INTEGRATION OF ADMINISTRATIVE AND SURVEY DATA TO REDUCE RESPONDENT BURDEN

The Italian experience in the field of agriculture production statistics

Monica Perez et al.
perez@istat.it
Istat – Directorate for Data Collection
Contents

• Istat modernization

• Agricultural statistics: livestock surveys

• Administrative sources for livestock

• Different use of administrative data for livestock statistics

• A first evaluation on a prospective use of administrative data for livestock statistics

• Conclusions
Towards a new organization of data collection

Since April 2016 a new organization model has been set up in Istat, which is now based on Business Architecture Model (BAM)

BAM has overcome “silo” approach. In the past each "silo" identified a specific field of statistics (census, social statistics, business statistics) and its corresponding production system

According to the new model, the production process relies on an integrated organization and centralized corporate cross-cutting support services (technical services, methodology, information technology, data collection, dissemination, …)
• **DCRD centralizes all functions and activities of data collection phase** by managing every single set of data collection process with regard to surveys, administrative data, registers, non-traditional data (Big data).

• **DCRD works in close cooperation with other Istat Directorate/sectors** providing and receiving services
A project for renewing agriculture statistics

• In the filed of agriculture a project has been set up in order to make production of statistics more coherent with Istat modernization guide-lines about redundancy reduction and data integration

• The goal is re-design surveys and making the survey process more efficient and less expensive

• In the data collection process the use of administrative data is a crucial issue

• Project is based on inter-Directorate collaboration, involving Directorate for environmental and territorial statistics in charge of livestock statistics and Directorate of Data collection in charge of finding solutions to reduce redundancy of data and response burden
Livestock statistics in Istat

- Istat carries out a sample survey twice a year about bovines and pigs (referring to the 1st of June and the 1st of December)
- Another sample survey is carried out once a year providing statistics on sheep and goats (referring to the 1st of December)

a) Sample units are the agricultural holdings with animal farms of one or more animal species

b) Sample units drawn for the December wave of “t” year (about 9,000 units) are maintained for the following wave carried out in June of “t+1” year (follow-up on the respondents of December, involving about 6,000 units)

c) Surveys are conducted by CATI technique

d) Surveys compliant with Regulation EC1165/2008
Bovine register (BR)

- In Italy, in 2002, the Ministry of Health Decree has established the Bovines Register (BR)
- The main aim of the BR is to preserve human and animal health by monitoring bovine meat production and marketing conditions

- A National Data Bank (NDB) has been set up in which all relevant information about different “actors” mentioned in the BR, need to be recorded (bovines, heard, farmers, etc.).

- In the NDB the following information are registered:
  - each animal identified by ear tags;
  - each event concerning the animals, within a given time;
  - each firm and herd;
  - veterinary service competent for each firm and heard
Livestock statistics: modernization and burden reduction

Plan of interventions consists of three main steps:

1) **reducing** surveys **sample size** preserving the quality of estimations

2) a **progressive shift** of survey technique **from Cati to Cawi**

3) the **assessment of quality of administrative data** available at Ministry of Health with the aim of (re)using them to produce statistics on livestock directly
Can NDB be used for livestock statistics?

Till now NDB has been used more to compare and validate data estimations produced by Istat surveys rather than as a real alternative data source

A better exploitation of NDB could bring important advantages:

(i) avoiding redundancy of information between two “official” sources (Istat and Ministry of Health)
(ii) making the statistical system more efficient
(iii) reducing the statistical burden on the respondents
(iv) reducing costs
(v) improving the overall quality of the data
Different use of administrative data

In this perspective, three different scenarios using NBD can be identified, with different degrees of impact on the project of renewing

I. The simplest usage of NBD could consist in data exploitation in the survey process as regards data editing/imputation, as far bovines are concerned ➔ low impact

II. Combining administrative data with sample survey data. A partial substitution of the survey with the administrative source as far as bovines and buffaloes are concerned ➔ medium impact

III. Produce statistical outputs directly from administrative data (with a breakdown of the time series) ➔ high impact
The main steps for assessment of NDB:

1. analyzing metadata in order to understand the relationship between BDN concepts and definitions and those used by Istat
2. studying the quality of information supplied by the NDB (data consistency, timeliness, etc.)
3. making comparison between NDB data and Istat’s survey data
4. making operational and sustainable proposals addressed to interventions that allow a better exploitation of NDB for statistical purposes

A teamwork of experts from Istat and Ministry of Health
Which use of NDB for livestock statistics?

Three main issues affect the use of NDB:

1) Data need to be pre-treated in order to reconcile the definition of the reference unit, which is different:
   - **reference unit in the survey** = economic unit of agricultural production, including one or more establishment or (open) place where animals are kept and bred→ reference unit (i.e. farm holder) can host more than one herd and be located in more than one place
   - **reference unit in NDB** = place where a herd is kept and bred → each place can host one or more herd; herds can be owned by different farmers
     - In the same farm can be found different holders
     - If an holder breeds two different herds (eg. bovines and pigs) he/she is counted twice in NDB

2) NDB could not be timely updated
   - risk of over-coverage of units, due to farms that have ceased business but are still registered in the database

3) Different classifications are used
About definitions of reference unit...

- Reconciliation of the reference unit is made by using the fiscal code of the owner
- Fiscal code identifies the herd belonging to the same owner but located in different sites (i.e. in different firm, according to NDB)
About NDB updating…

- In the register, data on livestock are checked and monitored.
- Farms with no signal of modifying data for a long time are identified as suspected to be «ceased».
- By linking the register with other administrative data, it can be verified if the farm is still active or not.
About classification...

<table>
<thead>
<tr>
<th>STATISTICAL CLASSIFICATION</th>
<th>ADMINISTRATIVE CLASSIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1 year slaughtered like calves</td>
<td>Bovines less than 8 months years old</td>
</tr>
<tr>
<td>&lt;1 year breeding males</td>
<td>Bovines males aged from 9 to 12 months</td>
</tr>
<tr>
<td>&lt;1 year males for slaughter</td>
<td>Bovines females aged from 9 to 12 months</td>
</tr>
<tr>
<td>&lt;1 year breeding females</td>
<td>Bovines males aged from 1 year to less than 2 years</td>
</tr>
<tr>
<td>&lt;1 year females for slaughter</td>
<td>Bovines females from 1 year to less than 2 years</td>
</tr>
<tr>
<td>1-2 years breeding males</td>
<td>Bovines males aged 2 years and over</td>
</tr>
<tr>
<td>1-2 males for slaughter</td>
<td>Heifers (bovines aged 2 years old and over without calves)</td>
</tr>
<tr>
<td>1-2 years breeding females</td>
<td>Cows in dairy farms (bovine aged years old with at least one calf)</td>
</tr>
<tr>
<td>1-2 years females for slaughter</td>
<td>Cows in other farms (bovine aged 2 years old with at least one calf)</td>
</tr>
<tr>
<td>&gt;2 years breeding males</td>
<td>TOTAL CATTLE</td>
</tr>
<tr>
<td>&gt;2 years males for slaughter</td>
<td>Foreign origin total bovines (not born in Italy)</td>
</tr>
<tr>
<td>&gt;2 years breeding heifers</td>
<td>Buffaloes less than 8 months</td>
</tr>
<tr>
<td>&gt;2 years heifers for slaughter</td>
<td>Young male buffaloes aged from 9 to 12 months</td>
</tr>
<tr>
<td>&gt;2 years dairy cows</td>
<td>Young female buffaloes aged from 9 to 12 months</td>
</tr>
<tr>
<td>&gt;2 years other cows</td>
<td>Male buffaloes aged from 1 year to less than 2 years</td>
</tr>
<tr>
<td>TOTAL CATTLE</td>
<td>Female buffaloes aged from 1 year to less than 2 years</td>
</tr>
<tr>
<td>Of which from foreign origin</td>
<td>Male buffaloes aged 2 years and over</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TOTAL BUFFALOES</th>
<th>TOTAL BUFFALOES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Of which from foreign origin</td>
<td>Breeding female buffaloes (buffaloes aged 2 years and over with at least a calf)</td>
</tr>
<tr>
<td></td>
<td>Buffaloes in not dairy breeding (buffaloes aged at least 2 years with at least a calf)</td>
</tr>
</tbody>
</table>

• to reconcile AdmCl and StatCl a «bridge» table has been build-up by using information on production target of the unit (also available in NDB)
### About classification...

<table>
<thead>
<tr>
<th>STATISTICAL CLASSIFICATION</th>
<th>ADMINISTRATIVE CLASSIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1 year slaughtered like calves</td>
<td>Bovines less than 8 months years old + Bovines males aged from 9 to 12 months + Bovines females aged from 9 to 12 months</td>
</tr>
<tr>
<td>&lt;1 year breeding males</td>
<td>Bovines less than 8 months years old + Bovines males aged from 9 to 12 months</td>
</tr>
<tr>
<td>&lt;1 year males for slaughter</td>
<td>Bovines less than 8 months years old + Bovines males aged from 9 to 12 months</td>
</tr>
<tr>
<td>&lt;1 year breeding females</td>
<td>Bovines less than 8 months years old + Bovines females aged from 9 to 12 months</td>
</tr>
<tr>
<td>&lt;1 year females for slaughter</td>
<td>Bovines less than 8 months years old + Bovines females aged from 9 to 12 months</td>
</tr>
<tr>
<td>1-2 years breeding males</td>
<td>Bovines males aged from 1 year to less than 2 years</td>
</tr>
<tr>
<td>1-2 males for slaughter</td>
<td>Bovines males aged from 1 year to less than 2 years</td>
</tr>
<tr>
<td>1-2 years breeding females</td>
<td>Bovines females from 1 year to less than 2 years</td>
</tr>
<tr>
<td>1-2 years females for slaughter</td>
<td>Bovines females from 1 year to less than 2 years</td>
</tr>
<tr>
<td>&gt;2 years breeding males</td>
<td>Bovines males aged 2 years and over</td>
</tr>
<tr>
<td>&gt;2 years males for slaughter</td>
<td>Bovines males aged 2 years and over</td>
</tr>
<tr>
<td>&gt;2 years breeding heifers</td>
<td>Heifers (bovines aged 2 years old and over without calves)</td>
</tr>
<tr>
<td>&gt;2 years heifers for slaughter</td>
<td>Heifers (bovines aged 2 years old and over without calves)</td>
</tr>
<tr>
<td>&gt;2 years dairy cows</td>
<td>Cows in dairy farms (bovine aged years old with at least one calf)</td>
</tr>
<tr>
<td>&gt;2 years other cows</td>
<td>Cows in other farms (bovine aged 2 years old with at least one calf)</td>
</tr>
<tr>
<td>TOTAL CATTLE</td>
<td>TOTAL CATTLE</td>
</tr>
<tr>
<td>Of which from foreign origin</td>
<td>Foreign origin total bovines (not born in Italy)</td>
</tr>
</tbody>
</table>
Use of NDB for livestock statistics

Ex-post comparison between Istat and NDB classification show that:

- the total number of bovine and buffaloes livestock have a high degree of comparability. The differences between the two sources are 0.2% for bovine and 0.4% for buffaloes.
- Both for cattles and buffaloes data are less consistent at level of animal age;
- important differences exists for some categories

The bridge table could be sharpened by using additional information – such as the race (not only cattle but also pigs, sheeps) and the precise age of the animal- which are into NDB but not yet available.
Final remarks

• Results suggest that the use of NDB in place of Istat surveys could be a possible way for producing statistics on livestock, containing the statistical burden

• To build up a more accurate “bridge” table, using further information

• A condition for adopting NDB for statistical purposes is that the quality of the statistics obtained can not be lower than those obtained by ad hoc surveys

• Transition from survey to NDB-based statistics requires an increase of coverage

• Last but not least, a strong cooperation between Istat and Ministry of Health is needed

• Practice in this field confirmed importance to find solution ex-ante, before administrative database is implemented
Questions and suggestions are welcome

Thank you

perez@istat.it