



# Integration of administrative data in statistical process

Aladar Belec

Budapest, October 2023



# **Modernisation of Agricultural Statistics - CRB OFB**

The general objective of the project is to establish and to analyse the possible improvement of the use of two administrative data sources:

- **Organic Farming Register (OFR)**
- **Central Register of Bovine animals (CRB)**



- **Organic Farming Register (OFR)**

The aim is to establish the technical environment for using the administrative Organic Farming Register (OFR) on a detailed data level in statistical processes of agriculture annual production statistics.

These newly developed environment and processes will be then integrated with other statistical data and administrative sources in order to be used for the production of relevant statistics on agriculture organic production.



- **Central Register of Bovine animals (CRB)**

The Central Register of Bovine animals (CRB) is already in use as a primary administrative source for statistics on the number of cattle and the slaughter of cattle. But we are still using it in combination with the December livestock survey (telephone data collection - CATI) in order to estimate the number of detailed cattle categories not directly available in the register.

Our aim is to analyse a possible extended use of the CRB as the only source for the number of cattle data (with a view to abolishing the direct collection of data related to detailed bovine animal categories).

This would reduce the burden for the respondents since the CATI collection method would no longer be needed for cattle related data.



- *For the use of the CRB:*
  - **Methodological analysis of the current statistics on the number of bovines with a view to improving and further reducing the burden on reporting units:** overview of the existing method of combining December livestock survey with data from the bovine register (strengths and weaknesses); preparation of available data from December livestock surveys and data from the bovine register.
  - **Analysis of the possibility of using additional administrative sources (as secondary sources, additional information) for the number of cattle:** a detailed analysis and special control of obtained administrative and statistical data sources for all concerned data of bovine animals (data on the purpose of breeding: cows (dairy or suckling), heifers for breeding or for slaughter, calves for breeding or for slaughter).
  - **Investigation to examine whether model estimates can be developed for statistics that cannot be obtained directly from administrative sources (CRB and other):** preparation of a review (overview of the existing methodological processes, determining the methodological issues and possible obstacles in all concerned sources of bovine data, studying the similar administrative source manuals, seeking cases of good practices).
  - **Preparation of a list of required activities for the integration of possible new processes into statistical production:** a proposal of the methodology to obtain as many high-quality data as possible not directly available in the bovine register (data on the purpose of breeding: cows (dairy or suckling), heifers for breeding or for slaughter, calves for breeding or for slaughter); assessment of which methods provides us with good enough data in order to eliminate those variables from the livestock survey.



- *For the use of the OFR:*
  - **Analysis of available data in the context of the new SAIO Regulation statistical demands:** analysis of the use of data within the requirements of the SAIO Regulation; overview of the existing methodological processes, determining the methodological issues and possible obstacles in all concerned organic modules. Studying the administrative source manuals, seeking cases of good practices.
  - **Analysis of the administrative data source:** Set-up technical and methodological aspects of new procedures.
  - **Establishing formal cooperation with the administrative data holder:** supplementing the annual program of statistical surveys and updating the technical protocol with elements that will formalize the use of detailed data in annual production statistics.
  - **Establishing the technical environment inside the statistical process which will allow transparent integration of these data into annual agriculture production statistics:** possible integration of administrative data into agriculture production statistic; detailed analysis of obtained data.
  - **Preparation of procedures for the integration of detailed OFR data into individual statistical surveys in the field of agricultural production:** analysis of the use of data within the requirements of the SAIO Regulation.



# Assessment of the current statistics on the number of cattle

SURS is currently producing statistics on the number of cattle by combining two sources:

- 1. Central Register of Bovine animals (CRB)**
- 2. December livestock survey (KME-DEC)**



With the current methodology of combining the register (CRB) and the December livestock survey (KME-DEC), we are already using the register for 10 main categories, while other categories are calculated by combining the register and livestock survey results at the NUTS2 level (purpose of breeding).





## Central Register of Bovine animals (CRB):

### **Strengths:**

- data available for any given reference date (with the direct access to the database),
- no additional burden for the respondent (compared to surveys),
- more reliable records due to cross checking and control (veterinary inspections and penalties in the case of non-reporting or misreporting or late reporting to the register),
- more exact data (since register is constantly updated all events are recorded with very small time delay) - this makes all data more accurate and precise,
- cheaper (for the statistical office): no additional costs for enumerators, informing the farmers ...
- more complete

### **Weaknesses:**

- not all data (defined with 165 livestock regulation/SAIO legislation is directly available in the registry)
- SURS (the statistical office) has limited influence on changes of the database



## December livestock survey (KME-DEC):

### **Strengths:**

- complete control over the variables that need to be collected

### **Weaknesses:**

- sample survey errors due to nonresponse, sample error ...,
- additional burden for the respondents (they are reporting data to the registry and to the SURS),
- less reliable data:
  - due to voluntary nature of data collection (except for enterprises),
  - collected data is not as controlled as the register data (although we do checks while we are surveying and also post checks of the data that is suspicious and also additional checks of enumerators),
- harder to provide the exact number (it is hard to estimate the age of the animal, especially for the border line cases: animals between 11 and 13 month and between 23 and 25 month) or in case of late reporting also problems with estimating the right values for reference date from more than a month ago (a survey is usually conducted in less than a month, but there were additional checking and also late reporting that do occur in more than a month after reference date,
- additional survey - related cost:
  - costs of enumerators,
  - costs related to the printed questionnaire that are sent to farmers,
- possible misinterpretations between enumerator and respondents



Of course, some problems occurred...



(CRB)

Additional data source was detected later in the Grant and gave the possibility to prepare a better model for using part of administrative data for statistical purposes...



(OFR)

Outsourcing of updating the administrative source (from three digit level to five digit level)...



## Never ever give up...

(CRB) In a wider scope, among all available administrative sources, we found out that one source (inseminations record) could provide us with data suitable for preparing a proper model for using administrative data instead of conducting a survey...

...and we finally get the access to the data source that is just about to be 'updated' (OFR).



Successful prolongation of the Grant based on firm facts; for nine months (at the beginning 9 months more time was available for the Grant, but we believed back then that we will be able to finish earlier, ...well, problems occurred...)



- Part with CRB is very successful and will replace the survey in total:
  - at the moment modelling of cows (dairy/other (suckling)) and calves (for slaughter/for breeding) is (more or less) solved,
  - final effort is now given to provide a model for heifers (for slaughter/for breeding); use of detailed data of the insemination record.





- Part with OFB is after long time standing still, now doing expected progress in building up the structure of technical environment (EKO base) with data of the Ministry to be used for statistical purposes => after building up the ,construction' of the database (three main subdivisions - tables):
  - animals
  - crop
  - agricultural products

there will be a detailed examination of administrative data for use for statistical purposes = change the old EKO base from a 3-digit code list to a 5-digit code list; comparing aggregates with microdata.



Instead of conclusions:

Administrative sources are great for providing us data for statistical purposes, but...

- pros

and

- cons



- CONS:
  - Not the same purpose
  - More burden for data producers
  - Not the same reference period



## PROS:

- No more (expensive) surveys
- Regular data
- Data in time
- Stable data
- Lower burden for farmers



Thank you very much for your attention.

If you have any further questions, you can ask them now

or

even later, via e-mails:

[ali.belec@gov.si](mailto:ali.belec@gov.si)

[ales.stele@gov.si](mailto:ales.stele@gov.si)