic and financial data (e.g., the value of production of the different crops, stocks, sales and purchases, production costs, assets, liabilities, production quotas).
- Subsidies including those connected to the RDP measures.

The standard results (known as SE variables) are a set of statistics calculated based on the 'Farm Returns' that are periodically produced and published by the European Commission. They describe in detail the economic situation of farmers by different groups and are usually sufficient to calculate the ATT effects for farms supported by a given RDP.

While physical and structural data as well as economic and financial data is available for longer time horizons, it should be relatively easy for evaluators to collect relevant data for both the period before a given RDP has started (e.g., year 2013) and after it (e.g., year 2021).

Furthermore, in most cases a sufficiently balanced micro-level panel dataset can be established. Balanced panel data, also known as cross-sectional time series data contains observations for each panel member (farm) for every year over a given time period[(vii)].

Due to the high data demands needed for a counterfactual approach, collected FADN economic data should include relevant information on both RDP beneficiaries and non-beneficiaries (separately for farms supported under FA 2A and FA 3A) concerning the farms' structure and performance and should also cover periods 'before' and 'after' the implementation of the programme. The largest part of data collected through the FADN (approx. 80%) is primarily related to the data block 'structure'. This data is essential for constructing meaningful control groups (e.g., through the application of matching techniques). Most of the data collected through the FADN which can be used for result and impact indicators fall under the data block 'performance' which should be collected both prior and after implementation of the RDP 2014-2020.

Further considerations

[(vii)] The most important advantages from using panel data are:

(i) Ability of more accurate inference of model parameters. Panel data usually contain more degrees of freedom and more sample variability than cross-sectional data.
(ii) Greater capacity for capturing the complexity of farm development and economic performance than a single cross-section or time series data, incl. better capacity for controlling impact of omitting variables.
(iii) Better ability to generate more accurate predictions for individual outcomes.
(iv) Panel data simplifies computation and inference analysis.

Therefore, in order to assess the effects of support provided to farms under FA 2A, which supports farm competitiveness, the evaluator must ensure that beneficiaries of FA 2A are compared with similar (in terms of farm structure and performance) agricultural holdings which have not received support from FA 2A. Similarities between these two groups of farms can be established through control variables,

which describe the structure and performance of these farms prior to the implementation of the RDP 2014-2020. When selecting relevant control variables (model covariates) only those FADN variables which are unaffected by the RDP programme should be included (i.e., variables which are fixed over time or which are measured prior to participation in the programme (e.g. farm economic size, employment, area, value of fixed assets))[13]. Furthermore, to measure the effects of FA 2A, which supports farms' competitiveness, various results and impacts indicators must be constructed. One important common result indicator linked to support under FA 2A is the complementary result indicator 2 (CRI 2) 'change in agricultural output on supported farms/Annual Work Unit (AWU)'. This indicator can be constructed using FADN variables SE131 (total output) divided by SE010 (total farm labour) which have been collected for a panel of farms (FA 2A beneficiaries and non-beneficiaries) both prior to and after implementation of the RDP 2014-2020. Other examples of common impact indicators which can be calculated from FADN variables to answer CEQ 27 are: I.01 (agricultural entrepreneurial income), I.02 (agricultural factor income) and I.03 (total factor productivity).

Though FADN data is usually sufficiently large and diverse to carry out a counterfactual analysis, some issues may arise when trying to validate which RDP has supported a given farm (applicable if the Member State has more than one RDP) and under which focus area a given farm received support.

Validating: Which RDP has supported a given farm

In the FADN database, the European Union is divided into FADN regions (e.g. 1 region in Czech Republic, 4 regions in Belgium, 4 regions in Poland, 16 regions in Germany, 17 regions in Spain, 21 regions in Italy, or 22 regions in France). In most cases, information concerning a specific FADN region (variable coded A1) enables setting up a direct link between a location of a farm and the RDP area. This is rather unproblematic for countries where the number of FADN regions is equal to or higher than the number of RDPs (e.g. Poland one national RDP, Germany 13 regional RDPs). However, in France the number of regional RDPs is larger than the FADN regions (i.e. 27 RDPs in programming period 2014-2020 compared with 22 FADN regions). If a direct link between the FADN region to a given RDP is not possible, evaluators should seek assistance from the Paying Agency to match farms included in the FADN with a given RDP.

Validating: Under which focus area a given farm received support

The rules applied to the implementation of RDPs changed in the programming period 2014-2020 in comparison to the previous period in that each individual rural development (RD) measure can now contribute to the implementation of more than one priority and focus area at the same time. For example, Measure M1 (Knowledge transfer and information actions) can be applied to the implementation of multiple RD priorities and focus areas (e.g., 1A, 1C, 2A, 2B), depending on the individual needs of the region. As a consequence, evaluators are now NOT expected to assess the effect of an individual RDP measure (as it was the case in the previous programming period), but instead they are asked to estimate the effects of a bundle of measures programmed and implemented under each specific focus area (and across all priorities). This change may create various issues if the evaluator only uses FADN data.

Some examples

On the basis of information on a type of subsidy received by a given farm available in the FADN dataset (e.g. Other RD payments (variable SE623 = JC830 'support to standards' + JC835 'support to advisory services' + JC840 'support for quality' + JC900 'subsidies for afforestation' + JC910 'subsidies other forestry' + JC953 'subsidies other rural development')) it is difficult to determine if a farm received this subsidy under the focus area 'x' or 'y'. Due to the fact that many subsidies can be programmed under various priorities and focus areas means that the knowledge of a positive entry for SE623 in the FADN database is usually not sufficient to place a given farm (RDP beneficiary) in the group of farms supported under a specific focus area (e.g. FA 2A).

[13] A comprehensive discussion of various criteria/methods to be considered while choosing control variables into the matching models (e.g., Hit or Miss method, Statistical Significance, Leave One out Cross Validation, etc.) can be found in: Caliendo and Kopeing (2005) http://ftp.iza.org/dp1588.pdf

As a solution to the above-mentioned issues evaluators should seek assistance from the Paying Agency regarding information on all RDP measures (programming period 2014-2 020) in which each farm included in the FADN participated (i.e., using administrative anonymised data).

An advantage of the current rules, which asks evaluators to assess the effect of RDP support provided under the respective focus area and not an individual RDP measure is that it leads to a higher number of observations in each relevant group (focus areas) of programme beneficiaries. This is because in most cases, more farms will benefit from support provided under a focus area than from an individual measure.

Some examples

Another example, FADN variable SE409 indicates a subsidy on agricultural investment which could indicate whether a given farm was supported under the FA 2A, however this variable is only available since 2017.

Key point to keep in mind when using the FADN variables

In order to utilise the FADN data for the assessment of RDP impacts on fostering the competitiveness of the farming sector the variables selected from the FADN database **should be supplemented with information from the Paying Agency concerning subsidies received by each farm from the RDP 2014-2020 as well as information on the focus area under which these subsidies were provided.** Linking the Paying Agency data with the FADN data can be done anonymously (e.g. using farm ID only) either in the FADN Liaison Agency or in the Paying Agency.

GUIDING QUESTION: Is the information and data available in the FADN on small farms sufficient to answer Common Evaluation Questions 4 and 6 and if not, how can one collect it?

As shown in Table 1, various EU Member States apply different thresholds (expressed in Standard Outputs (SO), a monetary equivalent of gross agricultural output at the farm gate prices) which qualify their farms to enter (or to stay outside) the FADN database. In Member States where farms are just above these thresholds they are considered to be economically 'small'. For example, a 'small farm' included in the FADN database in Germany, France (continental) or Luxembourg has a SO per year equal to a bit more than 25000 EUR, whereas in Poland, Portugal, or Slovenia the threshold is 4000 EUR and in Bulgaria or Romania only 2000 EUR. In each Member State a threshold is set to represent the situation of a country's agricultural market, especially with respect to commercial farms. The logical consequence of this narrative is that in each Member State the FADN database includes 'small' farms. Whether the same sample of 'small' farms included in FADN is representative for all small farms in the country may be question, however, the same issue of comparability of the economic performance of 'small' farms with SO (e.g., approx. 5000 EUR across the Member State (clearly Germany, France and Luxembourg and other countries with thresholds > 5000 EUR would NOT be included) remains. Nevertheless, when evaluating RDP effects (to be carried out at the programming area level) the fact that 'small' farms are included in the FADN database can be successfully utilised if there is enough observations for this group of farms (in order to increase the number of observations some clusters can be built (e.g. 'small and medium' farms vs. 'large' farms).

The problem may arise in a programme area or country that a large part of the RDP support is provided to farms whose SO is much lower than the respective FADN threshold (e.g. in Slovakia lower than 25000 EUR). In such a situation, it is quite clear that the FADN sample of farms is not sufficient to meaningfully estimate the population's ATTs and thus the answer provided for CEQ 4 would be biased.

In order to deal with such a situation, evaluators in their respective countries may revert to other sources of data (e.g., farm bookkeeping statistics, or surveys) which may include these 'very small' farms (supported and non-supported).

GUIDING QUESTION: Given that the FADN is the first choice as a source of data for the calculation of important policy parameters (e.g., ATT) needed for answering CEQ 4 and CEQ 6, how can the FADN be best utilised in order to mitigate any representation bias?

The answer to this question depends on the level of detail an evaluator is expected to provide when estimating the effects of the RDP on fostering competitiveness in farms supported under the Focus Area 2A.

Six possible cases based on different assumptions are presented and explained in detail below:

- **Case 1:** All farms from the FADN dataset which were identified as supported by a given RDP (under FA 2A or FA 3A) were randomly selected in the FADN.
- **Case 2:** Farms found in the FADN that are supported by the RDP under FA 2A or FA 3A are not representative[14] but there is a sufficient number of observations in the FADN database which allows for a separate estimation of Sample Average Treatment on the Treated (SATT) for the different groups of farms.
- **Case 3:** The same conditions as in Case 2, however, there is not enough observations in FADN to carry out a separate estimation for various farm types.
- **Case 4:** The number of farms supported under FA 2A or FA 3A found in the FADN database is higher than the number of non-supported farms. Additionally, the conditions of Case 2 or Case 3 apply (i.e., FADN sample of farms supported under FA 2A or FA 3A is not representative for the total population of supported farms (under FA 2A or FA 3A)).
- **Case 5:** Almost all farms found in the FADN database in a given programming region have been supported by the RDP under FA 2A or FA 3A. Additionally, Case 2 and Case 3 conditions apply (i.e., the FADN sample of farms supported under FA 2A or FA 3A is not representative for the total population of supported farms (under FA 2A or FA 3A)).
- **Case 6:** In the FADN database for a particular EU Member State there is not a sufficient number of observations enabling the estimation of the Sample Average Treatment on the Treated (SATT) effects.

CASE 1

All farms from the FADN dataset which were identified as supported by a given RDP (under FA 2A or FA 3A) were randomly selected in the FADN

The assumption in Case 1 is that all farms found in the FADN dataset which belong to a given RDP area (e.g., Andalusia (Spain), Bavaria (Germany), Slovakia (whole country)) and were supported under Focus Area 2A or FA 3A have been **randomly**(viii) selected to the FADN dataset. This assumption is quite reasonable since all farms in the FADN dataset have been selected using a **representative sampling** system with the objective to maintain the same general characteristics as **the country's target population of commercial farms** (by regions, farm size and farm typology) and **NOT** to specifically obtain a representative sampling of **farms supported by the** RDP under FA 2A or FA 3A.

Further considerations

(viii) A **random sample** is a group (or set) chosen in a **random** manner from a larger population. In a **simple random sample**, every member and set of members has an equal chance of being included in the sample. A simple random sample is statistically sufficient for a good estimate.

While **random sampling** helps to produce **representative samples**, we may conclude that the FADN sample of farms supported under FA 2A or FA 3A will only be subject to a self-selection bias. However,

[14] Farms in this case differ significantly in terms of their structure in comparison to the total population of supported farms (under FA 2A or FA 3A).

after reducing this bias (e.g., as a result of using a matching methodology) **the obtained ATT results will be representative for the whole population of farms supported in a given programming area under FA 2A or FA 3A.**

Under this assumption the answer to the CEQ 4 and CEQ 6 or the estimated total effect of the RDP on all supported farms under FA 2A or FA 3A is quite straightforward in that it is just the value of the ATT multiplied by the number of all farms supported under FA 2A or FA 3A in a given programming area.

Example illustrating case 1

In Slovakia two measures (3 sub-measures) were implemented under the Focus Area 2A in years 2014-2018. These were M04 with sub-measures M4.1, M4.3, and M06 with sub-measure M6.3. The overview of total finalised operations under FA 2A until the end of 2018, or the whole population of supported farms under FA 2A, is provided in Table 2.

Table 2. Slovakia: Summary of finalised operations under FA 2A in years 2014-2018

Measures	Paid support in EUR (2014-2018)	Number of supported holdings
M4.1	37 535 864	202
M4.3	1 815 705	7
M6.3	-	0
Total:	**39 351 569**	**209**

The estimated values of the ATTs from a FADN sample of **82 farms** supported under FA 2A are as follows:

- Change of agriculture output = 424407 EUR/farm
- Change of farm employment = 5.79 AWU/farm
- Change of labour productivity (R2) = **-3921** EUR/AWU

The amount of the change on R2 of the ATT is provided in Table 3.

Table 3. Slovakia: Estimated values of R2 for supported and non-supported farms

Variable	Sample	Treated	Controls	Difference
R2_diff	Unmatched	4750.54255	9904.0208	- 5153.47825
	ATT	4808.47348	8729.42858	- 3920.9551

Based on the results from Table 3 one can see that in period 2014-2018 labour productivity in farms supported by the RDP under FA 2A increased by +4808 EUR/AWU. This number was much **lower** compared to a similar group of farms which have not received any support under RDP FA 2A where it grew by +8729 EUR/AWU[15]. Based on the above calculation one can conclude that support received through the RDP under FA 2A resulted in a **decrease of labour productivity (R2)** in the group of RDP beneficiaries in average by -3921 EUR/AWU per farm.

By extrapolating these results to the group of **all farms which received support under FA 2A** one is able to obtain the following **aggregated primary direct effects**:

1. Change of agricultural output = 209 supported farms * 424407 EUR/farm = 88.7 Mill EUR
2. Change of employment = 209 supported farms * 5.79 AWU/farm = 1210 AWU
3. Change of labour productivity = 209 supported farms * 25 AWU/farm * -3921 EUR/AWU per farm = -20.5 Mill EUR

[15] Support received from other RDP measures was in both groups of farms the same – the level of subsidies received in both groups from other measures, e.g., M10, M11, M12, M13 and M14 were used as control variables.

CASE 2

Farms found in the FADN that are supported by the RDP under FA 2A or FA 3A are not representative, but there is a sufficient number of observations in the FADN database which allows for a separate estimation of SATT for the different groups of farms.

For Case 2, it is assumed that in the FADN database sample farms found as being supported by the RDP under FA 2A or FA 3A **differ significantly in terms of their structure** with the total population of supported farms under FA 2A or FA 3A (e.g. regarding % of small farms vs. % of large farms). As a consequence, the sample of farms supported under FA 2A or FA 3A in the FADN database **cannot** be treated as a **representative sample** for the total population of farms supported under FA 2A or FA 3A.

Under this assumption, in order to answer CEQ 4 and CEQ 6 and to estimate programme effects on all farms supported under FA 2A or FA 3A we cannot simply **estimate** a sample average of treatment on the treated (SATT) and then multiply it with the total number of supported farms because the estimated SATT will **NOT** be representative for the whole targeted group of the population (i.e. farms supported under FA 2A or FA 3A).

If there is a **difference in terms of share** of small and big farms in the sample in comparison to the total population of farms supported under FA 2A an evaluator can proceed as follows:

1. Based on the FADN dataset one can estimate SATT (1) for the group of **small and medium** farms only.
2. Based on the FADN dataset one can estimate SATT (2) for the group of **big farms** only.
3. Calculate **the aggregated effects for the total population of supported farms** as: {SATT(1) * % share of small and medium farms in total population + SATT(2) * % share of big farms in total population} * total number of farms supported under FA 2A.

Example illustrating case 2

Assuming that:

- The number of supported farms under FA 2A = 7100 (in total)
- **SATT(1)** for agriculture output in small farms = 10000 EUR/farm
- **SATT(2)** for agriculture output in big farms = 300000 EUR/farm
- % share of small/medium farms in the total population of farms supported under FA 2A = 78%
- % share of big farms in the total population of farms supported under FA 2A = 22%

In this case, the **aggregated effects** of the support under FA 2A on agricultural output (size weighted) for the total population of supported farms under FA 2A = 7100 * (10000 * 0.78 + 300000 * 0.22) = 7100 * 73800 EUR = 523.9 Mill EUR.

Calculations as described above in Case 2 can be carried out **without** taking into consideration corresponding shares of small/medium vs. big farms in the whole FADN sample because these shares are **irrelevant** if the estimation is done **separately** for both groups of farms. Also, it is **irrelevant** if the sample of all farms in the FADN database is **representative or not** for the whole population of farms in the given programming area.

However, such calculations can only be carried out if there is a **sufficient number of observations** in the FADN database allowing for a **separate** estimation of the SATT for different groups of farms (e.g. SATT(1) for small/medium and SATT(2) for big farms). However, if this is NOT the case and there are not a sufficient number of observations allowing for a separate estimation of RDP effects for different farm categories, we can proceed as in Case 3 (see below).

CASE 3

The same conditions as in Case 2, however, there is not enough observations in FADN to carry out a separate estimation for various farm types.

Case 3 makes two fundamental assumptions:

- That in the FADN database sample farms which are supported by the RDP under FA 2A or FA 3A are not representative, differing significantly in terms of their structure, in comparison with the total population of supported farms under FA 2A or FA 3A.
- That in the FADN sample there are not enough observations, a low number of farms supported under FA 2A or FA 3A that belong to a given programming area, to estimate econometrically the effects of the programme on the economic performance of these farms.

Under these two assumptions, in order to answer CEQ 4 or CEQ 6 and to estimate the programme's effects on all farms supported under FA 2A or FA 3A an evaluator may **increase the number of observations** by including in this sample those farms which were supported under FA 2A or FA 3A in other **FADN regions, within the same Member State** (e.g., Andalusia and Murcia (Spain), or Bavaria and Baden-Württemberg (Germany)). This can only be done when RDPs are implemented in regions, which **do not substantially differ** in their implementation details (e.g. regarding eligibility criteria) and agricultural holdings supported under FA 2A or FA 3A do not significantly differ in terms of farm characteristics between the 'original' and 'added' regions.

After extending the original dataset by including another 'similar' FADN region, additional variables (e.g. 0-1 dummies/shifters) can be built into the list of covariates enabling the evaluator to identify the location of a farm in the different RDPs or FADN regions, thus enabling an identification of the specific programme effects in those RDPs or region.

Example illustrating case 3

A construction of a **balanced panel** containing observations on the same farms over time for both programme beneficiaries and non-beneficiaries, in years prior to the RDP (e.g., 2006) and after its implementation (e.g., 2014) are a prerequisite for the calculation of the effects of the RDP on individual farms.

For example, in Bavaria, the balanced panel FADN data of beneficiaries and non-beneficiaries of the RDP by considering a FADN variable 'subsidies on investment' implemented in 2007-2013 consists of 1003 individual farms. The main characteristics of the available FADN panel data for the programme area of Bavaria (A1=90) concerning the number and distribution of programme beneficiaries and non-beneficiaries by a RD measure 'subsidies on investment' is shown in Table 4.

Table 4. Number of RDP beneficiaries vs. non-beneficiaries in Bavaria (DE90), FADN panel 2006-2014 (total: 1003 farms)

Type of RD subsidies	Beneficiaries	Non-beneficiaries
Subsidies on investment	62 (6%)	941 (94%)

The data shows that out of 1003 farms included in the FADN panel (years 2006-2014) only 62 farms (6%) in the same period of time received RD investment subsidies. If the composition of supported farms found in the FADN database **significantly differs in terms of a particular farm characteristic, for example, the farm type** (field crops, horticulture, other permanent crops, milk) from the composition of the total population of the supported farms in Bavaria, **and we expect that the effect of the programme depends strongly on the farm type**, a simple extrapolation of an estimated average effect (SATT) calculated on the basis of the sample of 62 farms on the total population of farms supported under this measure or FA 2A will lead to biased results. A solution would be to estimate programme effects **for specific farm-types**.

However, in this example, due to the insufficient number of observations, econometric estimation of programme effects **differentiated by farm type** (according to TF1-TF8 classification) at the level of the **programming area using FADN sample for Bavaria may not be feasible**. One possible solution could be a reduction of the number of considered farm types (TF) through clustering/aggregating them from TF8 to TF4 as shown in Table 5.

Table 5. FADN TF8 farming type classification and clustering in TF4

TF4 classification	TF8 classification	
TF1_4	TF1	Field crops
	TF2	Horticulture
	TF3	Wine
	TF4	Other permanent crops
TF2_4	TF5	Milk
TF3_4	TF6	Other grazing livestock
	TF7	Granivores
TF4_4	TF8	Mixed

By applying this solution, an evaluator can increase the number of observations of supported farms belonging to a given farm-type or cluster. In this example, the results are shown in Table 6.

Table 6. Number of RDP beneficiaries vs. non-beneficiaries in Bavaria (DE90), FADN panel 2006-2014, Grouped by 4 farm types

Type of RD subsidies	TF 4 classification	Beneficiaries	Non-beneficiaries
Subsidies on investment	TF1_4	12 (5%)	223 (95%)
	TF2_4	24 (8%)	264 (92%)
	TF3_4	14 (6%)	206 (94%)
	TF4_4	12 (5%)	235 (95%)

In an **optimal situation, in order to econometrically estimate a specific farm-type effect of the programme**, in **each farm type (TF)** the number of observations in the group of beneficiaries should be much higher than the number of covariates used in the Propensity Score Matching (PSM) analysis, and the number of observations in the group of non-beneficiaries should be at least two times higher[16] than the difference between the number of observations in the group of beneficiaries and the number of covariates used in the PSM analysis.

The data in Table 6, however, shows that even after aggregation of farm types from 8 to 4 categories estimation of farm type-specific programme effects would not be feasible. Clearly, whether this solution will be effective or not is an empirical question, so it may happen that in other programming areas after such aggregation the number of observations in each farm type cluster will already be sufficient to perform an econometric estimation.

If after aggregating various types of farms the econometric estimation of ATT is still not possible, the suggested approach would be to **increase the number of observations**, in total and for farms in different farm types, **by incorporation into the analysis those farms which are located in other similar regions or neighbourhood** regions of the programming area being analysed. In the case of Germany, we may incorporate into a sample of farms in Bavaria (A1=90) also those farms found in the FADN located in Baden-Württemberg (A1=80) which increases our sample by approximately 1200 additional observations. However, this can only be done when the RDPs implemented in neighbouring FADN regions do not substantially differ in their implementation details, eligibility criteria, as well as in terms of farms characteristics from those supported in the original region. After increasing the number of observations an econometric estimation of the ATT can be carried out after a **dummy variable** showing in which programme area a given farm is located has been included into the list of control variables.

Should the estimation of the RDP's effects for a specific farm type not be possible even after adding observations from neighbouring region, meaning there is still not enough observations for each type of farm, evaluators may decide to analyse in detail the effect of the programme support under FA 2A or FA 3A by a certain type of farm (e.g., field crops or milk) **only** and answer the CEQ 4 or CEQ 6 by considering this type of farm only. Such a choice will still be more effective than giving up the whole quantitative analysis.

[16] Based on other empirical studies which applied PSM-DID approach.

CASE 4

The number of farms supported under FA 2A or FA 3A found in the FADN database is higher than the number of non-supported farms. Additionally, the conditions of Case 2 or Case 3 apply.

For Case 4 it is assumed that the number of farms **supported** under FA 2A or FA 3A found in the FADN database is **higher** than the number of non-supported farms; and that the FADN sample of farms supported under FA 2A or FA 3A is **not representative** for the total population of supported farms under FA 2A or FA 3A.

Assumption: the number of farms supported under FA 2A or FA 3A found in the FADN database is higher than the number of non-supported farms

From a modelling point of view in a situation where the number of beneficiaries is higher than the number of possible controls, the application of a quasi-experimental approach (e.g., PSM methodology) is still possible if **replacement** is allowed. In nearest-neighbour matching with replacement the same unit can be used for more than one unit in the opposite status[17]. **Radius (or calliper)** matching is also applicable if the distance measured in terms of their individual propensity scores between units in the two groups is lower than a chosen tolerance limit[18]. Application of the PSM methodology when the group of supported units is much bigger than the group of non-supported units is also possible through **Kernel matching**. When Kernel matching is applied, every supported unit is matched with a weighted average of **all non-supported** units with weights that are inversely proportional to the distance between the supported and non-supported units (i.e., lowest distance → highest weight). From a **technical** point of view even with such an unfavourable data structure (a high ratio of beneficiaries in relation to possible controls) the **PSM-DID methodology is still feasible and can be applied.**

Assumption: the FADN sample of farms supported under FA 2A or FA 3A is not representative for the total population of supported farms under FA 2A or FA 3A

If the FADN sample of farms supported under FA 2A or FA 3A is not representative for the total population of farms supported under FA 2A or FA 3A, the evaluator may **apply procedures described in Cases 2 or 3** (see above).

CASE 5

Almost all farms found in the FADN database in a given programming region have been supported by the RDP under FA 2A or FA 3A. Additionally, Case 2 and Case 3 conditions apply.

Case 5 makes two fundamental assumptions:

- almost **all farms** found in the FADN database in a given programming region **have been supported by the RDP under FA2 A or FA 3A**; and
- the FADN sample of farms which have been supported under FA 2A or FA 3A are **not representative** for the total population of supported farms under FA 2A or FA 3A.

Under the above assumptions, in order to answer the CEQ 4 and CEQ 6 and estimate the programme's effects on all farms supported under FA 2A or FA 3A an evaluator may proceed with the following steps:

[17] When matching is done without replacement, the same unit can be used only once per each unit in the opposite status.
[18] Of course, a defining *a priori* a correct calliper may be difficult.

Step 1: Estimating the RDP effects based on the FADN sample at farm-level

The evaluator may employ the generalised propensity score matching (GPSM) method[19] using sample data available from the FADN. The GPSM approach can be applied in situations when all farms located in a given programming area received a direct support under the specific focus area (e.g., FA 2A or FA 3A) and the **intensity of this support per unit is known.** The GPSM also allows one to take into consideration farms which did not receive support (i.e., level of intensity =0).

While the FADN database allows an evaluator to select all necessary covariates needed to implement the GPSM method at farm level, it does not provide detailed information about the intensity of support received by a given farm for each individual measure and sub-measure. Furthermore, even if such information was available, it would not be sufficient to identify for each farm a focus area under which these measures were implemented because each individual RD measure can currently contribute to the implementation of more than one priority and focus area simultaneously (i.e., many different measures can be implemented under FA 2A or FA 3A). Under these circumstances, evaluators should **seek assistance from the Paying Agency** in order to link which focus area a given farm was supported by and at what intensity this support was allocated under FA 2A or FA 3A.

If the above information is received, then under the GPSM framework, programme effects on farm competitiveness can be analysed by means of a **dose-response function and derivative dose-response function**. The GPSM method not only allows one to estimate the **average effect** of support under FA 2A or FA 3A on the selected result/impact indicators (e.g., Gross Value Added (GVA) /farm or GVA/region) but also to assess the programme's **marginal effects** based on the support intensity level obtained.

Step 2: Extrapolation of the RDP effects at the programming area level

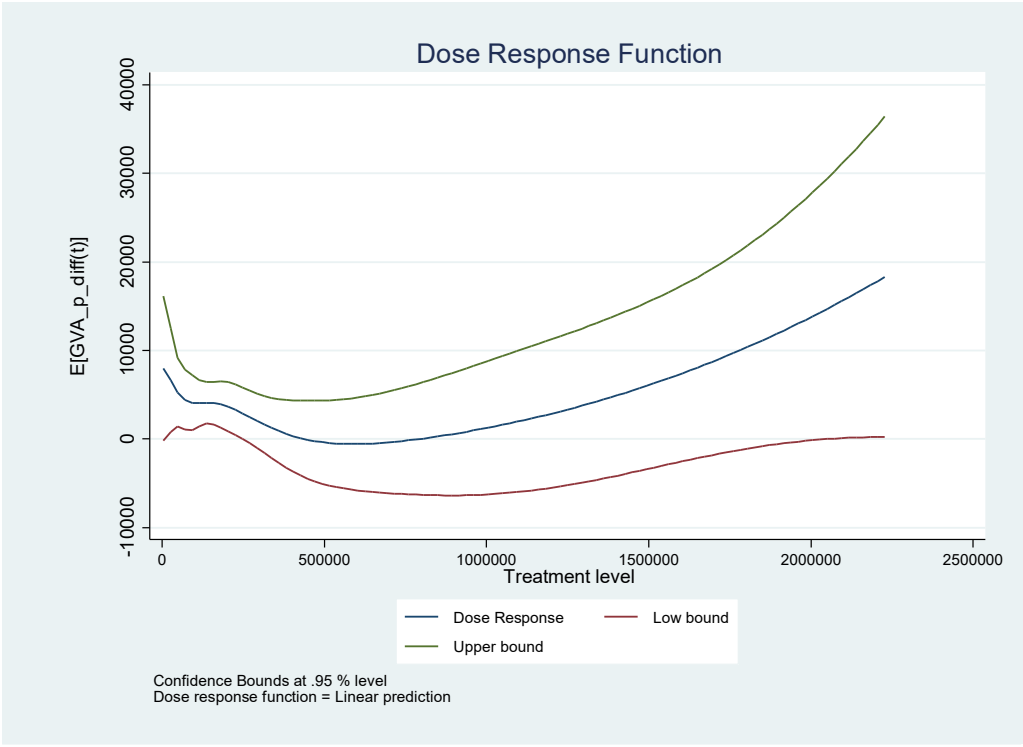
After estimating the programme's effects based on the FADN sample data (see Step 1) an evaluator can extrapolate these results to the whole population of supported farms under FA 2A or FA 3A. This can be done by mapping specific programme effects obtained for each **intensity interval** for each cluster of farms that were supported within those intervals.

Example illustrating case 5

Example 5 assumes the programme's effects for specific intensity levels (from Step 1) are as shown in Graph 1. From Graph 1 one can conclude that the estimated effects of the RDP on the **change** of GVA (a difference between GVA in 2006 and 2015) was a highly **non-linear function** of support intensity. At the lowest level of support intensity (i.e., between 0 and 17000 EUR) an increase of support resulted in **decreasing marginal effects** on GVA, which dropped from approximately 257 EUR per farm, at a minimal level of support, to almost 0 EUR, at the support level equal to approximately 17000 EUR. Only with support level 17001 EUR per farm and higher the GVA in supported farms started to increase and reached the local maximum of about 514 EUR at the support intensity level = 64000 EUR (2250000 SKK).

[19] See:
- Hirano, K. and Imbens, G., (2004) The Propensity score with continuous treatment, *Missing data and Bayesian Method in Practice: Contributions by Donald Rubin Statistical Family* in the book Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives. Edited by A. Gelman and X-L. Meng, 2004 John Wiley & Sons, Ltd;
- Imai, K. and van Dyk, D.A. (2004) Causal inference with general treatment regimes: Generalizing the propensity score. Journal of the American Statistical Association99:854-866;
- Bia, M. and Mattei, A. (2007) Application of the Generalized Propensity Score. Evaluation of public contributions to Piedmont enterprises, Department of Public Policy and Public Choice - POLIS;
- Michalek J., Ciaian P. and Kancs, d'A. (2014). Capitalization of CAP Single Payment Scheme into Land Value: Generalized Propensity Score Evidence from the EU, Land Economics, 90:260-289.;
- Kluve, J. et.al. (2012). Evaluating continuous training programs using the generalized propensity score, Journal of the Royal Statistical Society: Series A (Statistics in Society) 175 (2):587–617;
- Michalek J. (2012). Counterfactual impact evaluation of EU Rural Development Programmes - Propensity Score Matching methodology applied to selected EU Member States", Volume 2 – A regional approach, European Commission, JRC Scientific and Policy Reports, 1-83

Graph 1: Estimated dose response function with respect to intensity level of support on the change in Gross Value Added (GVA) per farm[20]



Source: Slovakia, ex post evaluation of RDP 2007-2013

If the evaluator has the knowledge of the distribution of farms for each support intensity cluster within the total population of supported farms under FA 2A then they can estimate the total effect of the programme. In this case the calculations would be as follows (see Table 7).

Table 7. Estimated effect of support under FA 2A on the change of GVA (per farm)

	Support intensity level between 0-17000 EUR or (0-595000 SKK)	Support intensity level between 17001-51428 EUR or (595001–1800000 SKK)	Support intensity level between 51429-64000 EUR or (1800001–2250000 SKK)
Number of supported farms in total*	1000	2000	500
Average change in GVA in EUR per farm	143 EUR	229 EUR	486 EUR
Total effect of obtained support on GVA change	1000 * 143 = 143000 EUR	2000 * 229 = 458000 EUR	500 * 486 = 243000 EUR

* number of supported farms (only as an example)

In total, due to received support, the GVA in all supported farms increased (between 2006 and 2015) by +844000 EUR = 143000 EUR + 458000 EUR + 243000 EUR.

Similar calculations can be carried out for all relevant common result indicators (e.g., R2) and thus provide answers to CEQ 4.

[20] In Graph 1, both the level of support intensity and GVA is expressed in SKK (1 euro = 35 SKK)

CASE 6

In the FADN database for a particular EU Member State there is not a sufficient number of observations enabling the estimation of the SATT effects.

In this case there are three possible solutions:

Solution 1: Evaluators can revert to other sources of micro-economic information about performance of supported farms (i.e., farm bookkeeping data or surveys (existing or planned)).

Solution 2: Evaluators may use disaggregated socio-economic regional data (e.g. available at LAU 1 or LAU 2 level) combined with information about the level of support under FA 2A or FA 3A distributed in each of the LAU 1 or LAU 2 regions in order to estimate the effects of the support based of this regional data.

Solution 3: Evaluators may use existing regional or sectoral models enabling them to simulate the RDP's effects on the agricultural sector.

GUIDING QUESTION: Given that the FADN is the first choice as a data source for the calculation of important policy parameters (e.g., ATT) needed for answering CEQ 4 and CEQ 6, how can the FADN be best utilised in order to answer CEQ 27?

Common Evaluation Question 27 (CAP objectives)

> *'To what extent has the RDP contributed to the CAP objective of fostering the competitiveness of agriculture?'*

As mentioned in Section 1 'CEQs to assess RDP effects on fostering farm competitiveness in agriculture', the CEQ 27 requires the estimation of RDP effects on the competitiveness of the **whole agricultural sector**[21], explicitly targeting both RDP direct beneficiaries supported under all focus areas (i.e., 1A-6C) as well as non-beneficiaries that may be indirectly affected by the RDP support.

First the evaluators should estimate the net impact of the RDP provided under FA 2A and FA 3A, as well as, under all focus areas by applying a quasi-experimental approach and using the relevant result (e.g., R2) and impact indicators (e.g. I01, I02 and I03) as measures of competitiveness. In this step calculation of the RDP's impacts can be based on FADN data, selected from FADN variables which allow for the construction of the appropriate comparison groups, as well as, used for the construction of the result indicators (common (R2), programme specific and additional) and impact indicators (common (I01-I03), programme specific and additional).

Then the evaluator should **extrapolate** the results obtained previously for the whole population of farms supported under FA 2A, FA 3A and other focus areas, by taking into consideration issues related to **representativeness** of individual groups of RDP beneficiaries in the whole population of farms receiving support from the RDP (see: Sections 2 'Use of farm-level data' and 3 'Specific considerations for using FADN'.).

The basic issue to be taken into consideration while estimating the effects of the FA 2A, FA 3A and all focus areas related to the competitiveness of the whole agricultural sector in a given country or particular programming area is to answer the question, 'how can the knowledge about the **direct and indirect programme effects on farms supported** be used for the estimation of **country-wide (or programme area) spill-over effects on other farms** (RDP non-beneficiaries)?'

[21] Although RDP can affect many sectors in economy, e.g. the agricultural sector, the food processing sector, forestry sector, other sectors located in rural areas, etc. under CEQ27 evaluator is asked to focus on the effect of RDP on the change of competitiveness of the agricultural sector only.

There are three possibilities as to how these effects can be considered and how CEQ 27 can be answered. It is important is to note, that in most of the cases, **answering CEQ 4 and CEQ 6 is a precondition for answering CEQ 27**.

Approach 1: Evaluators can try to estimate both direct effects of support provided under all focus areas or under selected focus area (e.g., FA 2A only) together with intra-regional spill-over effects from supported to non-supported farms using disaggregated regional data and applying a quasi-experimental methodology. In this approach control groups would be regions that have not received support under the RDP (e.g., FA 2A/FA 3A) or received support from the RDP, but at a different level of support intensity. Examples of such an approach can be found in various studies[22].

Approach 2: Evaluators can try to use and update existing regional or sectoral models. By adjusting the respective response parameters, or the whole supply function, on the basis of information from the estimated ATE (support under FA 2A, FA 3A and other focus areas) an evaluator can attempt to assess the effects of a shock from a given RDP (all focus areas) by comparing obtained results 'with support' to a baseline 'without the RDP support'. In this case both intra-regional as well as inter-regional spill-over effects can be accounted for.

Approach 3: Evaluators may apply the recently developed approach based on a quasi-experimental methodology taking explicitly into consideration spill-over effects[23].

Key points to take away for answering CEQ 27

- ⇨ Answering the CEQ 27 is **not** possible before answering CEQ 4, CEQ 6, and other questions requiring the calculation of the respective ATTs for farms who received support under various focus areas. Furthermore, information on the **ATT** obtained from the analysis of the FADN data **should be utilised** together with the analysis of the RDP **spill-over effects** as an **input** for the calculation of **aggregated** RDP effects **at the country level using macro-economic** models (e.g., partial or general equilibrium).
- ⇨ When answering CEQ 27 information on the **ATT**, obtained from the analysis of the FADN data, for **different types of supported units** (e.g., FA 2A) can also be used in the analysis based on the disaggregated regional data (e.g., at LAU 1 level) in the framework of a quasi-experimental evaluation methodology.
- ⇨ Estimation of the RDP's effects on the competitiveness of the whole agricultural sector is a rather difficult task and its quality depends on a number of key issues:
  - Availability of appropriate data (e.g., highly disaggregated socio-economic regional data at LAU 1 or LAU 2 levels (see Approach 1)).
  - Availability of adjusted analytical instruments (e.g., partial or Computable General Equilibrium (CGE) models (see Approach 2)).
  - Skills of evaluators (e.g., knowledge of recently developed advanced quantitative evaluation techniques (see Approach 3)).

In all cases the answers provided to CEQ 4 and CEQ 6, **estimation of net effects on supported farms by using ATT, is a vital precondition to answering the CEQ 27**.

[22] Michalek, J. et.al., (2020) Regional impacts of the EU Rural Development Programme: Poland's food processing sector 54(10):1389-1401 and Michalek J. (2012), *'Counterfactual impact evaluation of EU rural development programmes - Propensity Score Matching methodology applied to selected EU Member States', Volume 2 – A regional approach*", European Commission, JRC Scientific and Policy Reports, pp 1-83.

[23] Cerulli, G., (2017) Identification and Estimation of Treatment Effects in the Presence of (Correlated) Neighborhood Interactions: Model and Stata Implementation via Ntreatreg. The Stata Journal,17(4):803-833.

European Evaluation Helpdesk
Boulevard Saint-Michel 77-79
B - 1040 BRUSSELS
T: +32 2 737 51 30
Email: info@ruralevaluation.eu
http://enrd.ec.europa.eu