Not Enough or Not Right Data?

The Data used Evaluation of Slovene Rural Development Programme 2007 - 2013

Mojca Hrabar, Oikos, Slovenia
Budapest, 8 – 9 October 2012

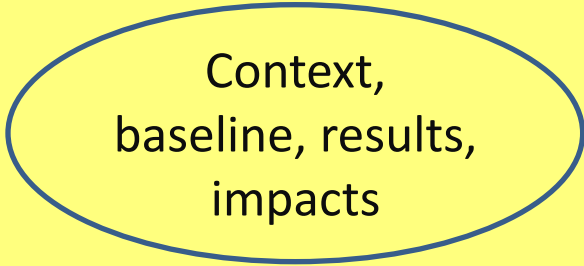# Just a brief introduction…

- RDP 2007 – 2013 for Slovenia: mid-term evaluation (2010) and on-going evaluation from 2011 onwards done jointly by Oikos and PRAC Partneship Co.

- EU accession of Slovenia in 2004
  - ⇒first implementation of entire programming period
  - ⇒first evaluation of this extent

- CMEF: theoretically exciting, but difficult to put in practice

Data Available for the Evaluation

First look: large sets of very detailed data

- Raw data collected by the Agency for Agricultural Markets and Rural Development
  - Applications
  - Reports $\Rightarrow$ obligatory monitoring
  - Contact details

  RDP

  - FADN, Registry of Agricultural Holdings
- Monitoring data of the Ministry of Agriculture and Environment
- Statistical Office
- AAS, Research institutions

Context, baseline, results, impacts

Data Available for the Evaluation
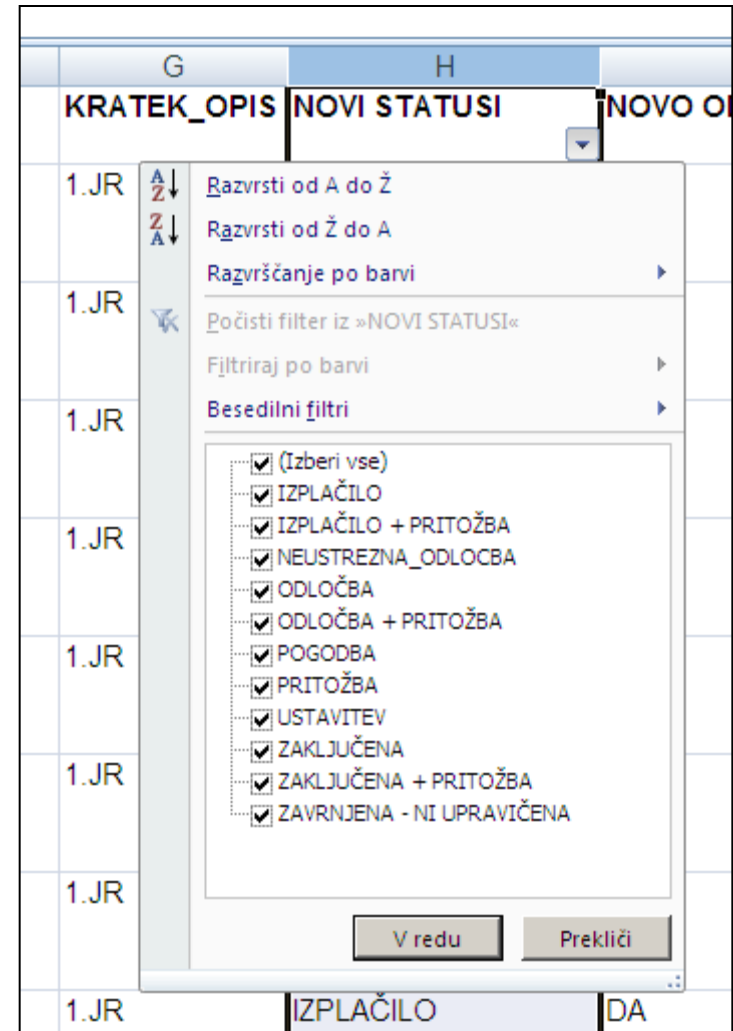
Second look: a lot of data, but difficult to understand and largely incompatible

RDP data management and "coding"

- No single system $\Rightarrow$ inconsistencies accross measures
- Subset of databases for each measure
  - High dependence on contractors for programming
  - No interoperability
  - Limited access
- No handbook $\Rightarrow$ difficult to understand for an "outsider"

Difficult data?

- No data: database bug or lack of data in the beneficiary's application/report?

- "manual" processing of data from subsets of databases: how much is lost/connected wrongly in the process?

	A	B	C
1	Name	Address	Post
2	Janez Novak	Dolga vas 15	1234 Gornje Selo

	E	F	G	H	I	J
1	Name	Surname	Address	House no.	Postcode	Town
2	Janez	Novak	Dolga vas	15	1234	Gornje Selo

- How many beneficiaries are there?

- Comparison of RDP data with other databases (e.g. Registry) almost impossible

Other information available

- **Statistical Office:** data useful for context, comparison
  - mostly impossible to have separate data on rural regions $\Rightarrow$ assessment of net effects?
  - Some data on agriculture collected in larger intervals $\Rightarrow$ lagging behind evaluation needs
- **Agency for Public Legal Records and Related Services**
  - Economic data
- **Agricultural Advisory Service:**
  - Data mostly gathered on ad hoc basis and therefore inconsistent
- **Research institutions**
  - Monitoring of some impact indicators: HNV, water quality
  - Data useful for context, comparison, but: sometimes very narrow, specific focus

Choice of methods: data quality

- **GIS analysis of axis 2**: high quality of data
  - Output and result indicators
  - Comparison with HNV, Farmland Bird Index
  - Effects on habitats, water quality: combination with data from Environment Agency, Institute of the Republic of Slovenia for Nature Conservation, research and other projects
- **Financial, economic output indicators** (e.g. GVA)
  - Available from reports of beneficiaries
  - Question of consistency of application on the side of beneficiaries
- **Statistical indicators and RDP** indicators are sometimes similar, but differ: e.g. Tourist infrastructure in rural areas

Choice of methods: data availability

- **Data from Statistical Office:** aggregated – distribution of individual data is not allowed
  - E.g. impossible to match economic with geographical data $\Rightarrow$ impossible to analyse spatial distribution of trends, relation to natural factors
- **Quality schemes**: no systematic data on quality schemes
- **Innovation:** qualitative assessment through interviews of a sample of beneficiaries
  - No data available
  - Difficult to define innovation, new technologies,…

Choice of methods: data availability

- **Data are costly**: this has so far not been reflected in financial allocations for evaluation
  - Data processing costs of Statistical Office, Agency for Public Legal Records and Related Services, Ordinance Survey,…
  - Costs of preparation of received data (matching, processing) for use in the evaluation team, especially without handbooks/guidance
- Valuable information: **contact details!**

Evaluation of result indicators

- In general, **data were readily available** for evaluation
  - Calculation on the basis of RDP data from the Paying Agency and the Ministry
  - In some cases, combination of RDP data and data from other sources
- in some cases, the meaning/ understanding of the indicator is questionable
  - e.g. no. of agricultural holdings that introduced new technologies or/and products: do all beneficiaries report in the same way?

Evaluation of impact indicators

- Often impossible to **single out data for rural areas**

  - Datasets from sources other than RDP do not have consistent definition of rural or are too aggregated

- deadweight, leverage, displacement, substitution and multiplier effects virtually **impossible to quantify**

- CMEF guidelines successfuly followed for HNV, Farmland Bird Index, quality of water

Evaluation of impact indicators

- CMEF not strictly followed for:
  - **Economic growth, employment creation, labour productivity**: summary of data from RDP, but no indirect effects taken into consideration
  - **Contribution to combating climate change:** statistical data
  - **Quality of life**: a selection of data was made, some information collected with interviews and surveys

Methodological Constraints

- **Qualitative evaluation** where not enough (good) quantitative data: surveys, interviews
  - Response rate?
  - Depend on interpretation
  - May be subject to temporary conditions (e.g. LEADER: LAGs initially very irresponsive due to a temporary conflict with the Ministry)
  - Example: information on trainings, improvement of economic situation, quality of life

Methodological Constraints

- **Statistical data on rural areas**: data available only on municipality level, including cities $\Rightarrow$ adjustment on data based on assumptions

- **Statistical data on target groups** not always available
  - Good data on farmers, less on companies, owners of cultural heritage,...

- **Difficult to match data from different sources**
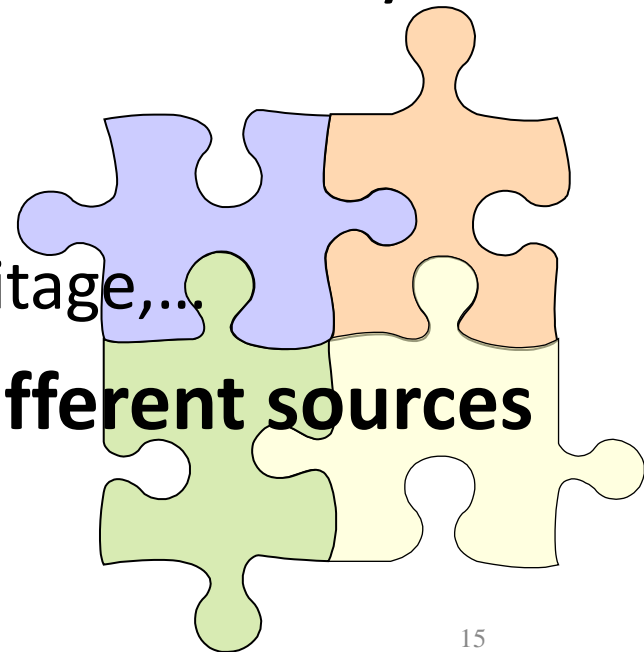
Methodological Constraints

- **Econometric analysis**: difficult to perform due to data inconsistency and poor experience
  - Time consuming data extraction and a lot of effort $\Rightarrow$ results not so useful, not many information for comparison
  - Almost no examples in evaluations in Slovenia
- Answering MTE evaluation questions: **matching qualitative data from surveys with statistical data, use of models**

Evaluation methods that could not be used

- Use of **control group** for analysis
  - vast majority of farmers participate in RDP, those that do not are atypical $\Rightarrow$ no control group
  - Privacy of personal data limits possibility for control group of other measures

- **Beneficiary assessment**:
  - impossible to evaluate individual beneficiaries' preferences, combination of measures, patterns of behaviour and effects

Conclusions and Recommendations

- there is enough RDP data: it is of right quality, but sometimes not in the right format

- statistical data: difficult, if not impossible to separate rural areas $\Rightarrow$ settlements, not municipalities as geographical "data units"?

- Integration and coordination of data on environment and agriculture is recommended
  – Monitoring sites
  – Usefulness for more than one sector

Conclusions and Recommendations

- Redesign of the Agency Database $\Rightarrow$ an analytical tool instead of a repository of data

- fewer indicators $\Rightarrow$ consistent monitoring

- Structuring of beneficiary data
  - What data is wanted?
  - What is actually done with data? $\Big\}$ What data is actually needed? In what format?

- many beneficiaries are willing to report/engage electronically $\Rightarrow$ opportunity for crowdsourcing?

Thank you for your attention!

Any questions?

Mojca Hrabar
Mojca.hrabar@oikos.si
@mojcahrabar
www.oikos.si