European Evaluation Helpdesk for Rural Development

European Commission

INTERACTIVE DECISION TOOL

DATA FOR THE ASSESSMENT OF RDP ACHIEVEMENTS AND IMPACTS

IMPACT INDICATORS:

I.01 AGRICULTURAL ENTREPRENEURIAL INCOME

I.02 AGRICULTURAL FACTOR INCOME

I.03 TOTAL FACTOR PRODUCTIVITY IN AGRICULTURE

CONTENTS LOGIC MODEL

JANUARY 2019

Contents

Introduction
How to use the tool
Logic Model

Acknowledgements
Disclaimer
Contacts

Copyright notice

© European Union, 2019

Reproduction is authorised provided the source is acknowledged.

Recommended citation:

EUROPEAN COMMISSION – Directorate-General for Agriculture and Rural Development – Unit C.4 (2019): Interactive decision tool - “Data for the assessment of RDP achievements and impacts” with a focus on data and information sources pertinent for the assessment of RDP achievements and impacts in 2019 and ex-post . Brussels, January 2019.

Disclaimer:

The information and views set out in these Guidelines are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in these Guidelines. Neither the Commission nor any person acting on the Commission’s behalf may be held responsible for the use which may be made of the information contained therein.

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).

ACKNOWLEDGEMENTS

The interactive **decision tool, 'Data for the assessment of RDP achievements and impacts'**, *has been developed by an international team of rural development evaluation experts including Jerzy Michalek, Demetrios Psaltopoulos, Dimitris Skuras, Jela Tvrdonova, Darko Znaor. The related thematic working group has been coordinated by the Evaluation Helpdesk under the guidance of Valdis Kudiņš and Hannes Wimmer. Giulia Bekk, Valérie Dumont, Matteo Metta and Myles Stiffler supported the development work and ensured the quality and visual appearance of the final interactive tool. Various experts have provided valuable input as peer reviewers (Juris Hāzners, Jaroslav Pražan, Gerald Schwarz). Representatives of DG Agriculture and Rural Development have ensured the coherence of the tool with the EU's policy framework.*

The interactive **decision tool, 'Data for the assessment of RDP achievements and impacts'**, is based on the logic model approach which was originally developed by the EU collaborative project ENVIEVAL (Grant Agreement No. 31207 in the EU's 7th Framework Programme for research, technological development and demonstration). The Evaluation Helpdesk has applied this approach for its thematic working group, which serves to support Member States in their assessment of RDP impacts in 2019 and the ex-post.

INTRODUCTION

The choice of a suitable evaluation approach is a critical step in the evaluation process. The wish to carry out a robust assessment of the policy's effects needs to be matched with those aspects which factor into conducting an evaluation (data and information availability, budget and resources, and the skills of the evaluators).

In the non-binding Guidelines, '*Assessment of RDP impacts and achievements in 2019*', published in August 2018, logic models have been presented for the 13 Common CAP impact indicators covering Pillar II. These logic models support Member States in discussing different criteria for the choice of evaluation approaches for assessing the RDP's impacts during the evaluation activities in 2019 and the ex-post (2024).

The **decision tool, 'Data for the assessment of RDP achievements and impacts'**, transports the logic models developed in the above Guidelines into an interactive format, while providing further detailed and practical information. The decision tool has been specifically **designed for RDP** evaluators who may wish to gain further insights into the criteria for each step of the decision making process when choosing an evaluation approach. This tool also provides practical recommendations on what to do in case of data gaps both in the short and long term, when solutions are needed.

OBJECTIVES

The **interactive decision tool** consists of a set of 7 logic models covering the 13 Pillar 2 CAP Impact Indicators. The 7 logic models can be read separately and aim to:

- **Assist evaluation stakeholders** in their decision on which evaluation approaches they can use for the assessment of the common RDP impact indicators, as well as providing the necessary data and information sources at the EU level for these approaches.
- **Provide recommendations on possible solutions for overcoming data-gaps at the national and regional levels** (e.g. by providing guiding questions, practical hints and links to external information sources).

The tool focuses on data and information sources pertinent for the assessment of RDP achievements and impacts in 2019 and the ex-post. The decision tool is based on the Guidelines '*Assessment of RDP impacts and achievements in 2019*'. Additionally, the tool provides:

- Explanations on data needs for proposed evaluation approaches including availability and suitability of data for RDP evaluations (frequency, delays, time series).
- Important questions to consider.
- Links to existing data sources and good practices.
- Complementary information on evaluation methods and their data needs.

HOW TO USE THE TOOL

This interactive decision tool contains a set of **7 logic models**:

Sector-related impacts

- I.01 Agricultural entrepreneurial income
- I.02 Agricultural factor income
- I.03 Total factor productivity in agriculture

Socio-economic impacts

- I.14 Rural employment rate
- I.15 Degree of rural poverty
- I.16 Rural GDP per capita

Environmental impacts

- I.07 Emissions from agriculture
  - I.07 – 1 GHG emission from agriculture
  - I.07 – 2 Ammonia emissions from agriculture
- I.08 Farmland Bird Index (FBI)
- I.09 High Nature Value (HNV) farming
- I.10 Water Abstraction in Agriculture
- I.11 Water Quality:
  - I.11-1 Gross Nutrient Balance (GNB) *(Gross Nitrogen Balance (GNB-N) and Gross Phosphorus Balance (GNB-P))*
  - I.11-2 Nitrates in freshwater
- I.12 Soil organic matter in arable land
- I.13 Soil erosion by water
  - I.13-1 Estimated rate of soil loss by water erosion;
  - I.13-2 Estimated agricultural area affected by a certain rate of soil erosion by water

HOW TO USE THE TOOL

Navigation within the clickable logic model:

Brings the user back to the starting page of the **logic model**

Are variables explaining participation known?

Takes the user to that specific **decision question** of the logic model

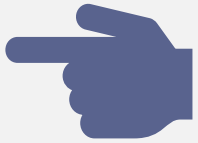
Starting decision question of the logic model

Hyperlinked text

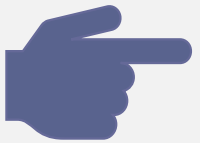
Takes the user to an external source or to another slide

Examples

Additional notes



Previous page



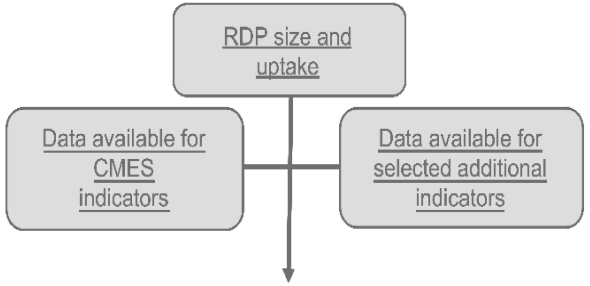
Next page

HOW TO USE THE TOOL

Structure:

Each logic model begins with a description of the:

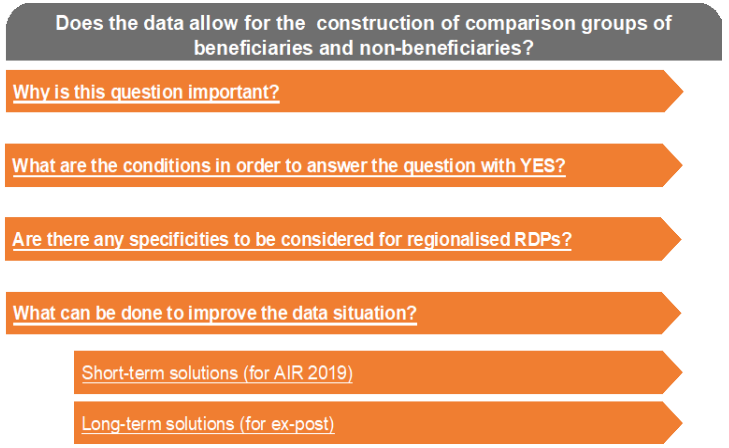
- **RDP size, uptake** and other aspects that have to be considered for the selection of the evaluation approach.
- **Data availability for CMES indicators** needed to assess net impacts at the micro and macro levels, as well as, the specificities in the data availability for regionalised RDPs.
- **Data availability for selected additional indicators**.



Each **decision question** is organised in a way that facilitates the answers to the following **sub-questions**:

- Why is this question important?
- What are the conditions in order to answer the question with YES?
- Are there any specificities to be considered for regionalised RDPs?
- What can be done to improve the data situation (In the short term (for AIR 2019) and long-term (for ex-post)?

Each sub-question can be explored by clicking on its link.

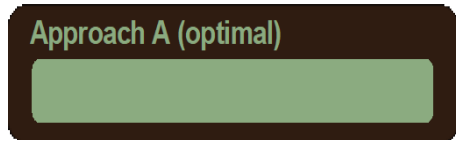


By answering each decision question in the tool with either a '**YES**' or a '**NO**' one will be taken to the next question, which will ultimately lead one to all possible evaluation approaches that can be applied given the specific criteria they have selected.

HOW TO USE THE TOOL

The tool will suggest various applicable approaches based on the data and other information:

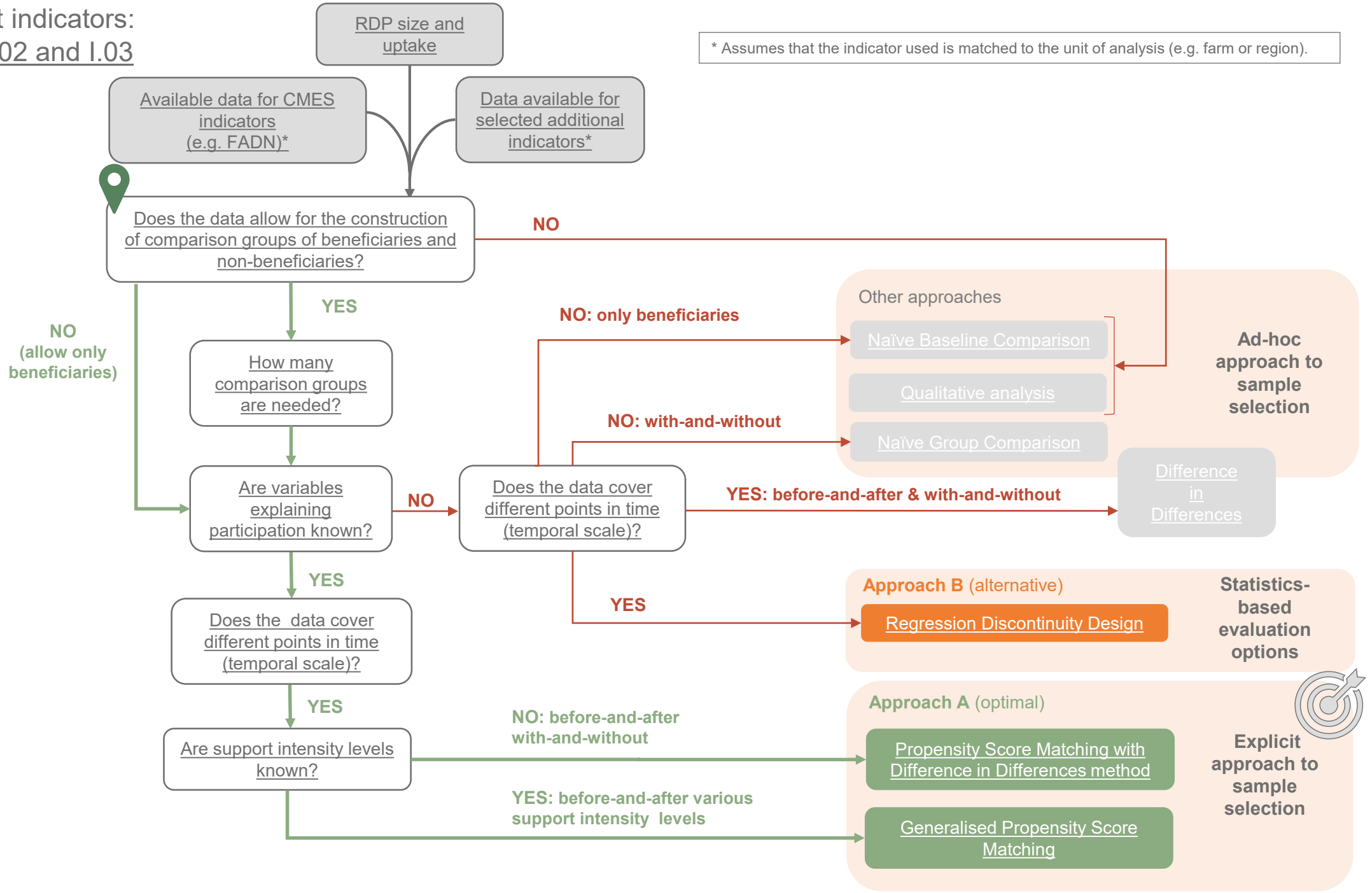
- Approach A **(an evaluation approach in an optimal data situation).**
  It can be used in 2019 and/or can be planned for the ex post evaluation.



- Approach B **(an alternative evaluation approach in case of data gaps).**
  In several cases, approach B contains a qualitative component.

In case of questions or any technical difficulties in accessing the files, please contact the European Evaluation Helpdesk for Rural Development:
E-mail: info@ruralevaluation.eu
T: +32 2 737 51 30

Impact indicators: I.01, I.02 and I.03



* Assumes that the indicator used is matched to the unit of analysis (e.g. farm or region).

Impact indicators

I.01 Agricultural entrepreneurial income

I.02 Agricultural factor income

I.03 Total factor productivity in agriculture

Impact indicators fiches

Guidelines *Assessing RDP achievements and impacts in 2019*, PART II, Chapter 2.2.1, Section: 'Intervention logic'

Related Common Evaluation Question:

CEQ 27: 'To what extent has the RDP contributed to the CAP objective of fostering the competitiveness of agriculture?

Guidelines *Assessing RDP achievements and impacts in 2019*, PART III, Chapter 3.6, Section: 'Clarification of general intervention logic linked to the CEQ'

RDP size and uptake

The size, uptake and the structure of the RDP are important factors in the selection of the evaluation approach. If only a few units are supported by measures under the rural development priorities and/or focus areas then significant RDP effects are not expected. In such a situation the evaluator might choose a less robust evaluation approach (e.g. naïve group comparison, or qualitative assessment). Restrictions concerning the interpretation of calculations based on the above simplified techniques should be taken in to consideration, especially regarding the magnitude of a possible selection bias. If there is sufficient uptake the evaluator can apply advanced evaluation approaches.

Data availability for CMES indicators

What is the unit of analysis and data available (EU-level)?

- At the micro level, the unit of analysis is the agricultural holding.
- At the macro level, the unit of analysis is the RDP area.
- Specifities in the data availability for regionalised RDPs.

Data availability for CMES indicators — Micro level

Data available:

- FADN database (recommended)
- Farm bookkeeping database (highly recommended if available)
- Surveys (less recommended)
- Additional sources for deriving price indices needed for calculation of impact indicator I.03 (Total factor productivity in agriculture) at the farm level can be found at Eurostat (Key European statistics and Economic Accounts for Agriculture) or national statistics.
    - A more detailed description of data necessary for the calculation of I.03 can be found in the Guidelines *Assessing RDP achievements and impacts in 2019*, PART IV, Chapter 4.1.6

Note

FADN database (recommended)

- FADN collects accountancy data from about 80,000 agricultural holdings in the EU. The FADN database includes several hundreds of various types of variables and is representative of commercial agricultural holdings in the EU. Additionally, FADN is the only source of farm micro-economic data that is harmonised (the bookkeeping principles are the same in every Member States). The main institutions responsible for the preparation and processing of FADN data (e.g. regional breakdowns) are the national Liaison Agencies (national FADN units).

- FADN data is available not only for each EU Member State, but also, for many FADN regions within a given country. The latest update of FADN data in 2019 will provide data for year the 2017.

Examples:
- Germany
- Poland
- France
- Belgium

Note for regionalised RDPs

- Concerning the coverage and applicability of FADN in the assessment of RDP's net impacts for 18 Member States which have only one RDP, the entire available FADN (or farm bookkeeping) sample can be used.

Note

Farm bookkeeping database (highly recommended if available)

- Farm bookkeeping data which in terms of the number of farms is much richer than FADN, may be available in many EU countries.

Examples:

- Germany
- Slovakia
- Austria
- Poland

- Farm bookkeeping data is collected yearly by professional national farm accounting organisations and/or associations, however, it is not harmonised across all EU Member States. The latest update in 2019 will present data for the year 2018.

Surveys (less recommended)

- Surveys data (specific surveys, e.g. Farm structure survey (FSS) EU 2013) is available every 3 to 4 years (depends on the country). The latest update in 2019 should contain data on the years 2016 or 2017.

Examples:

- Germany
- Austria

FADN databases:

- **Germany:** data is available for 16 Länder regions and every year from 1981 – present;
- **Poland:** data is available for 4 regions from 2004 - present;
- **France:** data is available for 25 regions, 22 regions of which from 1981 – present, and for 3 regions from 2012 – present;
- **Belgium:** data is available for 3 regions from 2004 - present.

Farm bookkeeping databases:

- **Germany:** Testbetriebsstatistik
- **Slovakia:** In Slovakia the number of farms included in the farm bookkeeping database (approximately 2,500 farms in a sample) is more than 4 times bigger in comparison to FADN (600 farms in the sample)
- **Austria:** Farm bookkeeping database
- **Poland:** Farm bookkeeping database

Surveys:

- **Germany:** German FSS 2016
- **Austria:** Austrian FSS 2016

Datasets used for the computation of CMES impact indicators I.01, I.02 and I.03 and for the subsequent analysis of net impacts of the RDP 2014-2020 must originate from the same source and concern the same farms. The data used should describe the major characteristic and economic performance of each individual farm included in the sample and not be an aggregate for any specific group/type of farming (TF) (e.g. in case of FADN, TF8 groupings are: field crops, horticulture, other permanent crops, milk, etc.; TF14 groupings; economic size classes, etc.)

Individual farm data (delivered to evaluators anonymously, i.e. only with ID farm number) should be in panel data form (i.e. the same units/farms are observed in various periods of time), for example, prior to and after receiving support from the RDP 2014-2020.

For further discussion see section: *Does the data cover different points in time?*

In countries (France, The United Kingdom, Portugal, Spain, Finland, Germany, Italy and Belgium) where there are more than one RDP to be evaluated, a closer look at the coverage of each RDP by FADN or farm bookkeeping data is required

See section: *Specificities in the data availability for regionalised RDPs*

It is important that FADN or farm bookkeeping data is supplemented with the following information to be obtained from the Paying Agency:

- Information specifying which individual farms included in the FADN or bookkeeping dataset were beneficiaries of the RDP 2014-2020 (in which years).
- Intensity of support received by each farm identified in the FADN dataset from individual RDP measures during the programming period 2014-2020.
- Ammount of support (in EUR) received by each farm identified in the FADN (beneficiaries and non-beneficiaries) shortly prior to the beginning of the RDP 2014-2020 programming period (e.g. year 2013).
- Total number of supported farms for each support category (measure, FA and priority) during the period 2014-2020.

Macro level:

- The up-scaling can be done by multiplying net effects calculated for an average supported farm included in the FADN (or bookkeeping) sample by total number of supported farms in a given RDP region.

  Note

- Computation of net programme impacts at a macro - (regional) RDP level does not require acquiring any specific data at a macro-level (except of a total number of supported farms in each support category).

- The indicators I.01, I.02 and I.03 calculated from the Economic Accounts for Agriculture are currently available only at the macro-level for each Member State and for several calendar years.

  Note

In the process of results extrapolation, additional information should be utilised regarding the distribution of individual types of beneficiary farms in the available FADN or bookkeeping sample in comparison with the overall population of farms supported from a given RDP.

By interpreting the obtained results, the calculated net results of the programme's impacts at the macro-level can be contrasted with the contextual data from Eurostat (Economic Accounts for Agriculture, Agriculture labour input, National Accounts) or from the Total factor productivity (TFP). This data describes the comparable observable trends calculated for impact indicators (I.01, I.02 and I.03).

The changes in these indicators over time (e.g. starting from 2014) represents a gross effect caused by a number of factors including the influence of other exogenous (i.e. RDP independent) factors.

Specificities in the data availability for regionalised RDPs

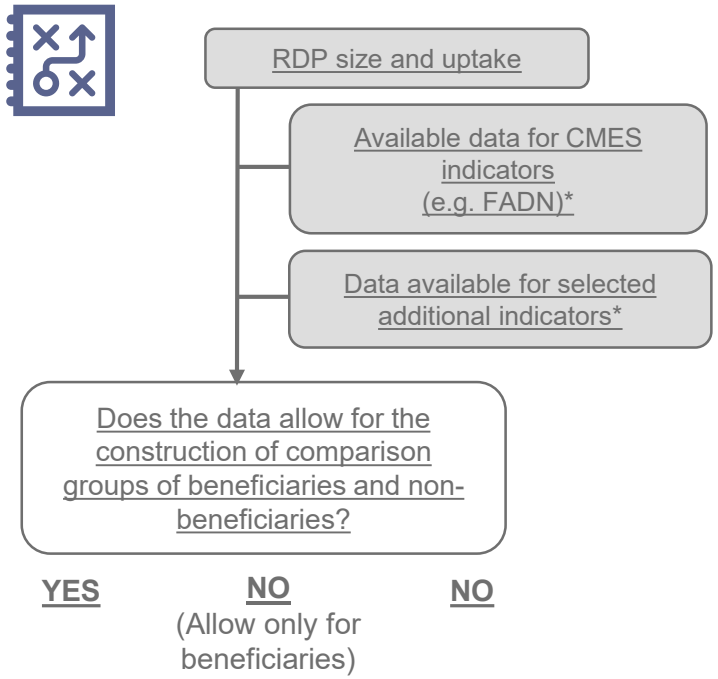
- For regionalised RDP programmes if the number of beneficiary farms identified in the FADN or bookkeeping data is low and the RDP programmes do not substantially differ from each other, it is recommended to combine FADN datasets covering more RDP programmes (FADN regions).

  For details see the section on regionalised RDPs: *Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?*

- If regionalised RDP programmes differ from each other and the number of RDP beneficiaries in each regional dataset (FADN or bookkeeping) is low, additional thematic surveys could be a solution.

Data availability for selected additional indicators

Examples of additional indicators, unit of analysis and data sources are provided in the Guidelines *Assessing RDP achievements and impacts in 2019*, PART IV, Chapter 4.1, Section 4.1.1 ‘Additional indicators (examples)’



Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

Why is this question important?

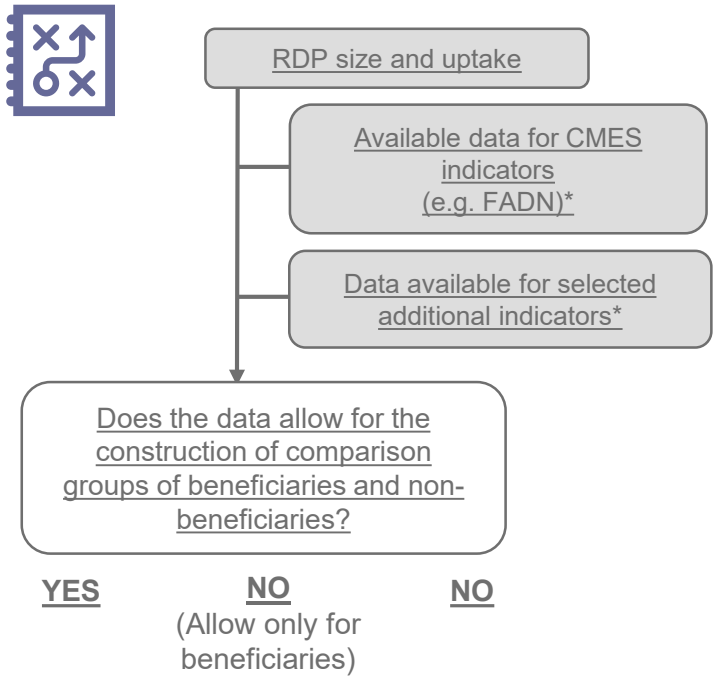
What are the conditions in order to answer the question with YES?

Are there any specificities to be considered for regionalised RDPs?

What can be done to improve the data situation?

- Short-term solutions (for AIR 2019)
- Long-term solutions (for ex-post)

Guidelines *Assessing RDP achievements and impact in 2019*, PART II, Chapter 2.1 and 2.2 and PART IV, Chapter 4.1.
Guidelines *Assessment of RDP results*, Chapter 2.1 and 6.2, and Annex 11, Chapter 2.4.
Guidelines for the ex post evaluation of 2007-2013 RDPs, Chapter 4.3



Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

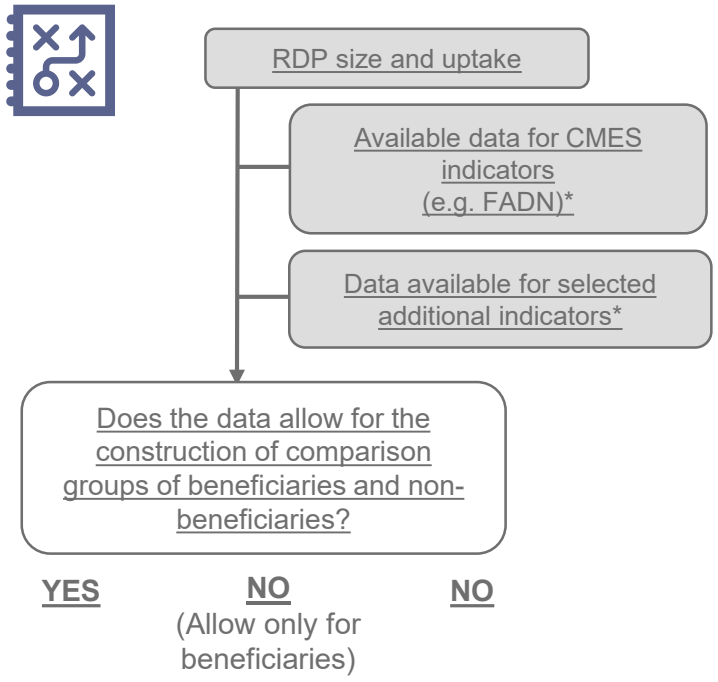
Why is this question important?

Analysis of net impacts of RDP programmes requires the use of a counterfactual approach and therefore the construction of comparison groups. First, FADN or bookkeeping data should be verified to see whether the data allows for clearly distinguishing farms which were supported by the RDP 2014-2020 (programme beneficiaries) from other farms which were not supported (non-beneficiaries). This can only be done in cooperation with the Paying Agency.

Whether a data structure enables one to construct a comparison group consisting of 'programme beneficiaries and non-beneficiaries' will predetermine the type of net impact analysis (a methodological approach) which can be carried out.

In case the FADN or bookkeeping data allows for the construction of comparison groups of programme beneficiaries and non-beneficiaries, the impact analysis can be carried out by applying a standard binary type of matching techniques (e.g. PSM-DiD). This is in contrast to a situation when the construction of two comparison groups is not possible (e.g. because all farms in the sample are identified as programme beneficiaries). If this is the case, other quasi-experimental techniques and methodological approaches (e.g. Generalised Propensity Score Matching (GPSM)) can be used.

From the fitted outcome model using GPSM, the dose-response function can be estimated for outcomes across all levels of the RDP support (incl. zero as a special case).



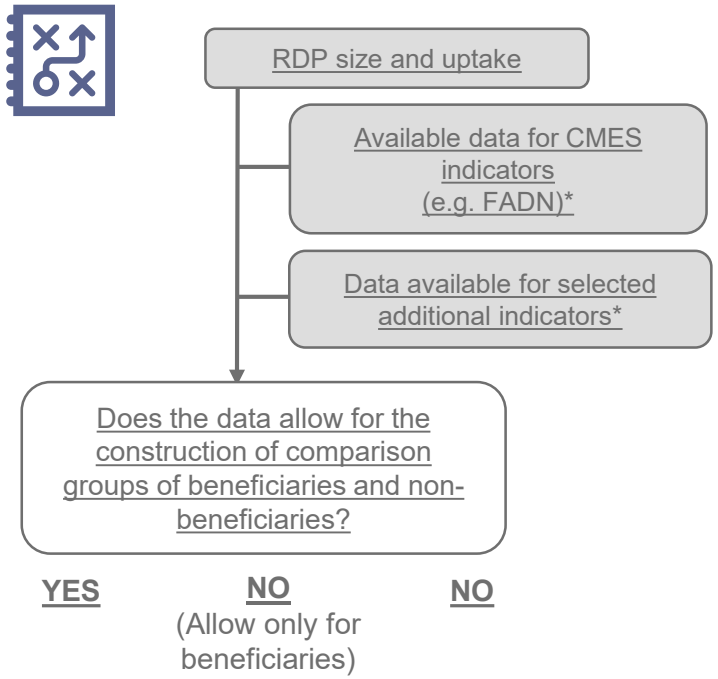
Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

What are the conditions in order to answer the question with YES?

- ✓ Database available contains farms of which one part was supported from the programme and the other part was non-supported.
- ✓ There is sufficient information available on both groups (adequate number of observations on farms which belong to these two groups).
- ✓ The dataset includes the most important control variables which will allow for the establishment of similarities between these two groups prior to the beginning of the programme.

FADN or bookkeeping databases distinguish both types of farms.

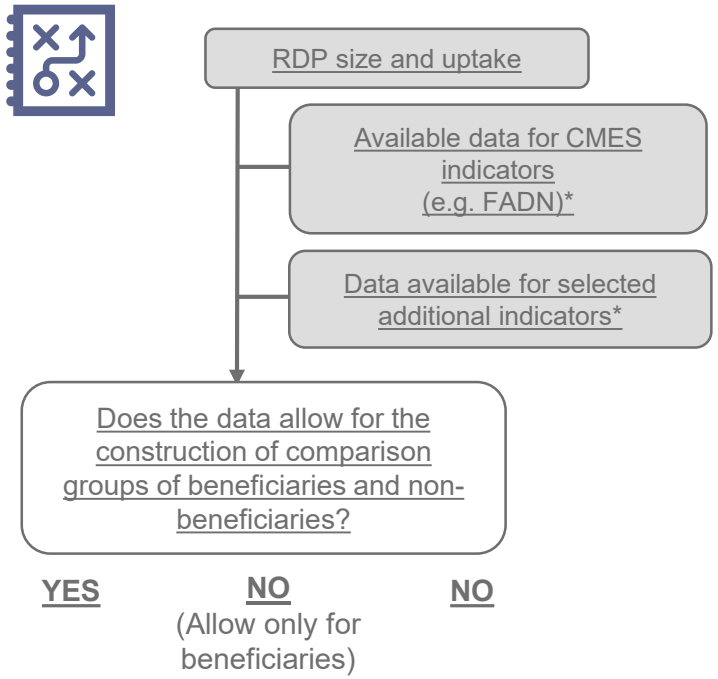
- Note
- Note
- Note
- Note



Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

What are the conditions in order to answer the question with YES?	Note(s)

- An evaluator should check if samples of programme beneficiaries available in FADN are indeed representative of all farms supported by the RDP 2014-2020 (if not, to what extent is represented?).
- Special attention should be given to the second condition, which requires that both samples of farms covered by FADN or bookkeeping databases (i.e. beneficiaries and non-beneficiaries) should include a sufficient number of observations which allows for further econometric net impact analysis (a statistical/econometric requirement).
- The issues of representativeness and econometric feasibility may be linked with each other.



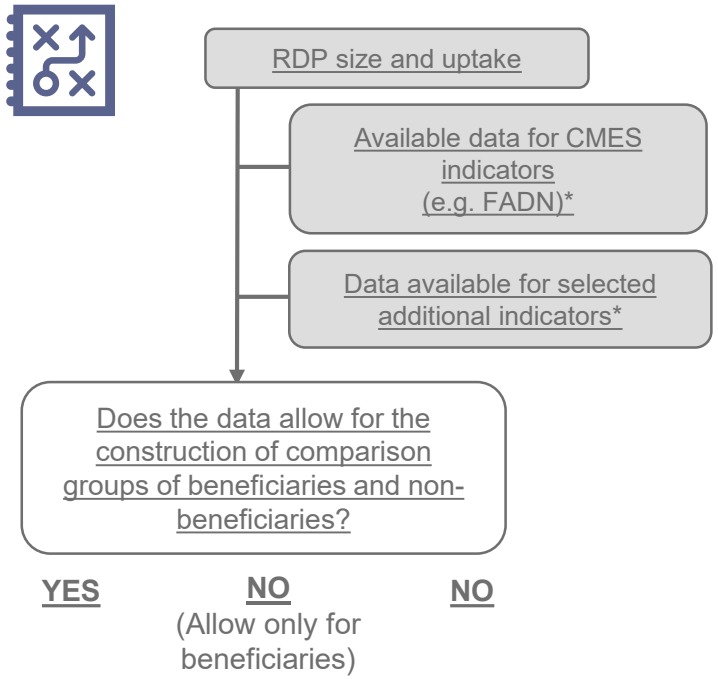
Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

What are the conditions in order to answer the question with YES?	Note(s)

Administrative problems may be encountered by evaluators during data collection. This may concern obtaining various administrative permissions necessary for using and receiving national farm bookkeeping and/or FADN data. Usually, after all these procedures are successfully completed (permissions are obtained) anonymous farm bookkeeping data combined with relevant records from the Paying Agency (indicating which farm is a programme beneficiary) are transferred to the evaluators without any further restrictions.

Example:

- EU report 'Investment Support under Rural Development Policy'



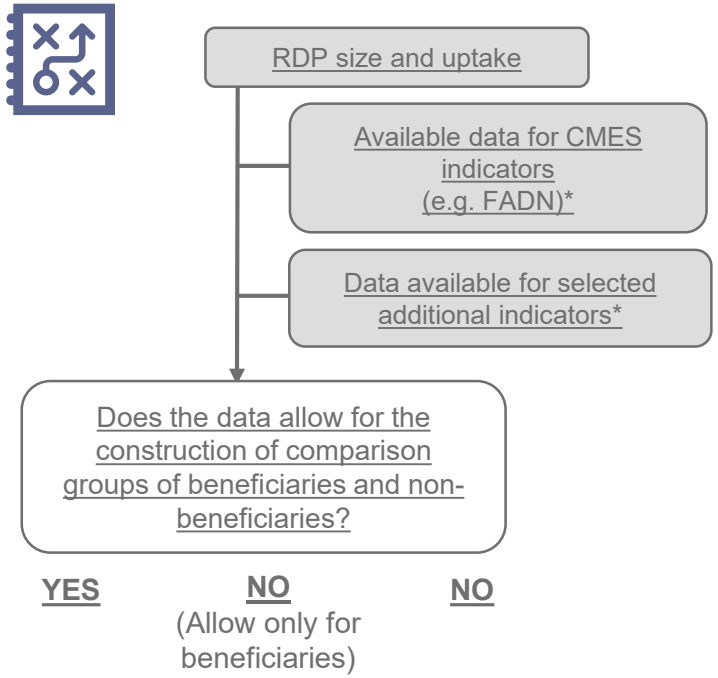
Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

What are the conditions in order to answer the question with YES?

Note(s)

Concerning the sufficient number of observations which allows for the construction of adequate control groups, the evaluator should be aware that:

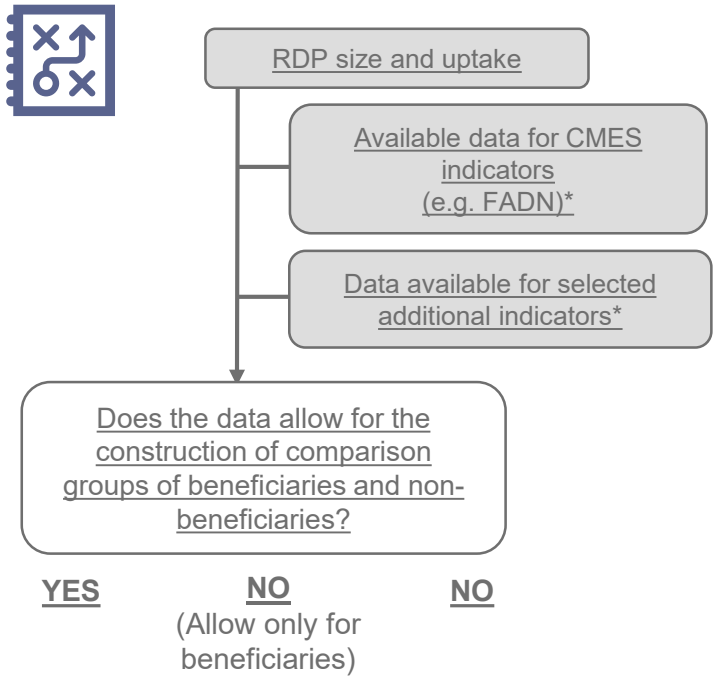
- The number of observations in the group of beneficiaries should be much higher than the number of model covariates (control variables) selected in the matching analysis.
- The number of observations in the group of non-beneficiaries should be higher than the number of observations in the group of beneficiaries (this requirement may be weakened by applying some specific binary matching techniques (e.g. Nearest Neighbours with replacement)).
- Small samples usually have larger error variances than large samples. A small sample size may lead to biased estimations of the marginal treatment effects. This means that in small samples a trade-off between bias and variance arises. The choice of the matching algorithm is therefore important. Therefore, evaluators should use the largest possible number of observations (supported and non-supported farms).



Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

What are the conditions in order to answer the question with YES?	Note(s)

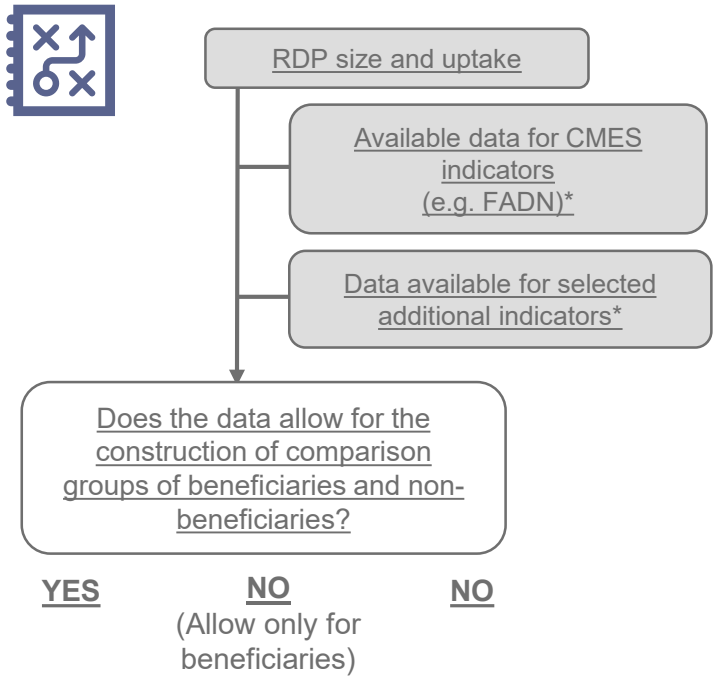
The evaluator should consider what matching techniques are used in the impact analysis for finding appropriate control groups. The binary matching approach further reduces the available number of non-supported farms to those having structural characteristics more homogeneous to those of supported units (outcomes are selected controls belonging to 'common support area'). While a low number of observations may negatively influence the statistical significance of results obtained, working with more farms is always more advantageous. This is particularly the case when using PSM-DiD models, which require large samples and good data.



Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

Are there any specificities to be considered for regionalised RDPs?

Due to more rigid data constraints the construction of comparison groups and subsequent analysis of impacts of regionalised RDPs may create additional problems. These may be linked both to the issue of representativeness as well as to the econometric requirements. For example, assuming that a particular type of farm (e.g. field crops from TF8 groupings) had been assigned a special policy importance (e.g. TF type X) analysis of the available FADN database structure may reveal that due to an insufficient number of observations the econometric estimation of programme effects by farm types (TF) is not feasible.



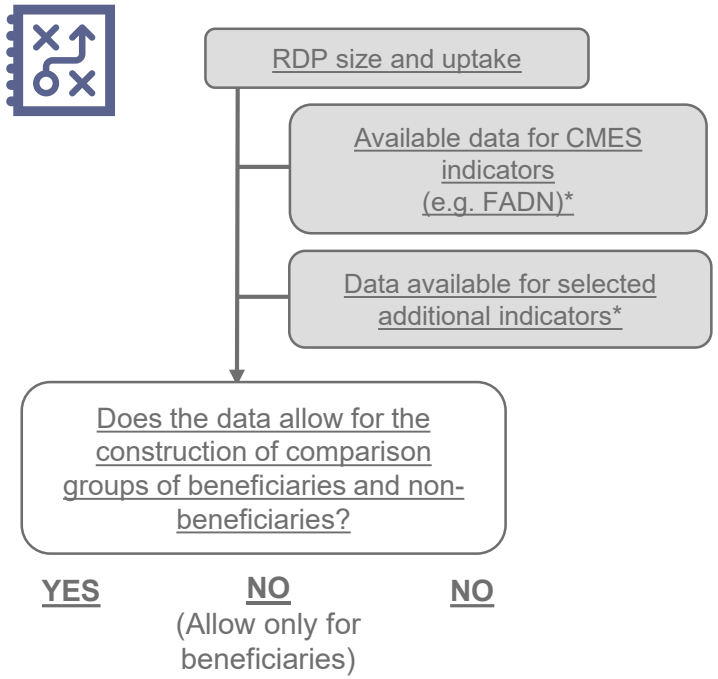
Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

What can be done to improve the data situation?

Short-term solutions (for AIR 2019)

If various RDPs within a country are designed in a similar manner (e.g. similar intervention logic or target group) but statistical information is missing to build a counterfactual at the regional level (e.g. too few observations on individual programme beneficiaries) one possible solution is to construct comparable control groups enabling the netting out of programme effects through combining multiple regionalised RDPs. To do so, firstly, one needs to identify at the farm level a set of common control variables among different programmes/territories. Secondly, additional variables (e.g. dummy variables) can be built into the list of covariates to identify the location of a farm in different RDPs within a country, which allows for the separation of programme effects carried out in various areas under consideration.

For more information on the use of territorial dummy variables see – Pufahl and Weiss, 2009 - Farm Structure and the Effects of Agri-Environmental Programs: Results from a Matching Analysis for European Countries, Evaluating the Effects of Farm Programmes: Results from Propensity Score Matching).



Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

What can be done to improve the data situation?

Long-term solutions (for ex-post)

- Increase the sample (number of observations) in FADN and farm bookkeeping statistics. Additionally, an option might be to extend the list of variables in FADN and farm bookkeeping datasets in order to enable a more detailed analysis of factors interesting from a policy point of view (e.g. determining external farm income). This can be done by distinguishing in the FADN subsidy section the level of support received in a given year from each individual RDP measure.

- Ensure that all beneficiaries of the RDP are fully integrated into the existing farm accountancy or bookkeeping systems with their anonymous records being easily retrievable by respective national authorities (e.g. FADN or farm bookkeeping offices) for evaluation purposes.

- Regular surveys adjusted to better meet the needs of the RDP evaluations.

RDP size and uptake

Available data for CMES indicators (e.g. FADN)*

Data available for selected additional indicators*

Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?

YES

How many comparison groups are needed?

NEXT

How many comparison groups are needed?

Why is this question important?

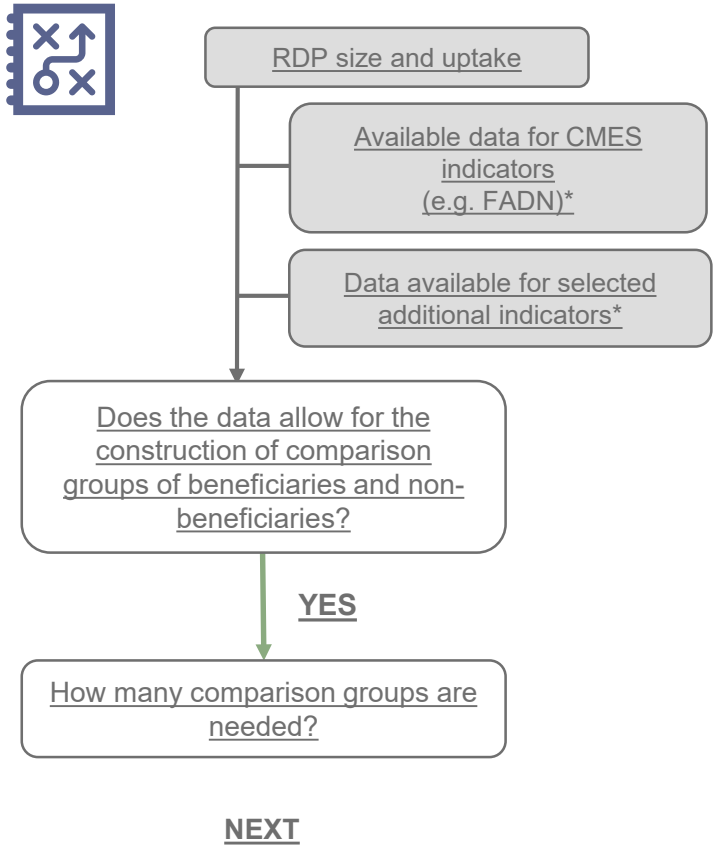
What are the conditions in order to answer the question?

Are there any specificities to be considered for regionalised RDPs?

What can be done to improve the data situation?

- Short-term solutions (for AIR 2019)
- Long-term solutions (for ex-post)

Guidelines *Assessing RDP achievements and impact in 2019*, PART II, Chapter 2.1 and 2.2 and PART IV, Chapter 4.1.
Guidelines *Assessment of RDP results*, Chapter 2.1 and 6.2, and Annex 11, Chapter 2.4.
Guidelines for the ex post evaluation of 2007-2013 RDPs, Chapter 4.3



How many comparison groups are needed?

Why is this question important?

The creation of control groups is an essential part of evaluating programme impacts.

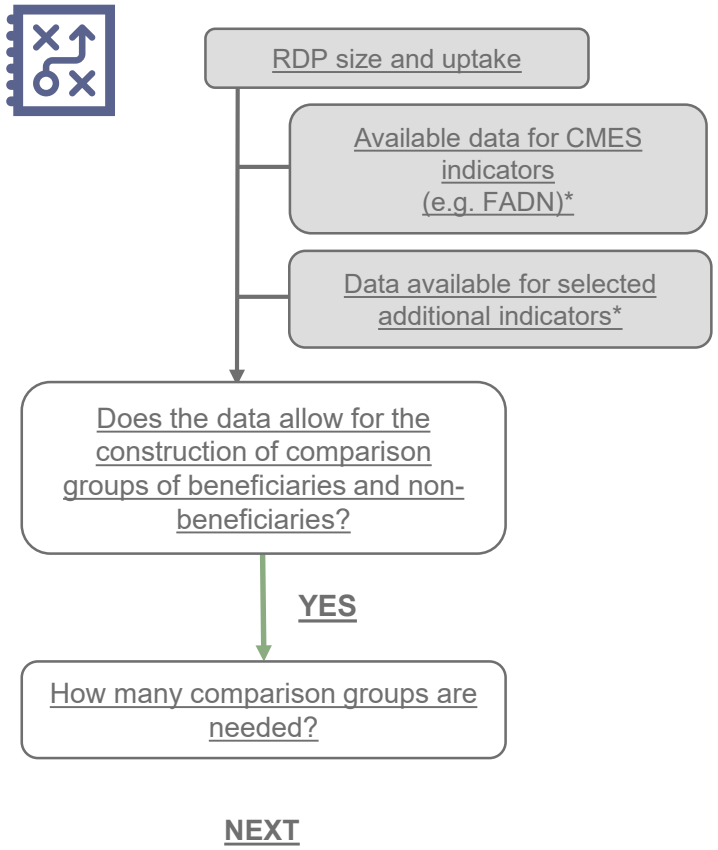
In this context answering the question, 'How many control groups are needed?' will depend on:

- the structure of the available data, and
- the detail level of the net impact analysis to be carried out.

In a standard situation when an impact analysis is to be carried out in order to calculate the Average Treatment Effects on Treated (supported farms) (ATT) only, a binary PSM-DiD method and a distinction in the FADN or bookkeeping data of two comparison groups (i.e. programme beneficiaries and non-beneficiaries) is sufficient.

However, when all farms in the sample are identified as RDP 2014-2020 beneficiaries (non-beneficiaries cannot be found) and/or an evaluator is interested in the estimation of Marginal Treatment Effect for RDP support intensity levels, the application of other quasi-experimental techniques (e.g. Generalised Propensity Score Matching (GPSM)) is necessary.

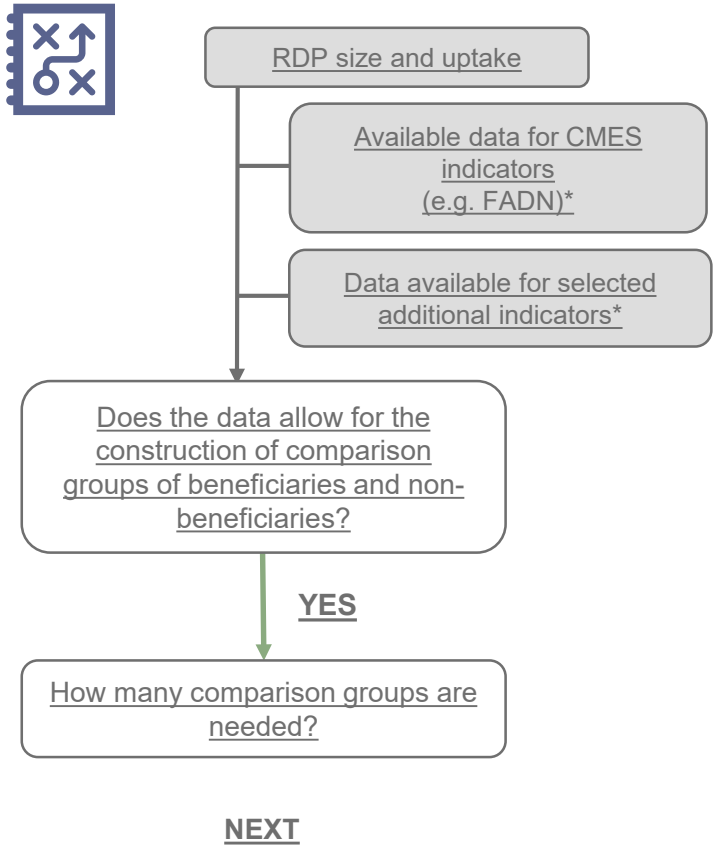
Note



How many comparison groups are needed?

Why is this question important? | **Note(s)**

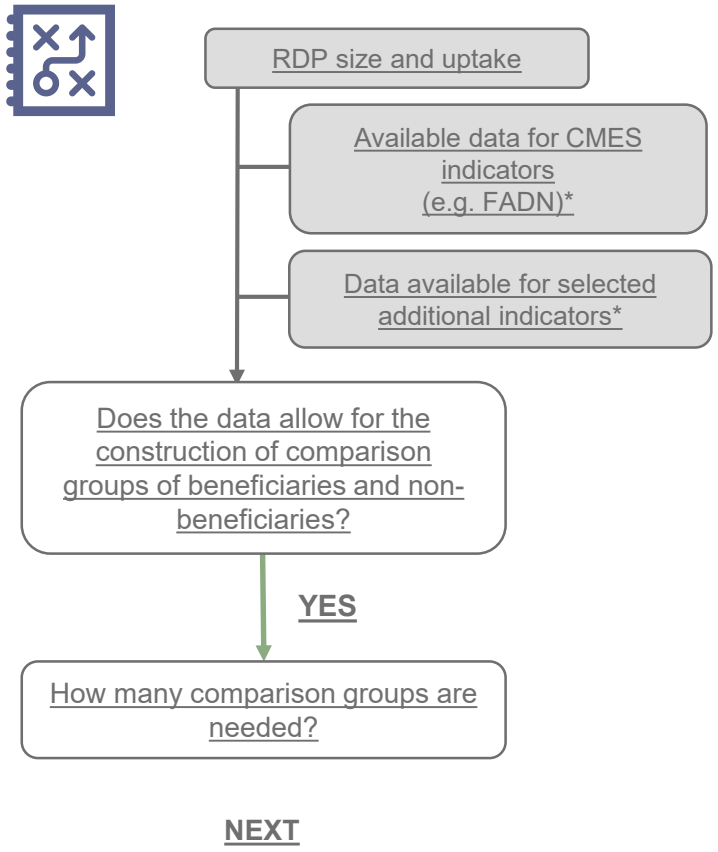
When applying GPSM, various control groups have to be created among programme beneficiaries only, whereby all RDP beneficiaries (given their estimated GPS) are divided into respective control groups characterised by similar GPS intervals and programme support levels (RDP support levels have to be divided into a set of pre-specified intervals).



How many comparison groups are needed?

What are the conditions in order to answer the question?

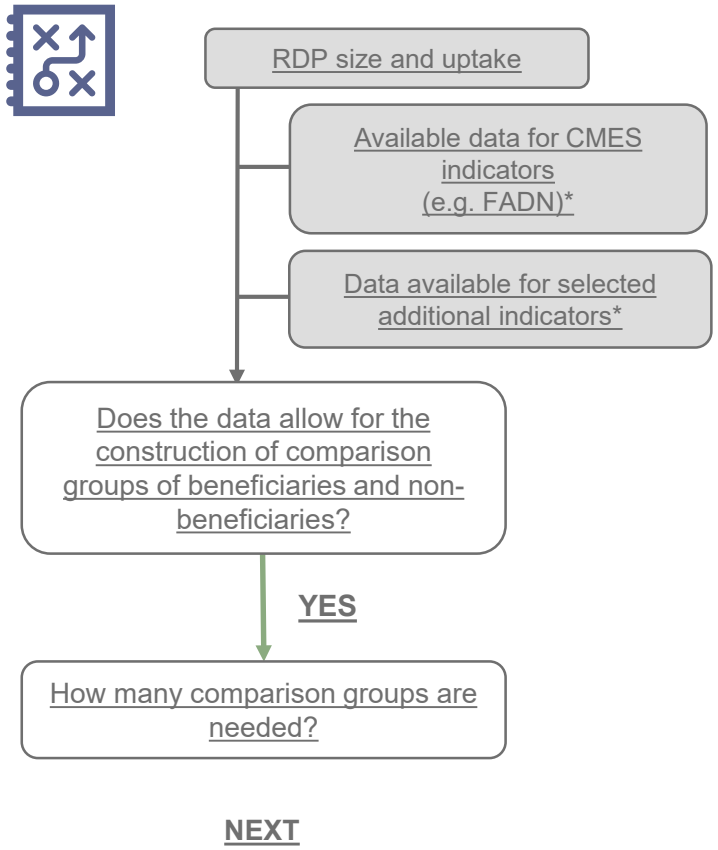
- ✓ The existing data-structure should distinguish programme beneficiaries and non-beneficiaries.
- ✓ Control groups prior to the programme have to be almost identical to each other (except that one received programme support and another did not).
- ✓ Matching techniques should be used to establish similarities between control groups.
- ✓ The use of FADN or bookkeeping data usually allows for the construction of at least two control groups consisting of programme beneficiaries and non-beneficiaries. However, in order to carry out a subsequent econometric net impact analysis another set of criteria as mentioned in the Guidelines *Assessing RDP achievements and impact in 2019*, PART IV, Chapter 4.1. have to be met.



How many comparison groups are needed?

Are there any specificities to be considered for regionalised RDPs?

For regionalised RDP programmes if: (a) the number of RDP beneficiary farms identified in FADN or bookkeeping data is very low, which prevents the construction of comparison groups, and if (b) the RDP programmes do not substantially differ from each other, then it is recommended to combine FADN datasets covering more then one RDP programme together (FADN regions). However, if in the same situation regionalised RDP programmes differ from each other significantly, then additional thematic surveys could be a solution.



How many comparison groups are needed?

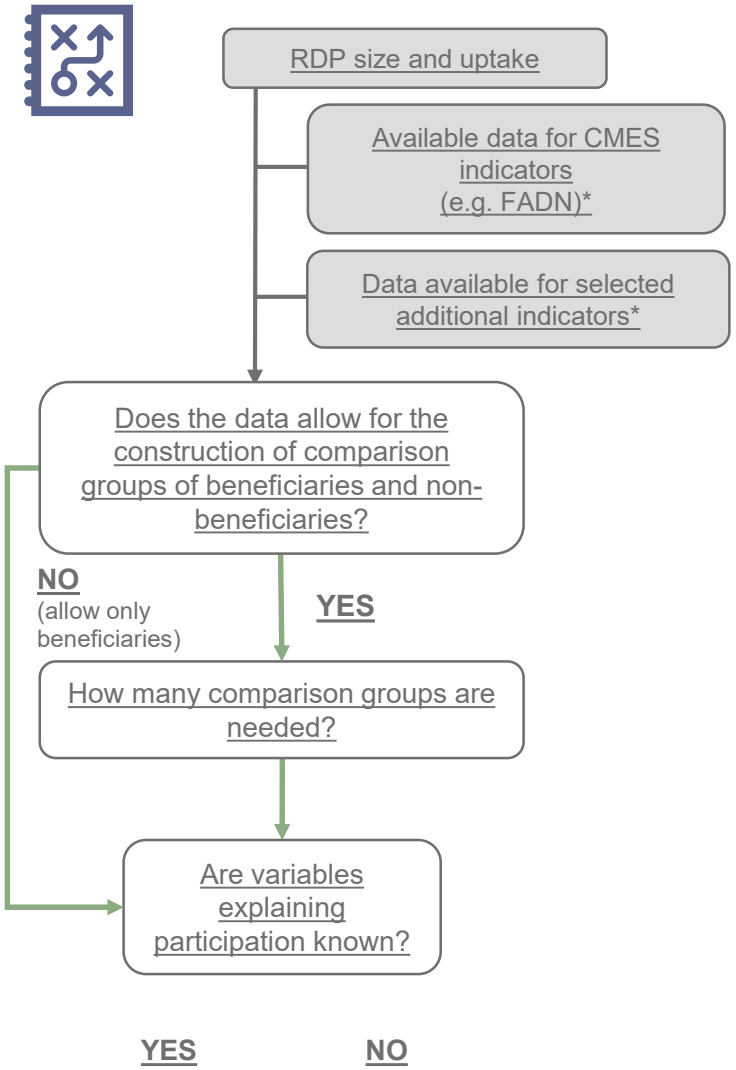
What can be done to improve the data situation?

Short-term solutions (for AIR 2019)

If various RDPs within a country are designed in a similar manner (e.g. similar intervention logic or target group) but statistical information is missing to build a counterfactual situation at the regional level (e.g. too few observations on individual programme beneficiaries), one possible solution to net out programme effects is to combine multiple regionalised RDPs. To do so, firstly, one needs to identify at the farm level a set of common control variables among different programmes/territories. Secondly, an additional variable (e.g. a dummy variable) can be built into the list of covariates to identify the location of a farm in different RDPs within a country, enabling the separation of effects of specific regional programmes or areas under consideration.

Long-term solutions (for ex-post)

- Increase the sample (number of observations) in FADN and farm bookkeeping statistics and/or rely more heavily on regular surveys better adjusted to the needs of the RDP evaluations.
- Ensure that all beneficiaries of the RDP programmes are fully integrated into existing farm accountancy or bookkeeping systems with their anonymous records being easily retrievable by each respective national authority (e.g. FADN or farm bookkeeping offices) for evaluation purposes.



Are variables explaining participation known?

Why is this question important?

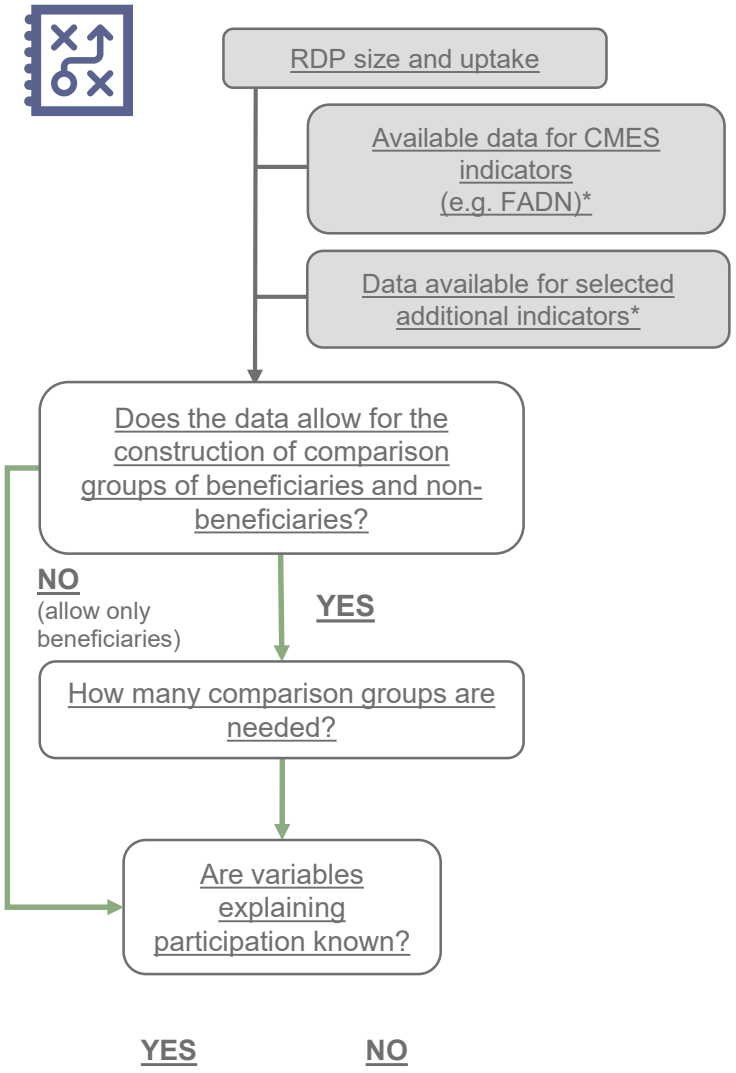
What are the conditions in order to answer the question with YES?

Are there any specificities to be considered for regionalised RDPs?

What can be done to improve the data situation?

- Short-term solutions (for AIR 2019)
- Long-term solutions (for ex-post)

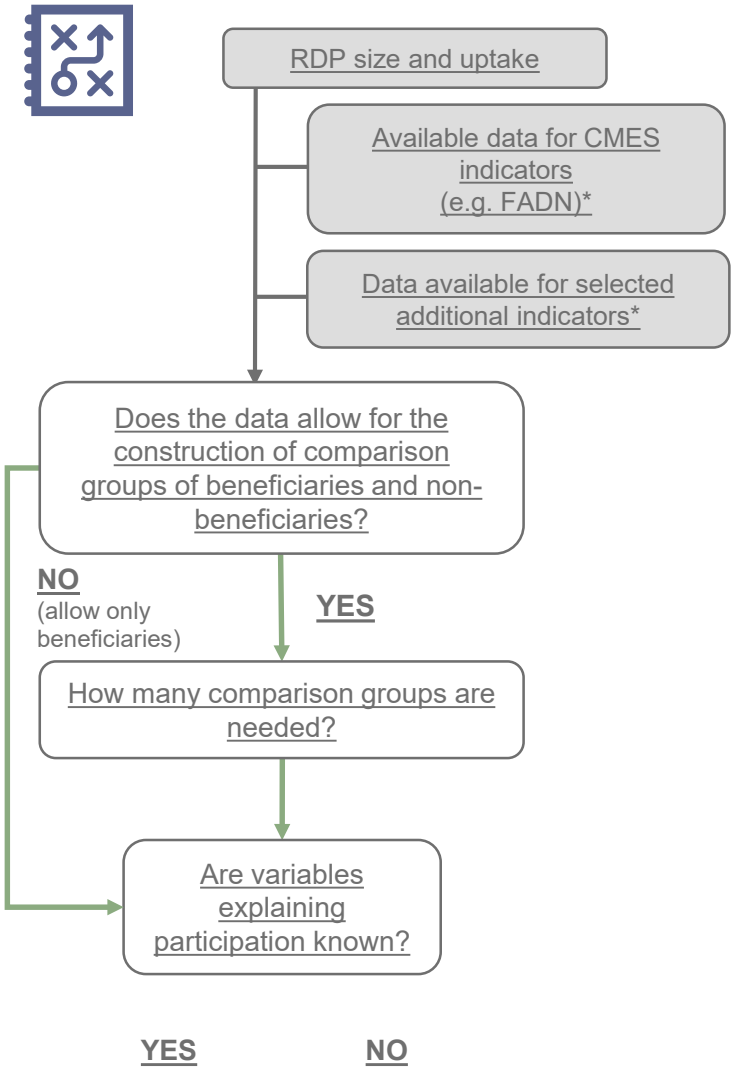
Guidelines *Assessing RDP achievements and impact in 2019*, PART II, Chapter 2.1 and 2.2 and PART IV, Chapter 4.1.
Guidelines *Assessment of RDP results*, Chapter 2.1 and 6.2, and Annex 11, Chapter 2.4.
Guidelines for the ex post evaluation of 2007-2013 RDPs, Chapter 4.3



Are variables explaining participation known?

Why is this question important?

In order to establish similarities between programme beneficiaries and non-beneficiaries a model which controls for the effects of other important confounding factors has to be constructed (e.g. PSM-DiD). Control variables (model covariates) which determine both the farm's participation in the RDP programme as well as the programme's outcomes must be fed into this model.



Are variables explaining participation known?

What are the conditions in order to answer the question with YES?

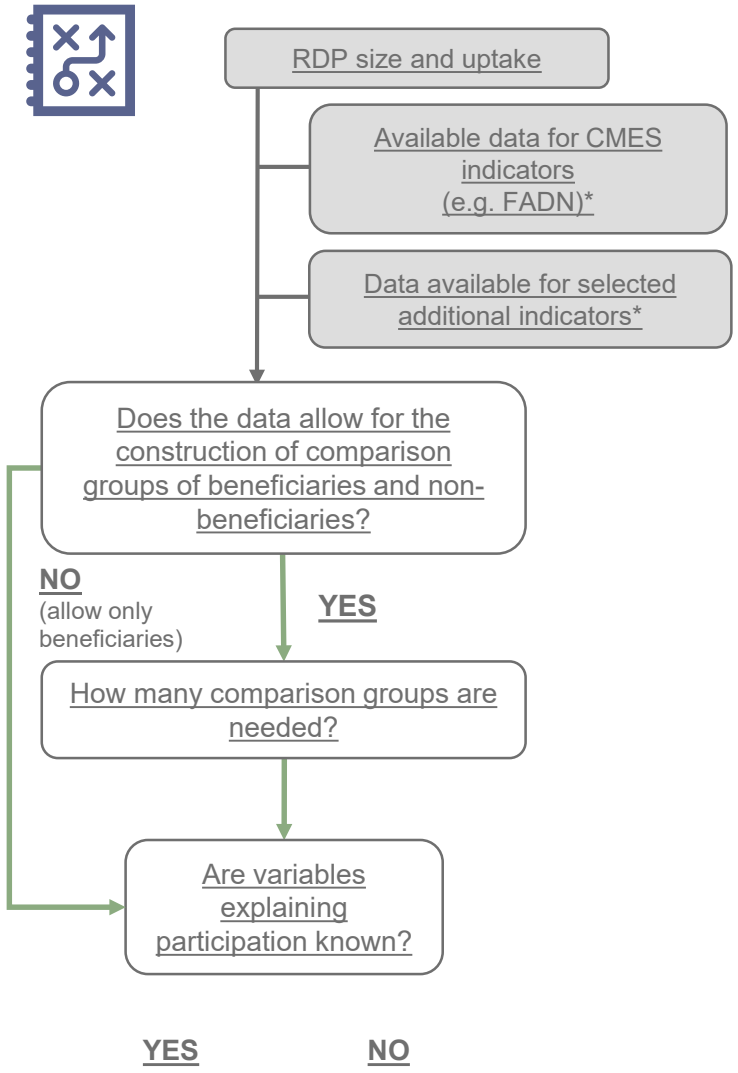
The database should contain data and information concerning the characteristics and economic performance of various types of farms over extended periods of time. This will allow for the selection of the most important control variables (model covariates) at the level of individual agricultural holdings facilitating the impact analysis at the micro level. FADN or farms bookkeeping provides data for variables which can be used as controls in PSM-DiD models.

Note

For the selection of relevant model covariates only those variables which are unaffected by the RDP programme should be included (i.e. variables which are fixed over time or which have been measured prior to participation in programme).

Example

In case there is no access to a database from which observable variables explaining programme participation and outcomes can be selected, controlling for important confounding factors is not possible. In this situation, a naïve comparison of changes in impact indicators in the group of programme beneficiaries vs. non-beneficiaries (i.e. without establishing any similarities between these groups) could be carried out. This can lead to significant selection bias and therefore should be avoided.

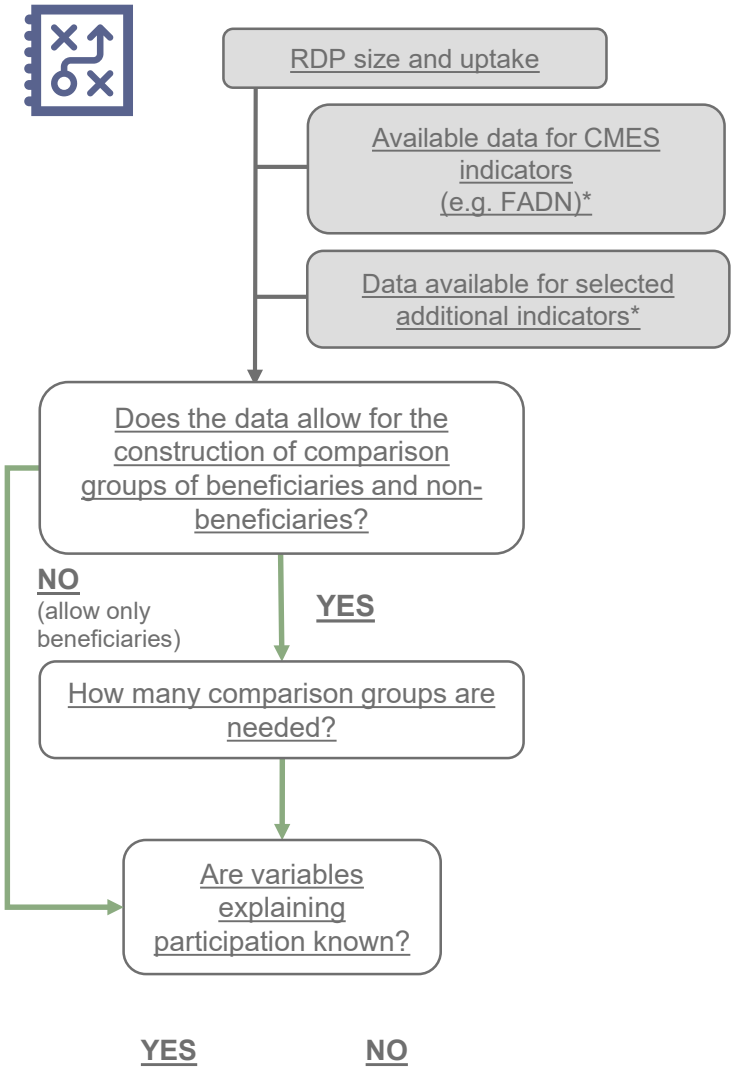


Are variables explaining participation known?

What are the conditions in order to answer the question with YES?

Note(s)

FADN does not cover all the agricultural holdings in the EU, but only those which are ranked as commercial holdings due to the economic size. If individual RDP measures specifically target non commercial agricultural holdings the FADN sample will not be representative. If this is the case, well designed surveys have to be carried out in order to collect the information needed (e.g. Impact and in-depth study of the Rural Development Grant Programme in Kosovo).

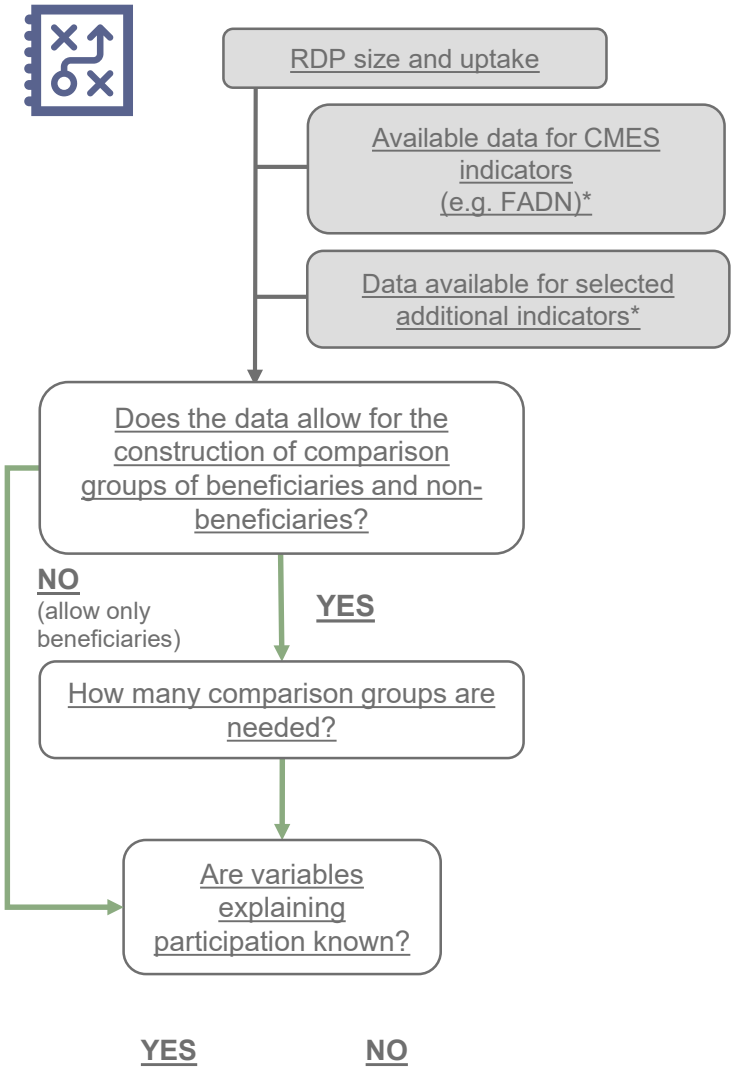


Are variables explaining participation known?

What are the conditions in order to answer the question with YES?

Example(s)

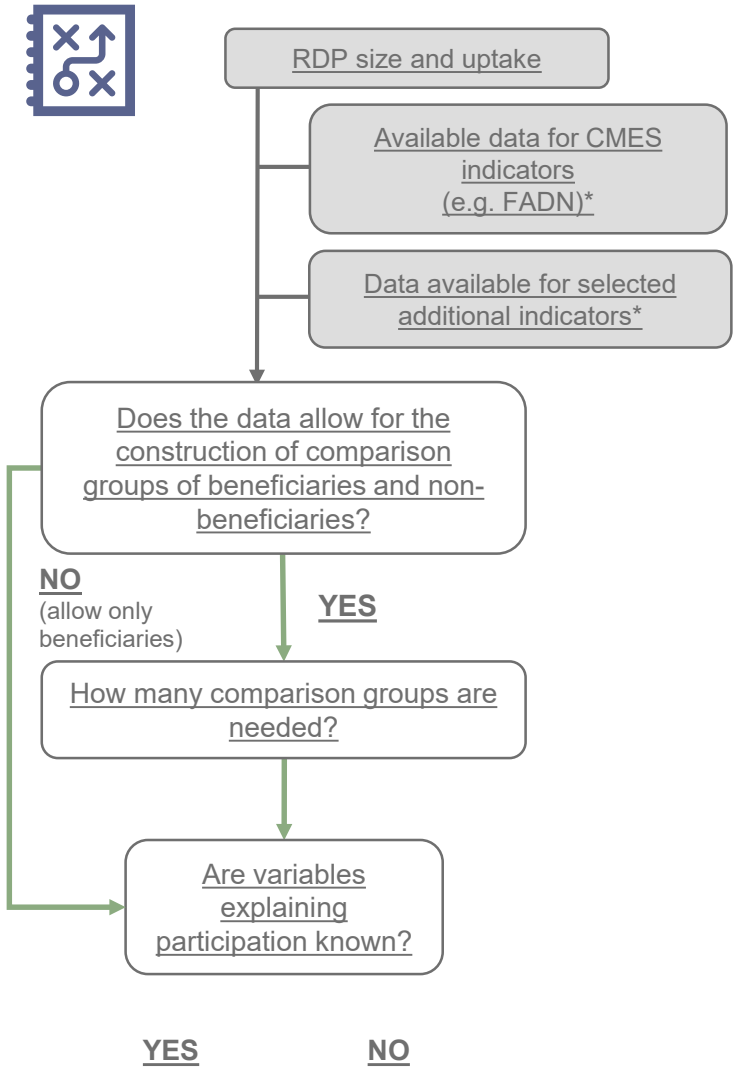
A comprehensive discussion of various criteria/methods to be considered while choosing control variables for the PSM-DiD models (e.g. Hit or Miss method, Statistical Significance, Leave One out Cross Validation, etc.) can be found in: Caliendo and Kopeing (2005), 'Some Practical Guidance for the Implementation of Propensity Score Matching'.



Are variables explaining participation known?

Are there any specificities to be considered for regionalised RDPs?

If the number of RDP beneficiary farms identified in FADN or bookkeeping data is very low, which prevents the construction of comparison groups, then the selection of variables explaining programme participation is not relevant. If this is the situation, other activities should be carried out to improve the data availability (see next slide).



Are variables explaining participation known?

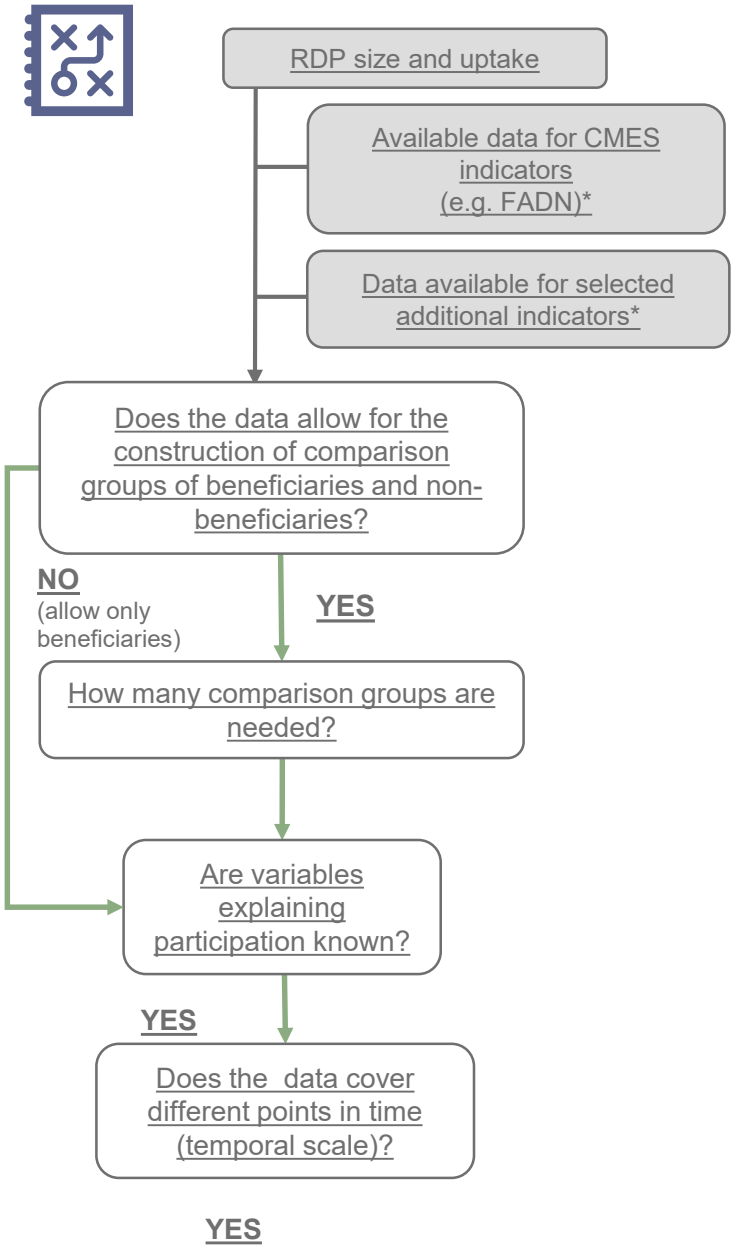
What can be done to improve the data situation?

Short-term solutions (for AIR 2019)

If various RDPs within a country are designed in a similar manner (e.g. similar intervention logic or target group) but statistical information is missing to build a counterfactual at the regional level (e.g. too few observations on individual programme beneficiaries), one possible solution to net out programme effects is to combine multiple regionalised RDPs. To do so, firstly, one needs to identify at the farm level a set of common control variables among different programmes/territories. Secondly, an additional variable (e.g. dummy variable) can be built into the list of covariates to identify the location of a farm in different RDPs within a country and enable the separation of different effects of each specific regional programme or areas under consideration.

Long-term solutions (for ex-post)

- Increase the sample (number of observations) in FADN and farm bookkeeping statistics and/or rely more heavily on regular surveys better adjusted to meet the needs of RDP evaluations. The suitability of FADN data for RDP evaluations can also be enhanced if different types of RDP support can be disaggregated and/or independently listed (e.g. differentiation of various agro-environmental or investment type measures).
- Ensure that all beneficiaries of the RDP programmes are fully integrated into the farm accountancy or bookkeeping systems with their anonymous records being easily retrievable by each respective national authority (e.g. FADN or farm bookkeeping offices) for evaluations.



Does the data cover different points in time (temporal scale)?

Why is this question important?

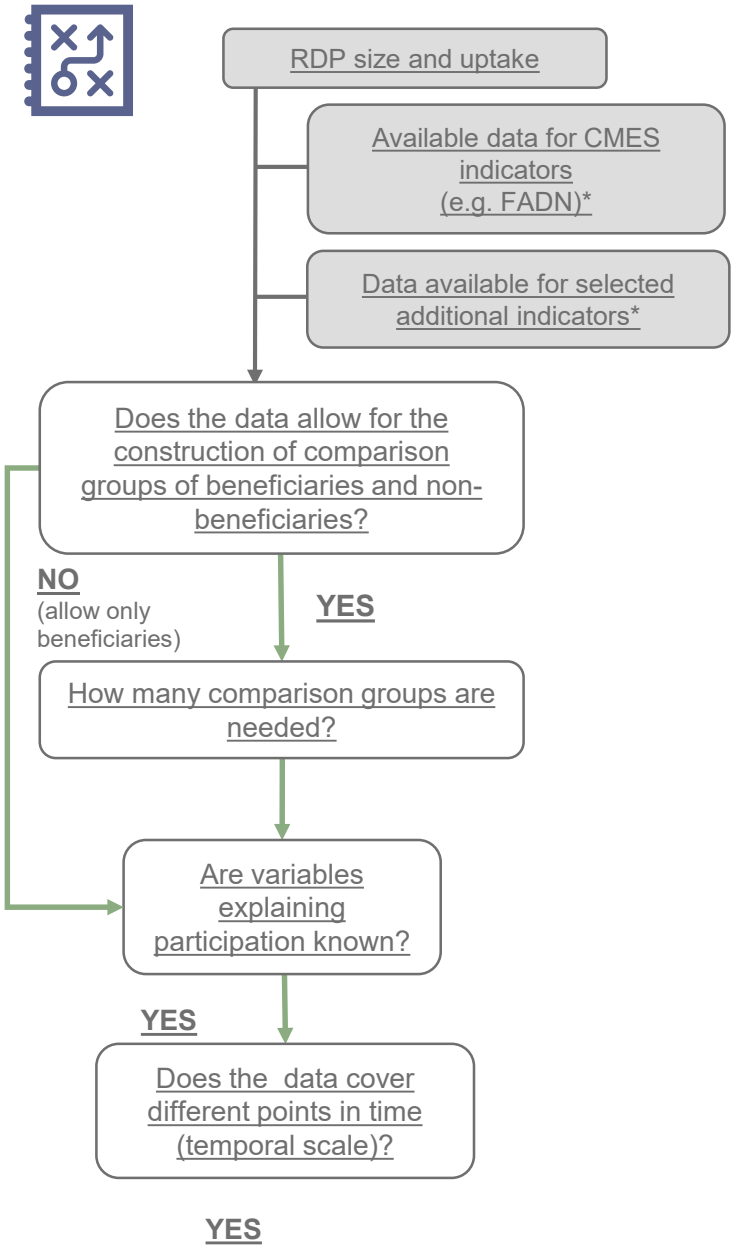
What are the conditions in order to answer the question with YES?

Are there any specificities to be considered for regionalised RDPs?

What can be done to improve the data situation?

- Short-term solutions (for AIR 2019)
- Long-term solutions (for ex-post)

Guidelines *Assessing RDP achievements and impact in 2019*, PART II, Chapter 2.1 and 2.2 and PART IV, Chapter 4.1.
Guidelines *Assessment of RDP results*, Chapter 2.1 and 6.2, and Annex 11, Chapter 2.4.
Guidelines for the ex post evaluation of 2007-2013 RDPs, Chapter 4.3

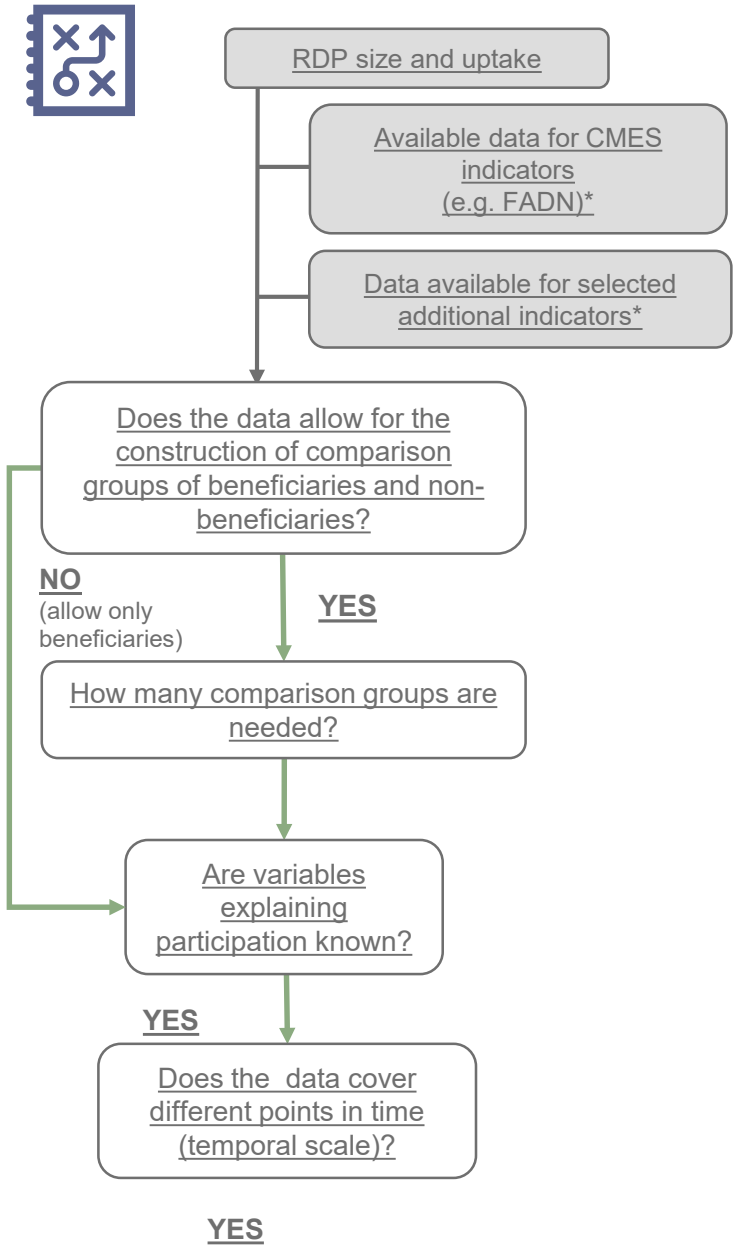


Does the data cover different points in time (temporal scale)?

Why is this question important?

Data should be collected at the farm level and be able to be used to show the development of individual impact indicators over time (at least before the programme implementation and at the time of evaluation during and after the programming period).

Note



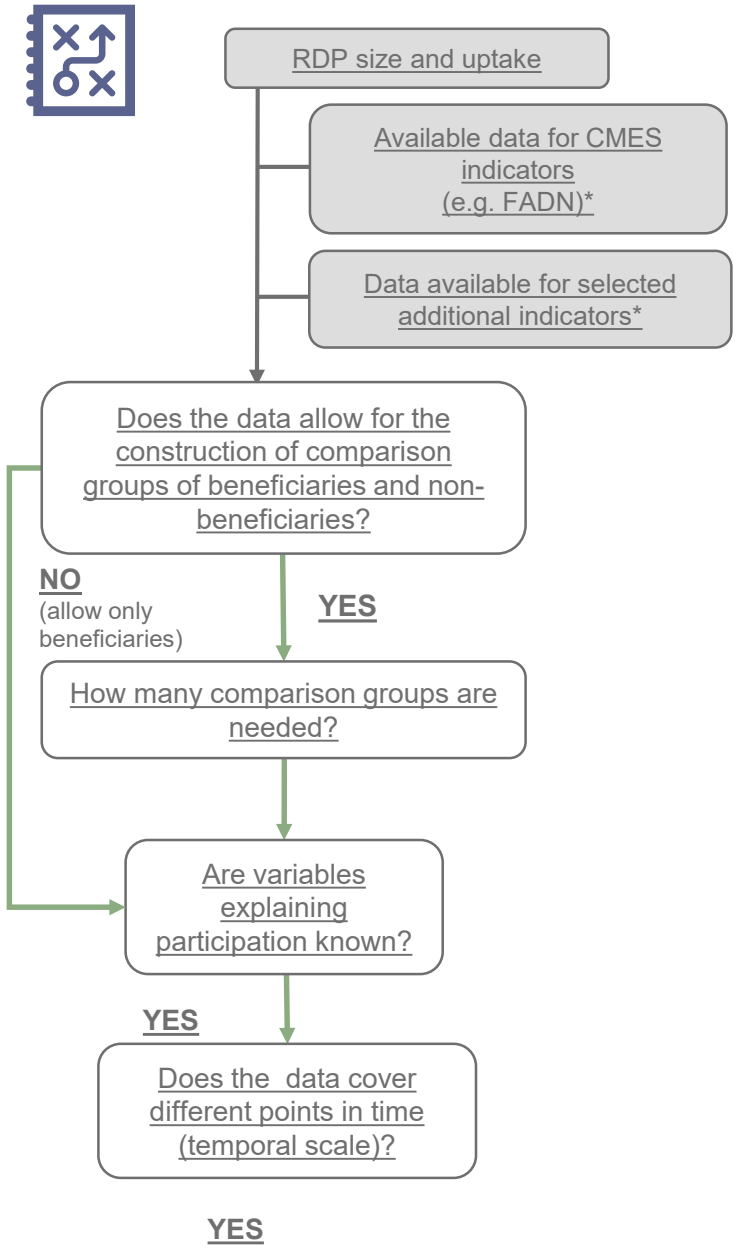
Does the data allow for the construction of comparison groups?

Why is this question important? | **Note(s)**

From an econometric point of view panel data is the most advantageous for impact evaluations of RDPs (i.e. data which contains observations of selected control variables and impact indicators for the same agricultural farms over multiple time periods (prior, during and after 2014-2020)). Panel data has several advantages over cross-sectional data (data for various individual units in a given time) or time-series data (data on one individual unit in different periods of time).

The advantages of using panel data are:

1. The ability to make more accurate inferences of model parameters. Panel data usually contains more degrees of freedom and more sample variability than cross-sectional data.
2. Greater capacity for capturing the complexity of farm development and economic performance than a single cross-section or time series data.
3. Allows for more precise controlling of the impacts of omitted variables;
4. Superior ability to generate more accurate predictions for individual outcomes.
5. Simplifies the computation and inference analysis.



Does the data cover different points in time (temporal scale)?

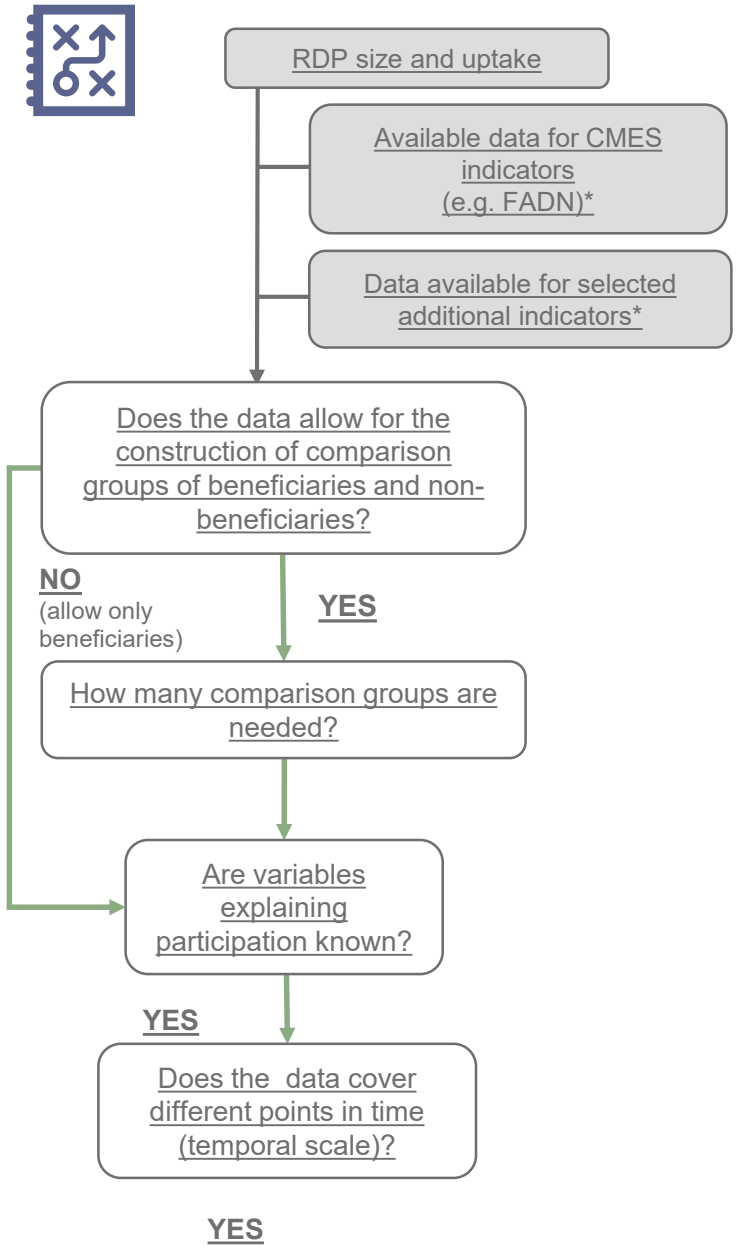
What are the conditions in order to answer the question with YES?

The database should collect data at different points in time. FADN and bookkeeping datasets are available in panel format, which provides detailed information on selected variables collected for hundreds of the same farms in each year over a certain time period (i.e. rotating panel).

Note

Are there any specificities to be considered for regionalised RDPs?

The data situation may be even more difficult (i.e. the number of observations can be low and the time series coverage can be more restrictive). A number of steps can be carried out in order to improve the data situation (see next slide).

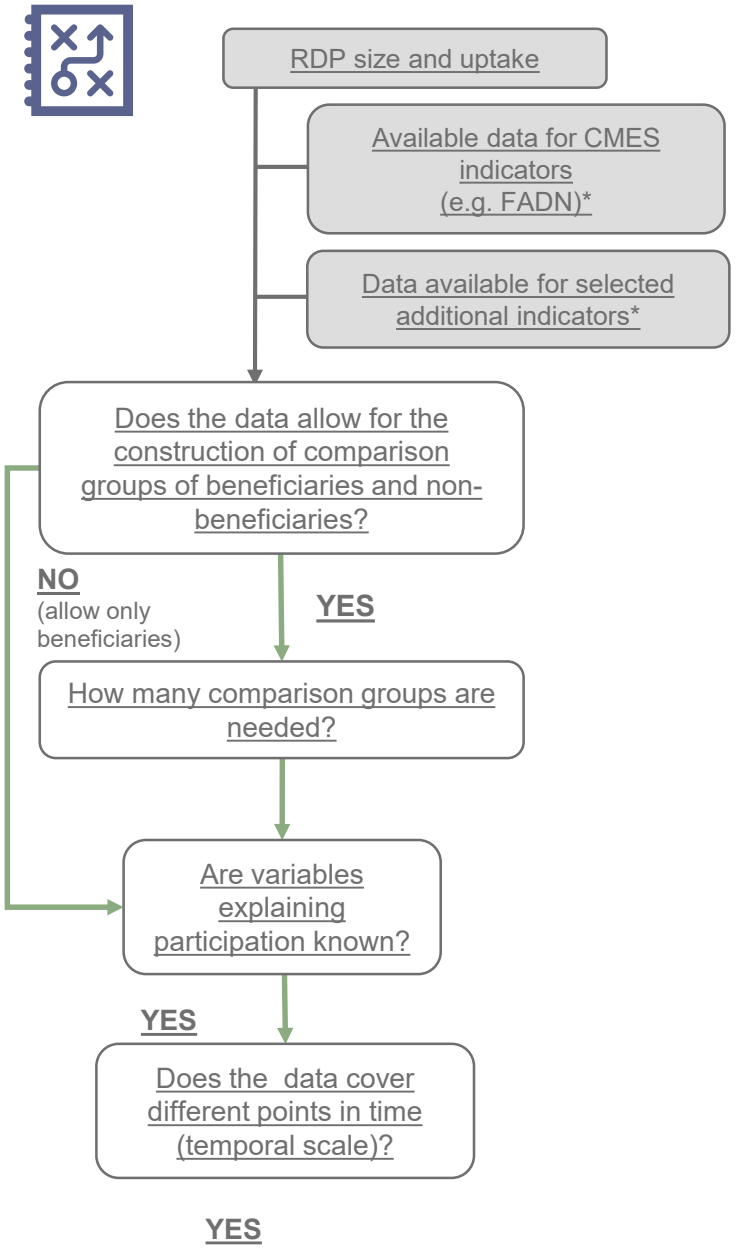


Does the data cover different points in time (temporal scale)?

What are the conditions in order to answer the question with YES?

Note(s)

Due to yearly farm sample rotations for the FADN database (up to 7-10% of the total sample) not all farms for which data exists in a given year can be entered into a panel constructed for the purpose of the RDP 2014-2020 evaluation. The evaluator should therefore verify for each EU country and each regionalised RDP assessed the availability of farm panel data for the period of time needed for the evaluation (e.g. prior to the programme and for last years of its implementation). If there is a missing observation for the same holding for the beginning or ending years, the panel data can be built on the bases of an available preceding year (a second-best procedure).



Does the data cover different points in time (temporal scale)?

What can be done to improve the data situation?

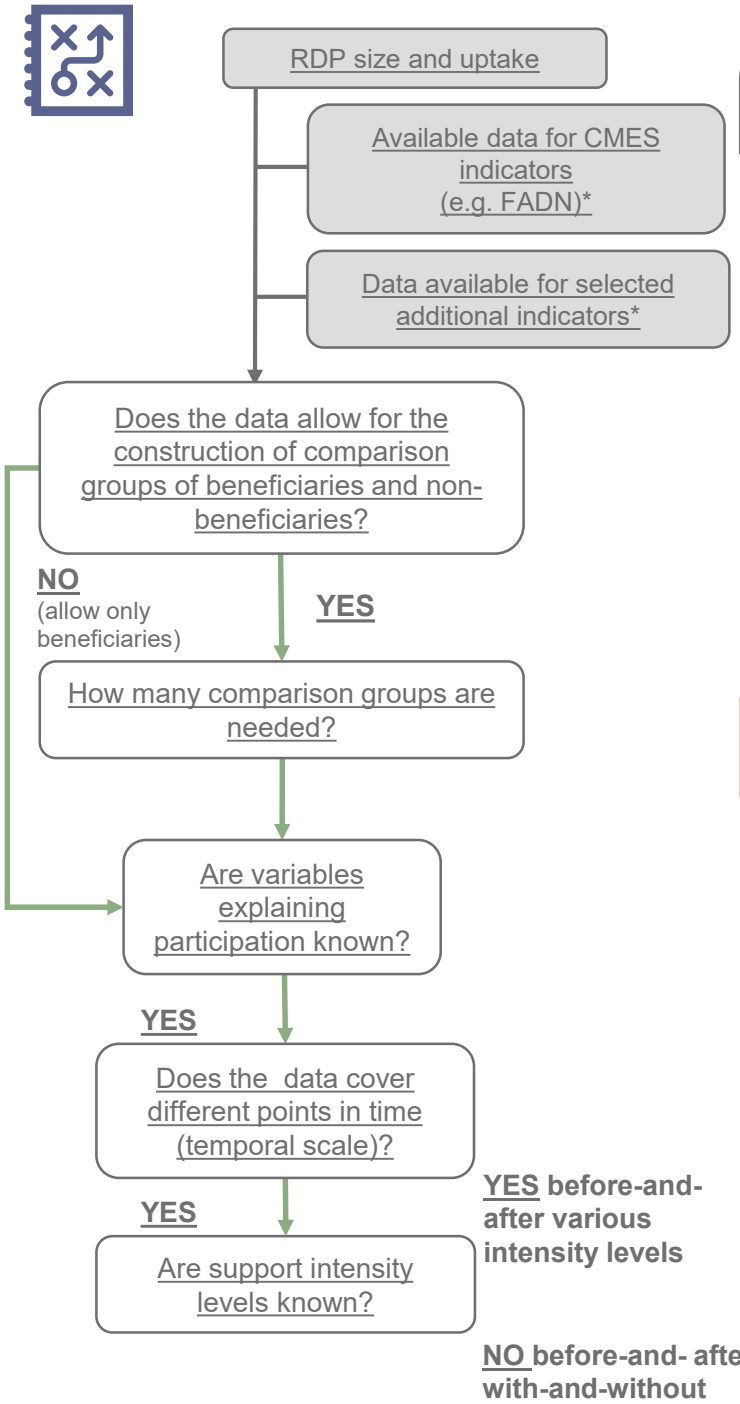
Short-term solutions (for AIR 2019)

For regionalised RDP programmes, if the data is insufficient for a number of RDP beneficiary farms identified in the FADN or bookkeeping data and the RDPs do not differ substantially from one another, then one can combine FADN datasets covering more than one RDP (FADN regions).

For details concerning regionalised RDPs see section: *Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?*

Long-term solutions (for ex-post)

- Increase the sample (number of observations) in FADN and farm bookkeeping statistics and/or rely more heavily on regular surveys better adjusted to the needs of RDP evaluations. The suitability of FADN data for RDP evaluations can also be enhanced if different types of RDP support can be further disaggregated and/or separately listed (e.g. differentiation of various agro-environmental or investment types of measures).
- Ensure that all beneficiaries of RDPs are fully integrated into the farm accountancy or bookkeeping systems with their anonymous records being easily retrievable by each respective national authority (e.g. FADN or farm bookkeeping offices) for evaluation purposes.



Are support intensity levels known?

Why is this question important?

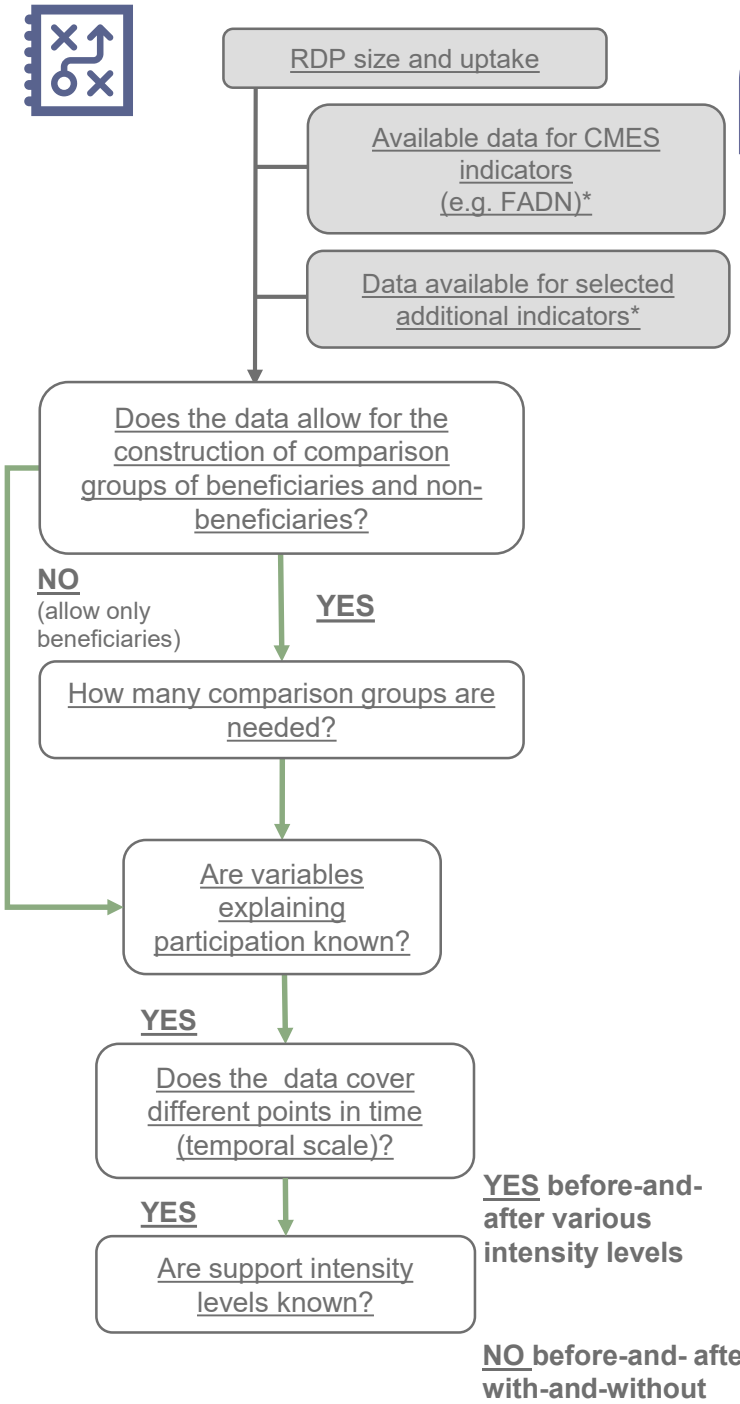
What are the conditions in order to answer the question with YES?

Are there any specificities to be considered for regionalised RDPs?

What can be done to improve the data situation?

- Short-term solutions (for AIR 2019)
- Long-term solutions (for ex-post)

Guidelines for the ex post evaluation of 2007-2013 RDPs, Chapter 4.3.3.2

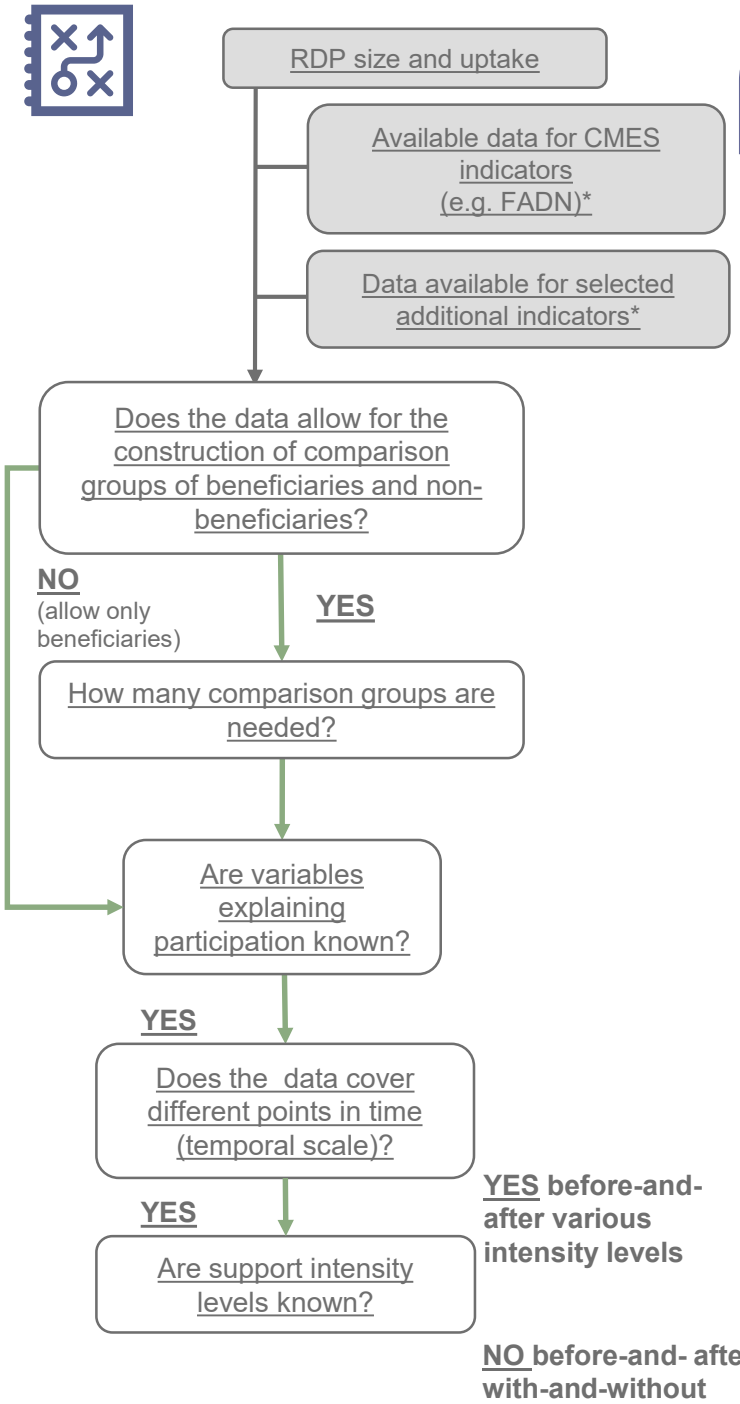


Are support intensity levels known?

Why is this question important?

Information about the intensity level of programme support is essential for choosing an evaluation approach of the RDP's effects in a situation when all units are classified as programme beneficiaries and/or an evaluator is interested in assessing the marginal effectiveness of RDP funds provided to the agricultural sector.

If information pertaining to the intensity of investment support (e.g. financial flows going to individual farm beneficiaries for each of the RDP's measures in the period 2014-2020) is known, programme impacts can be analysed using GPSM, which enables the calculation of dose-response functions and derivative dose-response functions.



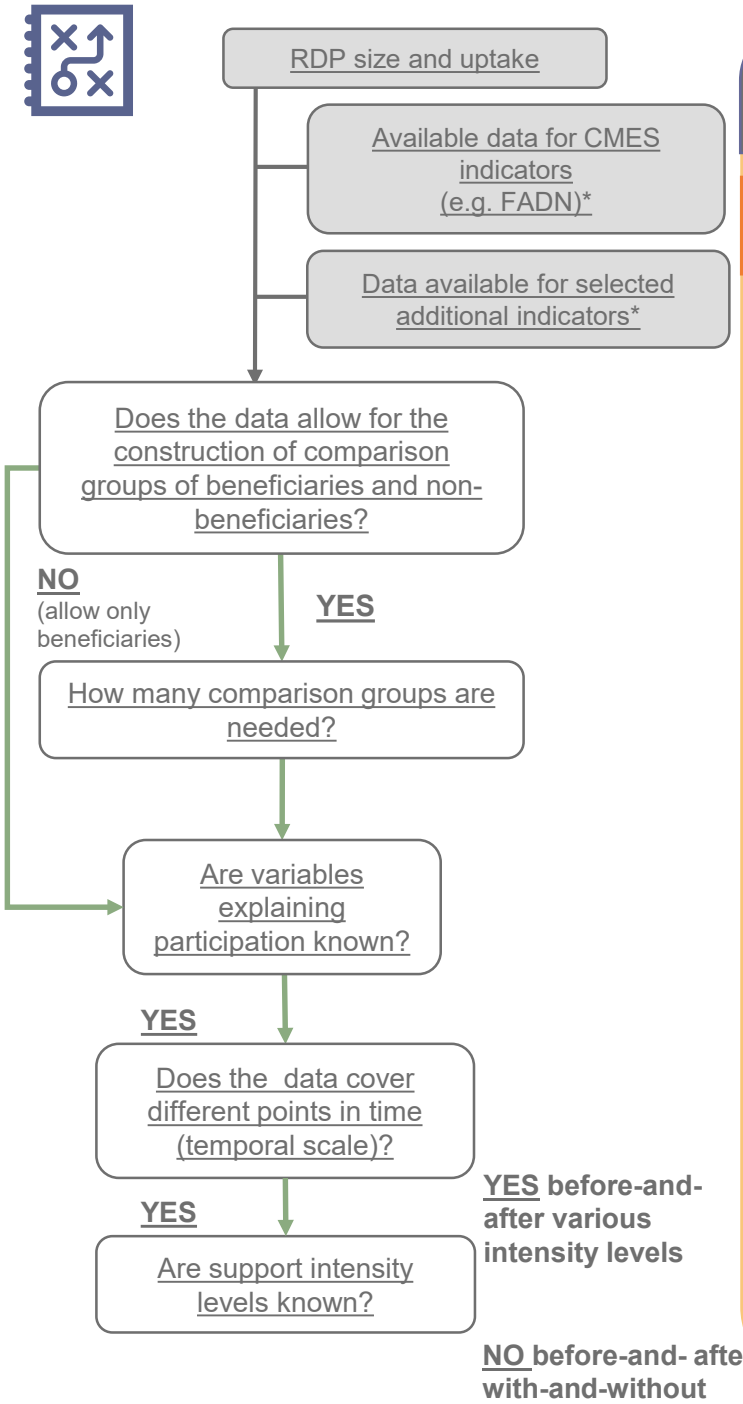
Are support intensity levels known?

What are the conditions in order to answer the question with YES?

Data concerning the level of RDP support for each agriculture holding supported by RDPs 2014-2020 must be delivered to the evaluator from the respective Paying Agency of the country where the evaluation is taking place. Data on the intensity of programme support for each RDP beneficiary in the years 2014-2020 (for all RDP measures or group of measures) is not available in the FADN dataset.

Are there any specificities to be considered for regionalised RDPs?

There is no specificities for regionalised RDPs as the relevant data is to be obtained from each respective Paying Agency, both for the whole country (in case of 1 RDP) as well as for each regionalised RDP.



Are support intensity levels known?

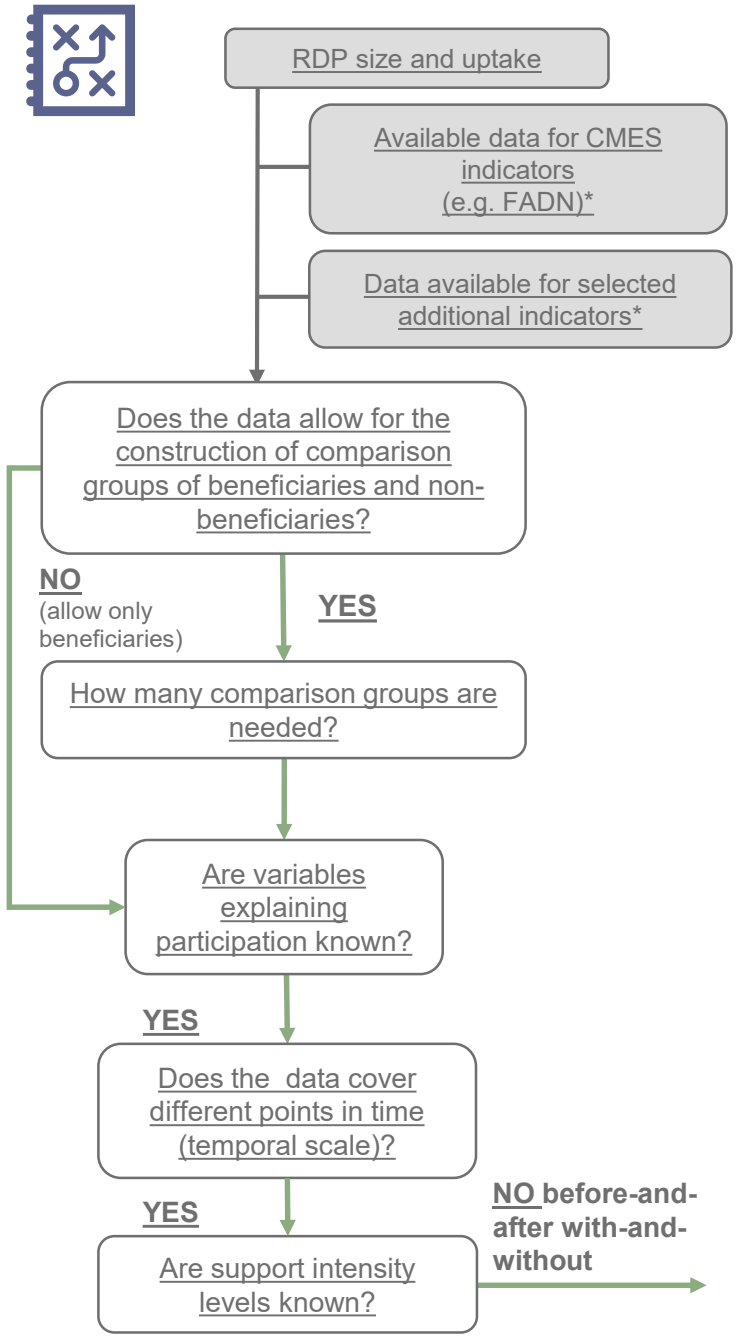
What can be done to improve the data situation?

Short-term solutions (for AIR 2019)

In order to improve the data situation, the Managing Authority and evaluator should inform the Paying Agency of the evaluation activities planned at an early stage.

Long-term solutions (for ex-post)

- Closer cooperation between evaluators and Paying Agencies should be pursued from the beginning of the programme evaluation.
- Ensure that all beneficiaries of RDPs are fully integrated into the farm accountancy or bookkeeping systems with their anonymous records being easily retrievable by each respective national authority (e.g. FADN or farm bookkeeping offices) for evaluation purposes.



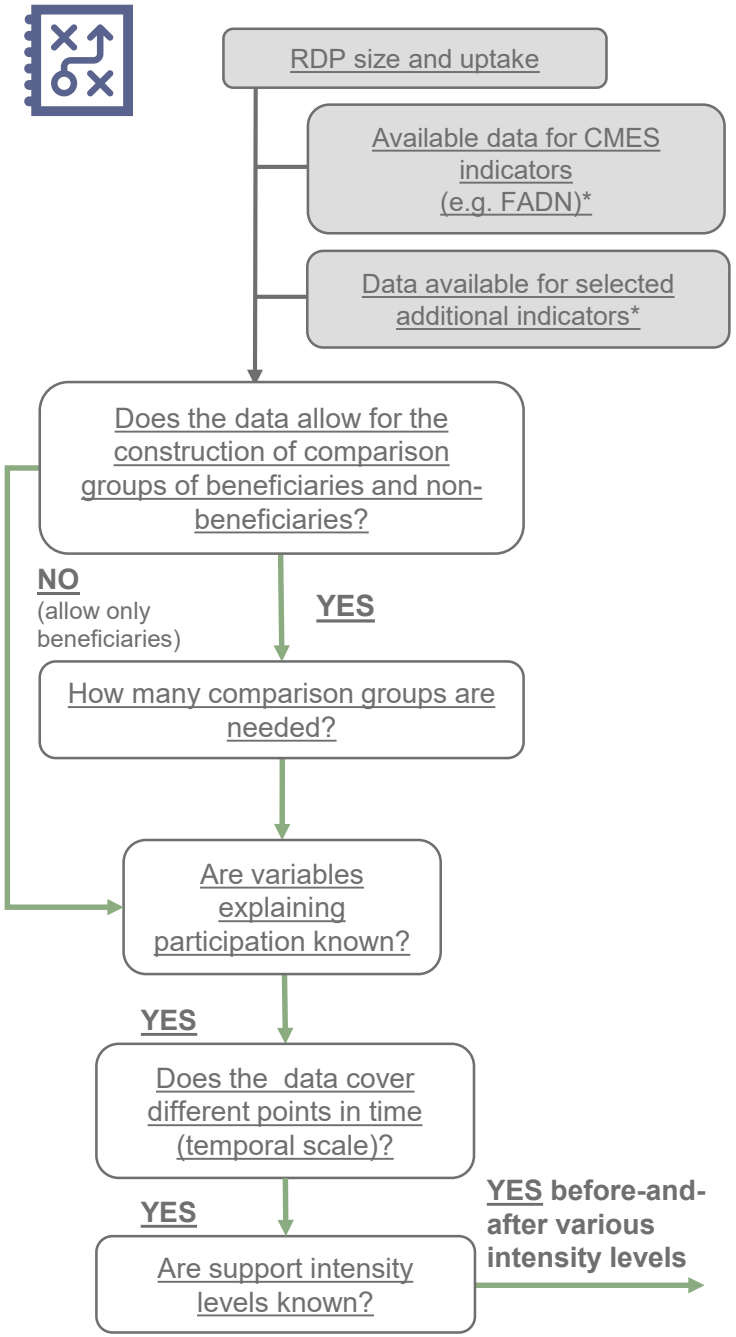
Propensity Score Matching with Difference in Differences method (PSM – DiD)

At the micro level, the assessment approach in case of good data availability at the level of the agricultural holding is based on the comparison of established similar control groups (beneficiaries and non-beneficiaries) prior to the beginning of the programme and at the time of the evaluation. The main objective is to net out the RDP's effects on the competitiveness of the agricultural sector, through the use of advanced econometric methods (e.g. PSM combined with DiD).

PSM-DiD is a highly applicable estimator when the outcome data of programme participants and non-beneficiaries is available for both the 'before' and 'after' periods. PSM-DiD measures the effect of the RDP by using the differences between comparable programme beneficiaries and non-beneficiaries in the before/after periods. Observed changes overtime for the matched (using PSM) programme non-beneficiaries are assumed to be appropriate counterfactuals for programme beneficiaries. A decisive advantage of the PSM-DiD estimator, compared to a conventional DiD estimator, is that this method allows for the better control of selection bias in both observables and unobservables.

At the macro level, the net effects of the RDP on the competitiveness of the agricultural sector are obtained by up-scaling the results from the micro-level assessment to the RDP area (macro-level).

Read more in guidelines *Assessing RDP achievements and impact in 2019*, PART II, Chapter 2.2.3 and PART IV, Chapter 4.1.2.

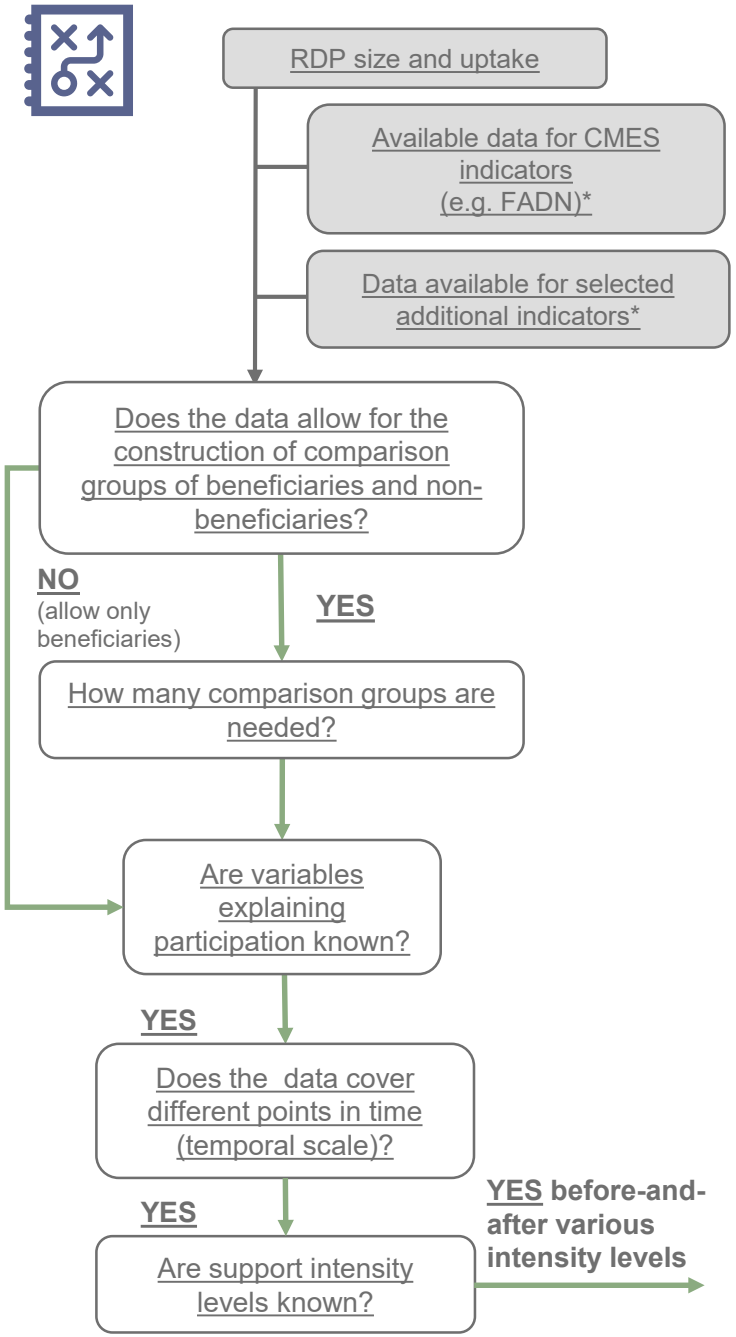


Generalised Propensity Score Matching (GPSM)

The GPSM method is used when all units are programme beneficiaries. If the evaluator has explicit information on the intensity of the programme support (e.g. financial flows from a given programme per farm) and programme effects (results/impacts) then the data can be analysed by means of a dose-response function and derivative dose-response function.

The GPSM allows one to estimate the average effect of public investment support on the selected result/impact indicator, as well as assess the marginal effects of the programmes or measures depending on the intensity of support. Disaggregated programme evaluation results cannot be obtained by employing traditional techniques (e.g. the binary propensity score matching methodology, regression discontinuity design, or any other techniques utilised in standard evaluation studies). GPSM can also be used as an extension of a binary PSM method and can be used to eliminate any bias associated with differences in the covariates included in the evaluation model.

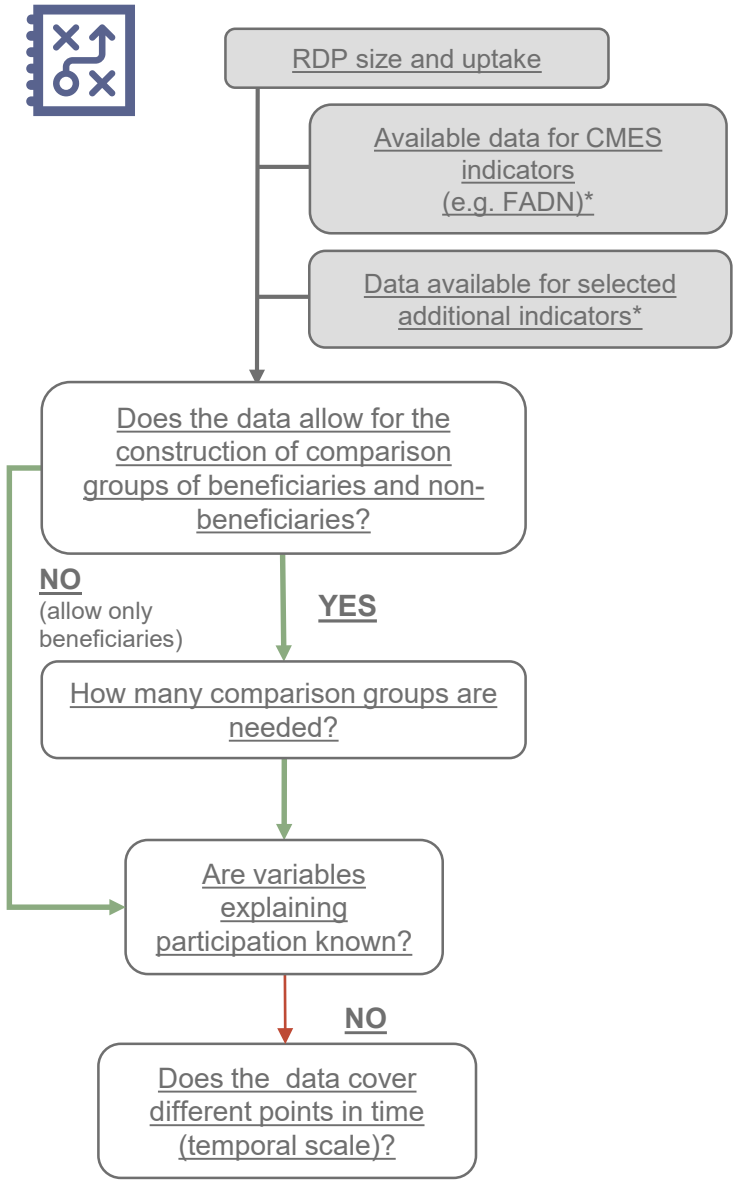
Examples



GPSM

Example(s)

- 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;
- Imai, K. and van Dyk, D.A. (2004),"Causal inference with general treatment regimes: Generalizing the propensity score" au *Journal of the American Statistical Association*, 2004, 99, pp. 854-866;
- Bia, M. and Mattei, A. (2007),"Application of the Generalised Propensity Score. Evaluation of public contributions to Piedmont enterprises, al *P.O.L.I.S. department's Working Papers 80*, Department of Public Policy and Public Choice - POLIS, 2007;
- Michalek J., Ciaian P. and Kancs, d'A. (2014), "Capitalisation of CAP Single Payment Scheme into Land Value: Generalised Propensity Score Evidence from the EU", *Land Economics*, May 2014, 90:260-289.;
- Kluve, J. et.al., (2012), "Evaluating continuous training programs using the generalised propensity score, Journal of the Royal Statistical Society: Series A (Statistics in Society) Volume 175, Issue 2, pages 587–617, April 2012;
- 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



Does the data cover different points in time (temporal scale)?

Why is this question important?

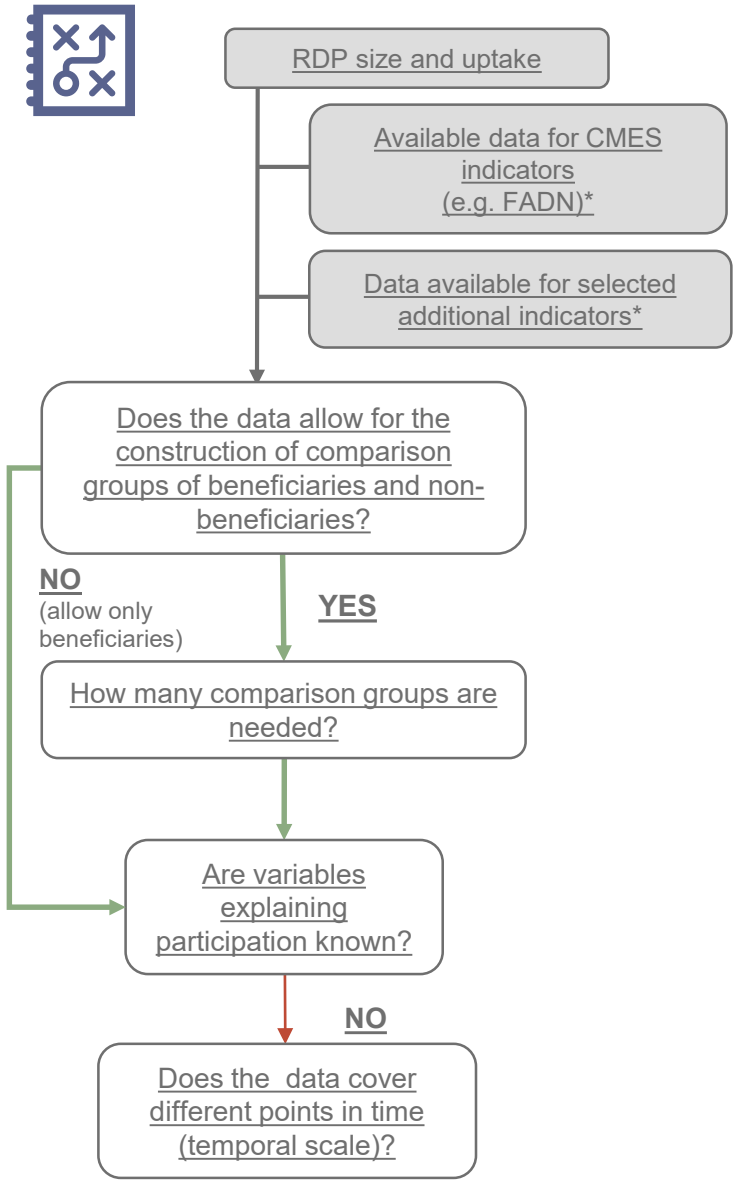
What are the conditions in order to answer the question?

Are there any specificities to be considered for regionalised RDPs?

What can be done to improve the data situation?

- Short-term solutions (for AIR 2019)
- Long-term solutions (for ex-post)

Guidelines *Assessing RDP achievements and impact in 2019*, PART II, Chapter 2.1 and 2.2 and PART IV, Chapter 4.1.
Guidelines *Assessment of RDP results*, Chapter 2.1 and 6.2, and Annex 11, Chapter 2.4.
Guidelines for the ex post evaluation of 2007-2013 RDPs, Chapter 4.3.

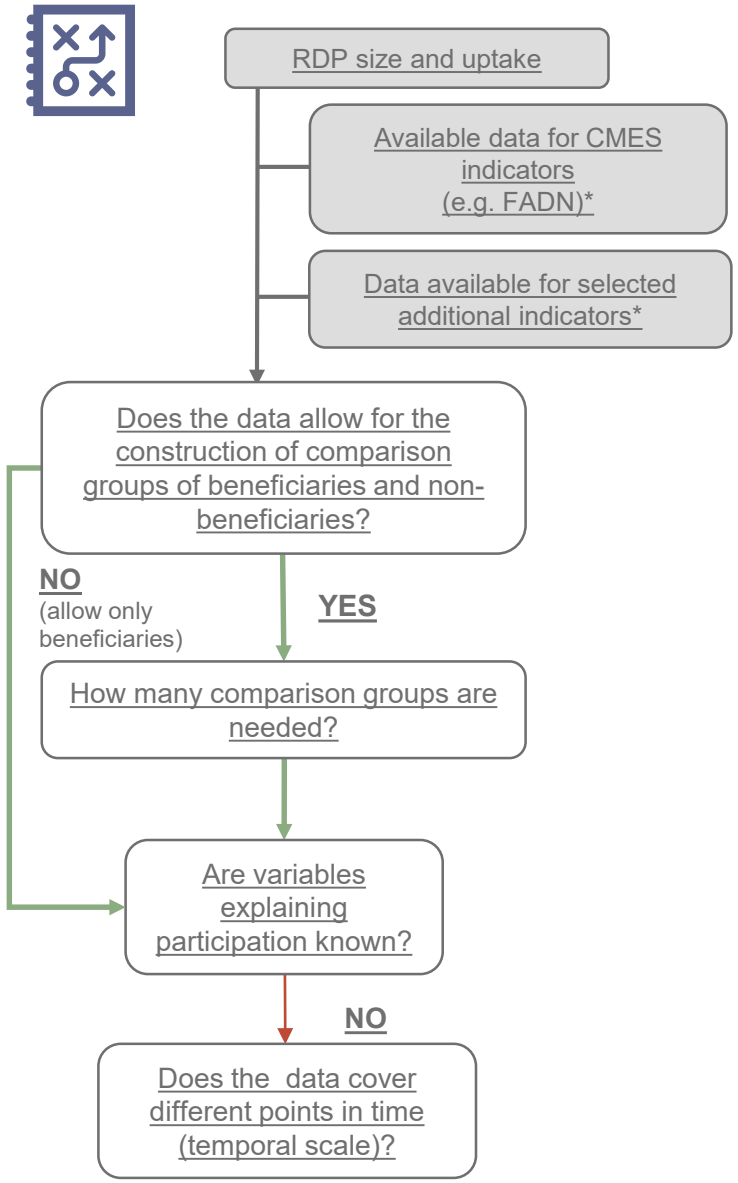


Does the data cover different points in time (temporal scale)?

Why is this question important?

Information should be collected at the farm level in order to show the development of individual impact indicators over time (at a minimum before the programmes implementation and at the time of the evaluation during and after the programming period).

Note



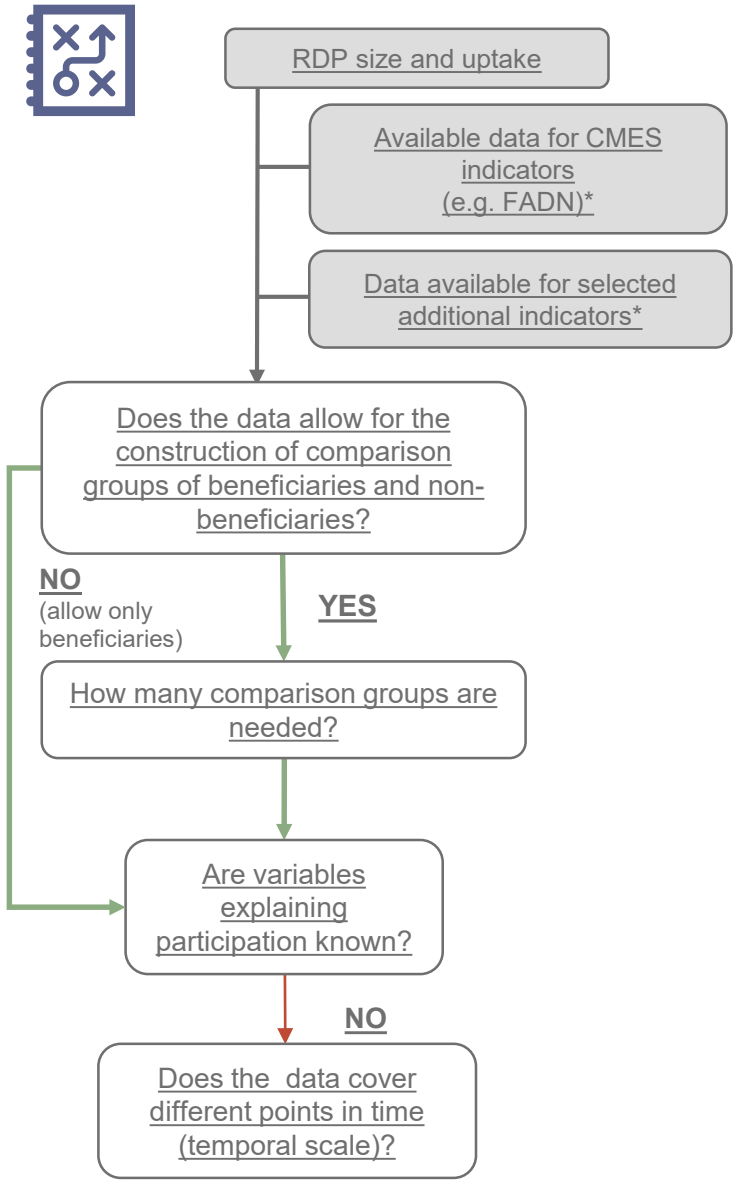
Does the data cover different points in time (temporal scale)?

Why is this question important? | **Note(s)**

From an econometric point of view panel data is the most advantageous for impact evaluations of RDPs (i.e. data which contains observations of selected control variables and impact indicators for the same agricultural farms over multiple time periods (prior, during and after 2014-2020)). Panel data has several advantages over cross-sectional data (data for various individual units in a given time) or time-series data (data on one individual unit in different periods of time).

The advantages of using panel data are:

1. The ability to make more accurate inferences of model parameters. Panel data usually contains more degrees of freedom and more sample variability than cross-sectional data.
2. Greater capacity for capturing the complexity of farm development and economic performance than a single cross-section or time series data.
3. Allows for more precise controlling of the impacts of omitted variables;
4. Superior ability to generate more accurate predictions for individual outcomes.
5. Simplifies the computation and inference analysis.



Does the data cover different points in time (temporal scale)?

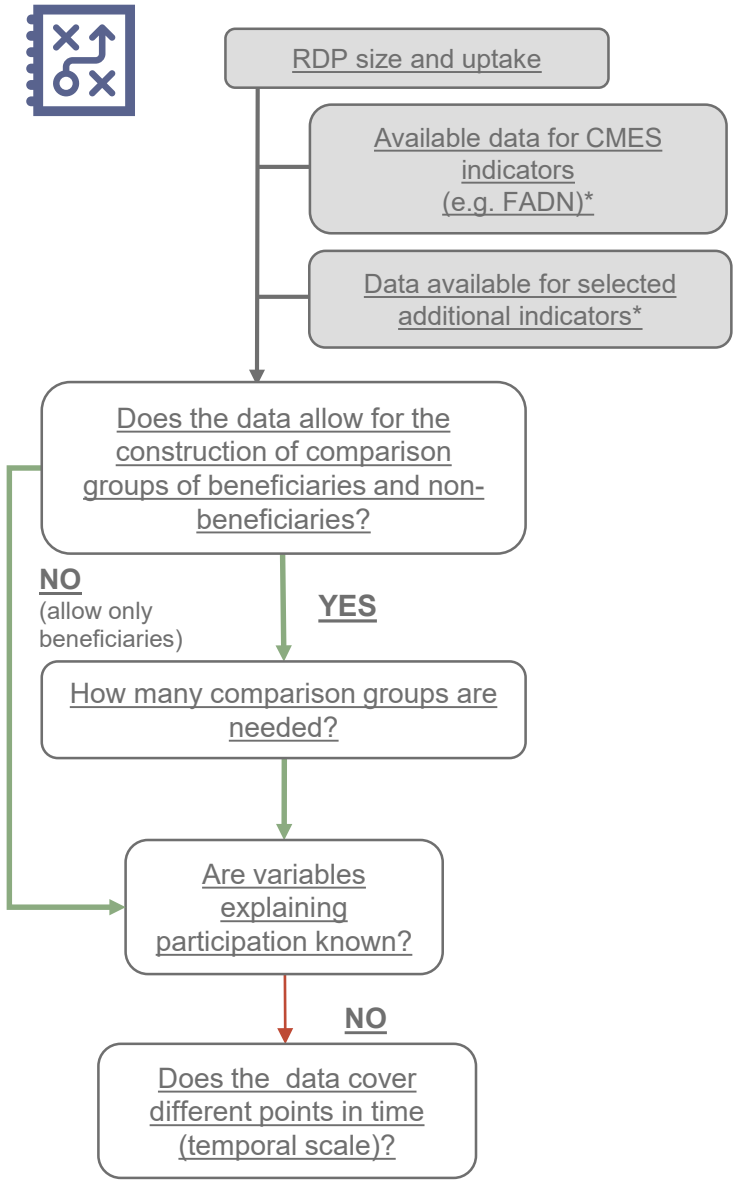
What are the conditions in order to answer the question?

The database should provide data at different points in time, ideally on an annual basis. FADN and bookkeeping datasets are available in panel format (i.e. they provide detailed information on selected variables collected for hundreds of the same farms in each year over a certain time period (i.e. rotating panel).

Due to yearly farm sample rotations for the FADN database (up to 7-10% of the total sample) not all farms for which data exists in a given year can be entered into a panel constructed for the purpose of the RDP 2014-2020 evaluation. The evaluator should therefore verify for each EU country and each regionalised RDP assessed the availability of farm panel data for the period of time needed for the evaluation (e.g. prior to the programme and for last years of its implementation). If there is a missing observation for the same holding for the beginning or ending years, the panel data can be built on the bases of an available preceding year (a second-best procedure).

Are there any specificities to be considered for regionalised RDPs?

The data situation may be even more difficult (i.e. the number of observations can be low and the time series coverage can be more restrictive). A number of steps can be carried out in order to improve the data situation (see next slide).



Does the data cover different points in time (temporal scale)?

What can be done to improve the data situation?

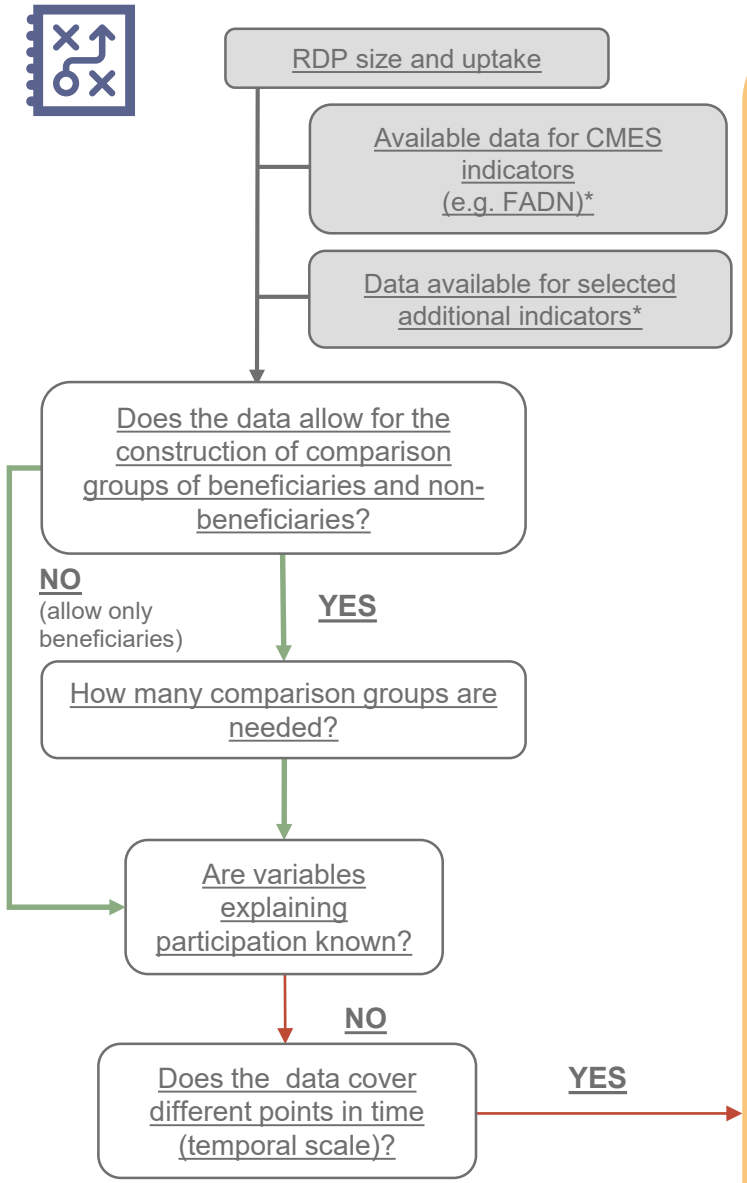
Short-term solutions (for AIR 2019)

For regionalised RDP programmes, if the data is insufficient for a number of RDP beneficiary farms identified in the FADN or bookkeeping data and the RDPs do not differ substantially from one another, then one can combine FADN datasets covering more than one RDP (FADN regions).

For further details on regionalised RDPs see: *Does the data allow for the construction of comparison groups of beneficiaries and non-beneficiaries?*

Long-term solutions (for ex-post)

- Increase the sample (number of observations) in FADN and farm bookkeeping statistics and/or rely more heavily on regular surveys better adjusted to the needs of RDP evaluations. The suitability of FADN data for RDP evaluations can also be enhanced if different types of RDP support can be further disaggregated and/or separately listed (e.g. differentiation of various agro-environmental or investment types of measures).
- Ensure that all beneficiaries of RDPs are fully integrated into the farm accountancy or bookkeeping systems with their anonymous records being easily retrievable by each respective national authority (e.g. FADN or farm bookkeeping offices) for evaluation purposes.

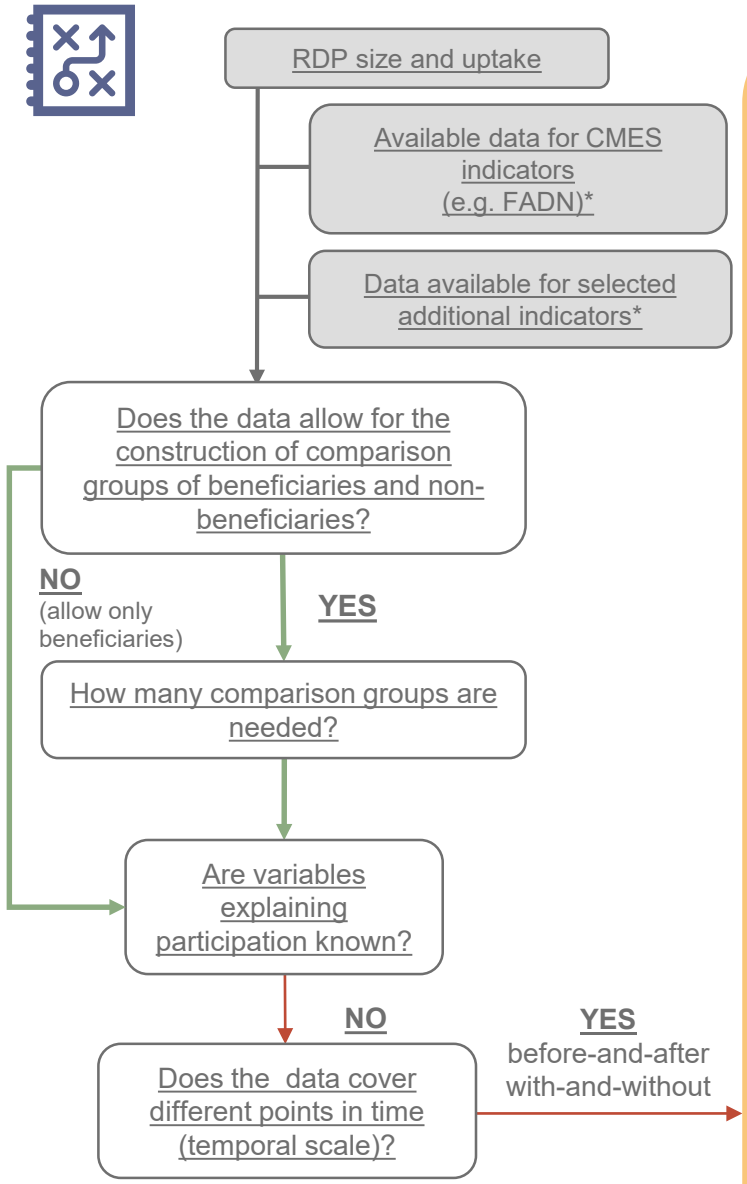


Regression Discontinuity Design

The second best choice to assess the RDP's net effects on the competitiveness of the agricultural sector at the micro-level is Regression Discontinuity Design (RDD). RDD can be used to assess the net effects of programmes that have a continuous eligibility index with a clearly defined cut-off score determining which farms are eligible and which are not. The assessment of the net impacts of the intervention is done by comparing the average outcomes for the beneficiaries just above the cut-off with non-beneficiaries just below it. Under certain comparability conditions, the assignment near the cut-off can be seen almost as random. The RDD method assumes that individual units around the eligibility cut-off point (on both sides) are similar, thus the selection bias should be minimal.

The net effects of the RDP on the competitiveness of the agricultural sector are obtained through up-scaling the findings from the micro-level assessment to the RDP area (macro-level).

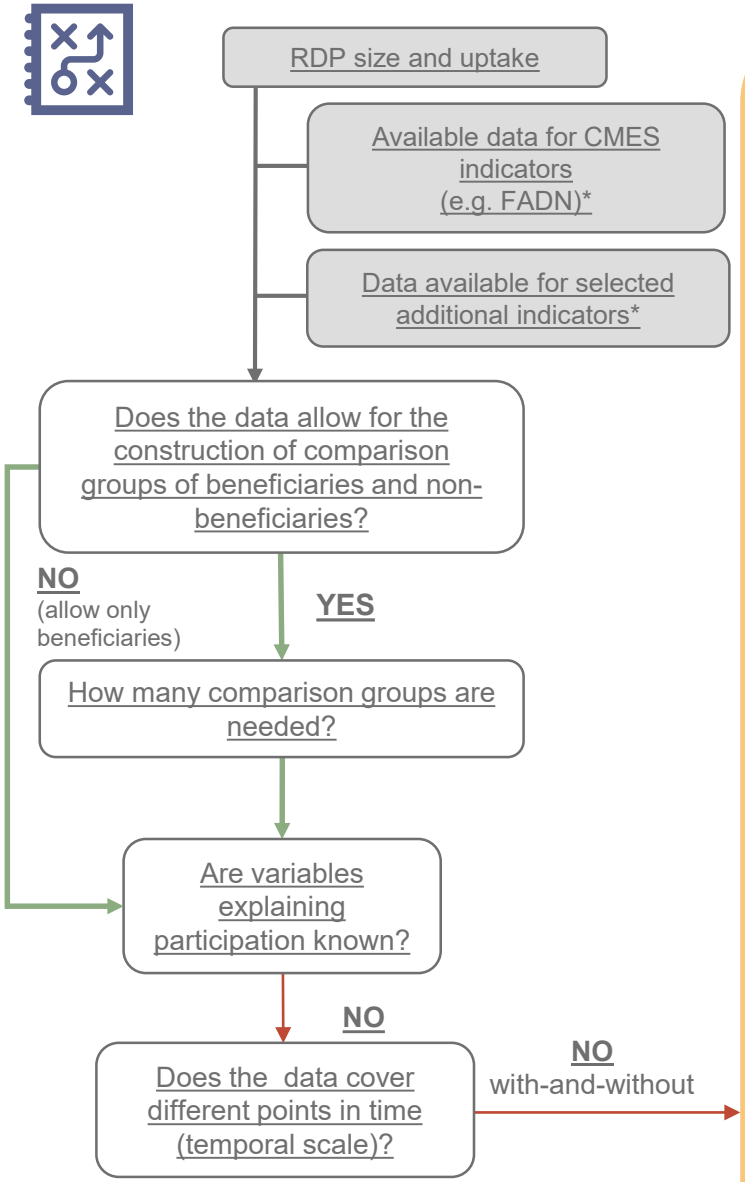
Read more in guidelines *Assessing RDP achievements and impact in 2019*, PART II, Chapter 2.2.4 and PART IV, Chapter 4.1.3.



Other approaches: Difference in Differences

Difference in Differences (DiD) is the evaluation method which compares before/after changes to the situation of beneficiaries with the before/after changes in the situation of the selected non-beneficiaries (control group). For DiD to be valid, the group of programme non-beneficiaries must accurately represent the change of outcomes that programme beneficiaries would have experienced in the absence of the programme. The difference-in-differences approach combines two "naïve" techniques (i.e. before and after comparisons of programme beneficiaries and comparisons between programme beneficiaries with programme non-beneficiaries) to produce a better estimate of the counterfactual. DiD is used for netting out the RDP's effects on the competitiveness of the agricultural sector, usually in combination with other methods (e.g. PSM). It can be used for both the micro and macro level assessment.

Read more in guidelines *Assessing RDP achievements and impact in 2019*, PART II, Chapter 2.2.3 and PART IV, Chapter 4.1.2.

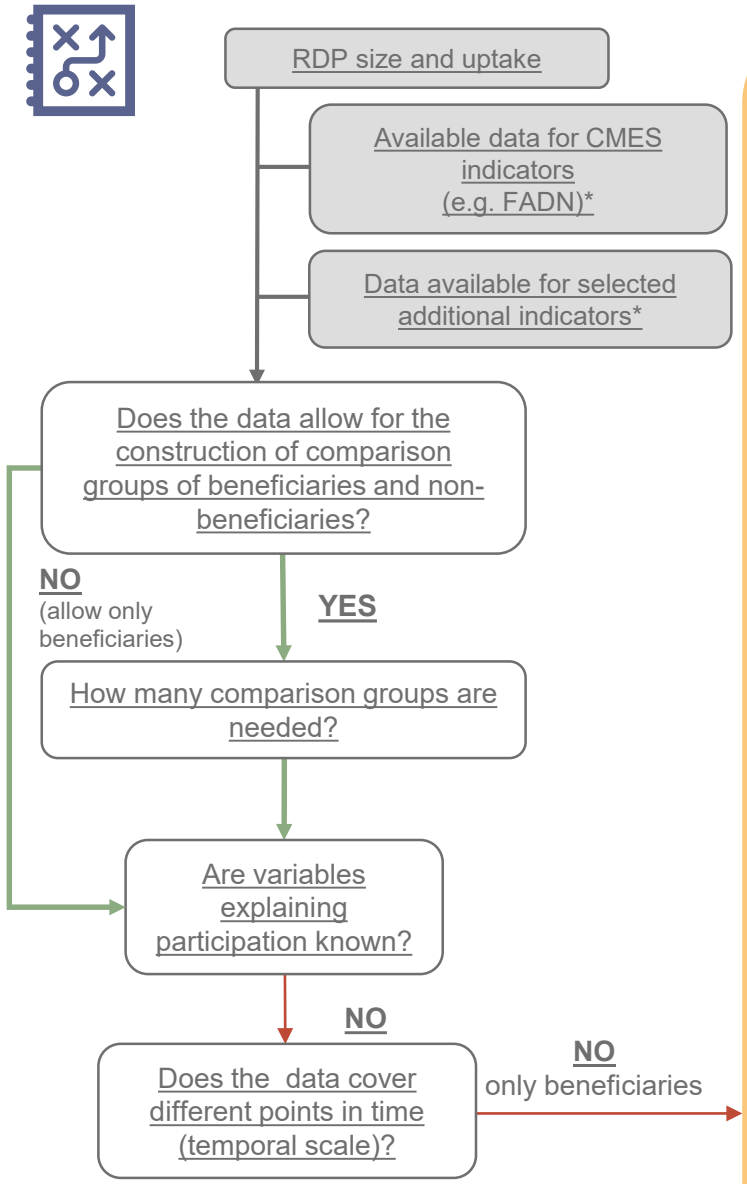


Other approaches: Naïve Group Comparison

If data is insufficient, the evaluator can apply other approaches to net out the RDP's effects on the competitiveness of the agricultural sector, such as naïve group comparisons (comparing the average value of an indicator of the population of RDP beneficiaries with the average value of the entire sector).

The naïve 'with' vs. 'without' approach relies on the assumption that in the absence of the programme, the outcome indicator of programme participants would be the same as for programme non-beneficiaries. Yet, this would only be justifiable if the systematic performance of the programmes participants was identical with the outcome performance of programme non-beneficiaries. Had this not been the case, the selection bias that results from using outcomes of non-beneficiaries as proxies for the outcomes that programme participants would have experienced had they not participated can be very substantial.

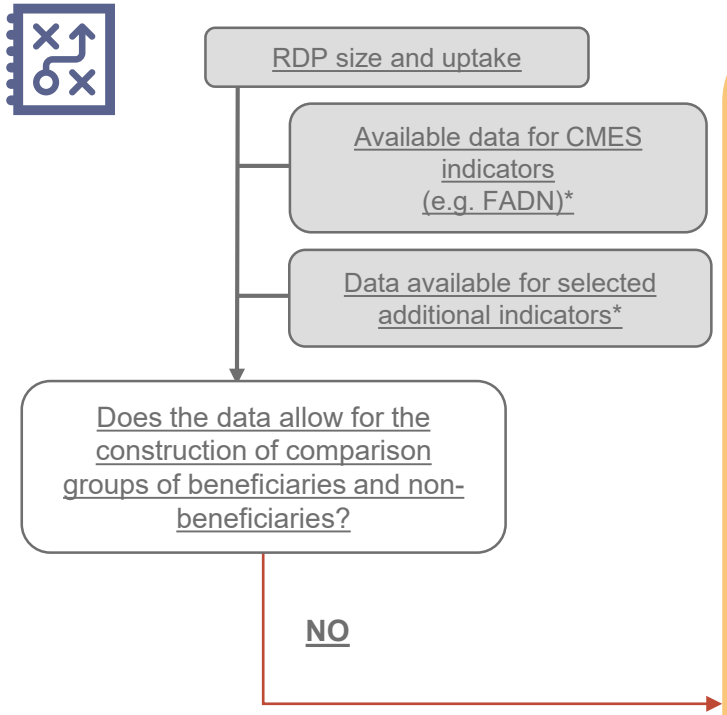
Method can be applied at the micro and macro level of the assessment.



Other approaches: Naïve Baseline Comparison

If there is insufficient data, the evaluator can apply other approaches for netting out the RDP's effects on the competitiveness of the agricultural sector, such as naïve baseline comparisons (comparing the average value of an indicator calculated for the population of RDP beneficiaries with the average value of the entire sector).

In this evaluation technique necessary data on average outcome indicators in the group of 'non-beneficiaries' is usually obtained from various national surveys (or aggregated national data). The approach relies on the assumption that in the absence of the programme the impact indicator of programme participants would be the same as the average of a joint group of programme participants and non-beneficiaries. This however would only be justifiable if systematic performance of the group of programme beneficiaries (measured by any arbitrary impact indicator, e.g. income, profit or employment) was identical with the performance of the joint-group of programme participants and non-beneficiaries (population average).



Other approaches: Naïve Baseline Comparison with Qualitative analysis

If the data is insufficient, the evaluator can apply other approaches for netting out the RDP's effects on the competitiveness of agriculture, such as naïve baseline comparisons and qualitative analysis (MAPP, Delphi etc.). This can be applied at the micro and macro levels of the assessment.

The selection bias that results from using outcomes of unmatched non-beneficiaries (or total farm population) as a proxy for the outcomes that programme participants would have experienced had they not participated can be very substantial.

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