**European structural farm statistics — new quality rating system**

Denisa Florescu,Eurostat, [Denisa.Florescu@ec.europa.eu](mailto:Denisa.Florescu@ec.europa.eu)[[1]](#footnote-1)

**Abstract**

*Eurostat, together with the statistical bodies belonging to the European Statistical System (ESS), has adopted a quality rating system to guide the dissemination of structural farm statistics derived from farm structure surveys. The system does this by showing when the estimates are sufficiently reliable to be published, either with or without a warning.  It is based on:*

* *coefficients of variations for totals and means of continuous variables;*
* *standard errors for proportions and counts.*

*This paper also presents the work carried out to harmonise variance estimation methods and their application within the ESS. To apply the new quality rating system consistently, Eurostat and the national statistical bodies must compute roughly the same variance estimates.*

*Future structural farm statistics will come from the data collected from ‘Integrated Farm Statistics’, based on a modular approach. This will lead to more complex national sampling designs. The paper also outlines ongoing developments towards integrating additional sampling design information specific to national multi-stage sampling in the estimation of variance.*

*The paper also introduces new quality reporting based on the European Standard Quality Reporting System (ESQRS) template. This is of great help in assessing all quality dimensions, thereby improving the quality of EU data and metadata.*

*Farm structure surveys are the main source of information on the current state of agriculture and the trends it is undergoing, required to monitor the common agricultural policy and other EU policies. High-quality data are essential for decision-makers.*

**Keywords:** farm structure surveys, quality, quality rating system, variance estimation, quality reporting

**1. Introduction**

The European Statistics Code of Practice states that statistical authorities should regularly monitor the quality of statistical data.

Regulation (EC) No 1166/2008 on farm structure surveys (FSS) and the survey on agricultural production methods (for the survey years 2010, 2013 and 2016) and the proposal for a Framework Regulation on Integrated Farm Statistics (IFS) (for the years 2020, 2023 and 2026) require estimates to be representative at the level of NUTS 2 regions, by farm type and farm size. Estimates are therefore expected to be reliable so they can be disseminated without quality concerns as regards these breakdowns. However, Eurostat publishes FSS estimates broken down by many variables that are not necessarily included in the national sampling design stratification. In 2014, the FSS Working Group asked for a quality rating system in FSS/IFS to avoid publication of poor quality estimates which would send misleading messages to users. The quality rating system should guide the decision on which estimates to disseminate, with or without a warning, and which are not to be disseminated, for all types of population breakdowns. Until the 2013 FSS, Eurostat published FSS estimates applying confidentiality rules. Since the 2016 FSS, Eurostat has also applied quality criteria to the disseminated data.

**2. The quality rating system**

Eurostat has reviewed existing practices for assessing the quality of estimates. These practices use various quality measures, most notably ***sampling errors***, the***numbers of sample observations (sample sizes)*** and ***response rates***, all of them computed per population breakdown (table cell).

On ***sampling errors***, the ESS guidelines (Eurostat, 2013) recommend using coefficients of variation (relative standard errors) for totals and means of continuous variables, and using standard errors for proportions, given the ‘symmetry’ i.e. equality of standard errors for a proportion and for its ‘opposite’ proportion[[2]](#footnote-2), which does not hold for coefficients of variations for proportions. Although confidence intervals are quite suggestive to users, their validity is conditioned by the normal distribution assumption which might not hold for small samples of holdings falling into a breakdown. On the other hand, standard errors and coefficients of variation are not conditioned by the normality assumption, and they are meaningful even if the sample is small.

Most of the FSS/IFS estimates are totals of continuous variables. Some are counts and a few are means of continuous variables and proportions.

The question has been which sampling errors should be used for counts. National statistical authorities use standard errors or coefficients of variation. Technically, the total of a count variable is the same as the proportion, especially as Eurostat disseminates the number of agricultural holdings in the population (which is an approximation to the number of holdings in the frame). Moreover, the coefficients of variation for a count and for its corresponding proportion are equal, which entails the ‘non-symmetry’ of the coefficient of variation for a count and for the ‘opposite’ count[[3]](#footnote-3). It was decided to use standard errors for counts, at least until there are further arguments in support of the alternative. However, this approach is controversial; some statisticians claim that the counts should be treated as continuous variables and coefficients of variation should be used for them.

The role of ***the sample size*** is to indicate potentially large sampling errors, which does not bring any additional value to the use of sampling errors.

***The response rate*** is not enough to detect bias, as bias is basically determined by the relation between the response probability and the variables of interest. In practice it is very difficult, if not impossible, to assess this relation for a large number of cells. Eurostat generally considers that weight calibration reduces bias. The main role of the response rate, therefore, is to indicate the risk of large sampling errors, which again brings no additional value to the use of sampling errors. Moreover, Eurostat knows the response rate only for the overall national population (from the quality reports). In the absence of the necessary information in the microdata, it cannot calculate the response rate by cell.

In conclusion, the FSS Working Group agreed on a FSS/IFS quality rating system based on coefficients of variation for totals and means of continuous variables and on standard errors for proportions and counts. Given that a standard error has the same measurement unit as the indicator it refers to, this leads to defining separate quality rating systems based on standard errors for proportions and for counts. To avoid such separate quality rating systems, Eurostat has decided to convert counts to proportions (defined in relation to the overall population) and then apply a single quality rating system based on standard errors to proportions.

In 2016, Eurostat analysed some published 2013 FSS datasets, with a view to seeking a balance between acceptable quality level (confirmed by existing practices), relevance (share of cells to be disseminated in FSS) and resources (affordable sample sizes).

The FSS Working Group adopted the following **quality rating system**.

For ***totals and means of continuous variables***, the system is based on the values of their coefficients of variation (CV):

* an estimate is released if its CV is below 24.99%;
* an estimate is released with a warning if its CV is between 25.0% and 34.99%;
* an estimate is not released if its CV is at least 35.0%.

For ***proportions***, the system is based on the values of their standard errors (SE):

* an estimate is released if its SE is below 12.49 percentage points;
* an estimate is released with a warning if its SE is between 12.5 and 17.49 percentage points;
* an estimate is not released if its SE is at least 17.5 percentage points.

The annex shows the quality of 2016 FSS estimates in cells for new data tables published on Eurostat's website. It shows that while the number of the breakdown variables increases gradually from one table to the next, the quality of data in breakdowns gradually decreases, as expected. The computations cover 23 EU countries[[4]](#footnote-4) for which the 2016 data are available at the time of writing this paper and which based their survey on a one-stage stratified sampling design.

**3. Harmonisation of variance estimation in FSS / IFS**

A condition for applying the quality rating system is that the sampling error estimates computed by each country and by Eurostat must be approximately the same. Otherwise, the quality rating system leads to different conclusions as regards the dissemination of the same point estimates at national and EU level.

*3.1. Quality reporting*

Since the 2013 FSS survey, the national authorities have been drawing up a quality report (which is also a national methodological report) based on the European Standard for Quality Reports Structure (ESQRS) template in the ESS Metadata Handler application. Using such a template makes it easier to assess comparability across countries and over time. The concepts *Sampling design*, *Data compilation,* *Sampling error* and *Sampling error – indicators* include information on the national sampling design, estimators, variance estimation method and the respective estimates of the relative standard errors for the main crop and livestock variables for which thresholds are set in Regulation (EC) No 1166/2008. The proposal for a Framework Regulation on Integrated Farm Statistics sets similar precision requirements for future data collections. Eurostat compares the estimates of the relative standard errors reported in the national quality reports with its own computed values and checks whether there are any differences, taking into account the national sampling design, estimators and variance estimation method. While analysing the data (other than microdata validation) and drawing up the quality report, Eurostat completes a detailed validation checklist for each country with methodological findings. The validation checklist is a tool used both to support formal approval of the data and the quality report, and to store methodological information.

*3.2 Reasons identified for differences between relative standard errors, and solutions to reconcile the differences*

Table 1 presents the reasons for the differences in relative standard errors between national computations and those performed by Eurostat, along with the possible ways in which such differences may be reconciled. The FSS Working Group agreed on the solutions, except on the ones corresponding to the reasons numbered 1, 2 and 3, which require further analysis.

**Table 1. Reasons for and solutions to differences in relative standard errors**

| **No** | **Reasons** | **Solutions** |
| --- | --- | --- |
| 1 | ***Different sampling designs***  Eurostat uses a variance estimation method for one-stage stratified sampling, while a few countries apply a more complex sampling design. | Eurostat to adopt a variance estimation method which accounts for more complex designs than one-stage stratified sampling. The countries concerned to report additional sampling design information. See section 3.3. |
| 2 | ***Systematic sampling with implicit stratification not accounted for***  Eurostat disregards the effect of systematic sampling with implicit stratification on the variance estimates. | The countries concerned to report the order of selection of the sampling units in additional dataset fields, where that order is determined by a variable correlated with the main variables of interest. Eurostat to add the field ‘computational stratum’ to the dataset in which the sampling units within each original stratum are to be grouped (e.g. paired) (ensuring that each ‘computational’ stratum has at least two units). See section 3.3. |
| 3 | ***Imputation*** ***is either*** ***not accounted for at all, or is not properly accounted for***  Eurostat and many countries treat imputed values as collected values. A few countries exclude imputed values when estimating variance. | Several different procedures have been developed for variance estimation in the context of imputation, but they are rarely used in practice in official statistics, as they are quite demanding.  For small percentages of imputations (to be defined), Eurostat and the countries concerned to treat imputed values as collected values when estimating variance.  For massive imputations, Eurostat and the countries concerned to use a method to account for the variability caused by imputation; however, this requires countries to transmit, for each variable, a separate field in which the imputed values are marked, and some countries do not keep record of the imputations.  To be further analysed. |
| 4 | ***Calibration not entirely accounted for***  Eurostat and some countries concerned consider the final calibrated weights, but do not include the variability of the calibration residuals when estimating variance. A few countries concerned incorporate the calibration variability into the variance estimate. | Correct procedures to include calibration variability are quite complex, and it is not practicable to apply them to about 30 countries with different calibration methods and variables. For example, a study for the Labour Force Survey (LFS) reached the same conclusion (Eurostat, 2015).  Under an FSS simulation study, based on Monte Carlo repeated sample simulations, variance was computed using the Taylor linearisation estimator of calibrated total by both considering and ignoring the residuals. The conclusion was that the bias introduced by not considering the residuals is not significant (Eurostat, 2015-2018).  The agreed way forward is for Eurostat to consider only the final calibrated weights, while the countries concerned should, if possible, also consider the calibration residuals. This entails the risk that there may still be some differences in the relative standard errors. |
| 5 | ***Different approach concerning the values of ineligible holdings***  Eurostat and many countries exclude the values of variables of ineligible units, while a few countries include them when estimating variance. | Eurostat and countries to exclude the values of variables of ineligible holdings when estimating variance, as this is also done in point estimation. |
| 6 | ***Different approach concerning the zero values of eligible holdings***  Eurostat and many countries include the zero values of variables of eligible units, while one country excludes them when estimating variance. | Eurostat and countries to include the zero values of variables of eligible holdings when estimating variance, as this is also done in point estimation. |
| 7 | ***Different approach to the real outliers***  Eurostat and many countries include the holdings whose main variables have high values in the reference period and low values in the sampling frame, while one country excludes such holdings when estimating variance. | Countries to adjust to one or trim the weights of real outlier holdings. Eurostat and countries to include the values of real outlier holdings when estimating variance, as this is also done in point estimation. |
| 8 | ***Different approach concerning allocation of holdings to strata***  Some holdings change attributes related to stratification between the sampling design and the reference period. Eurostat and many countries estimate variance over the original strata recorded in the dataset, while a few countries do that over updated strata. | Countries to move outlier holdings to a take-all stratum and keep original strata for the other holdings. Countries to perform overall calibration and record the final calibration weights. Eurostat and countries to estimate variance over the original strata (and updated stratum for outliers) using the final calibrated weights. |

*3.3. New sampling design information in datasets*

To correctly estimate variance as regards the first two reasons for the differences in relative standard errors presented in Table 1, additional sampling design information should be included in datasets starting with 2020. Depending on the national sampling design, the information concerned should cover all fields, a number of fields or just one. The fields concern primary strata, primary sampling units (PSUs), secondary sampling units (SSUs) and the order of selection of the units. It is planned that the primary strata should also record self-representing PSUs[[5]](#footnote-5) and collapsed strata with single primary sampling units. Information on the order of selection of sampling units is important for variance estimation, as systematically drawing from a judiciously ordered sampling frame might substantially reduce sampling errors. Information on the order of sampling units is relevant only if the order is determined by a variable correlated with the main variables of interest.

The proposal for a Framework Regulation on Integrated Farm Statistics establishes a framework for European statistics at the level of agricultural holdings where variables to be collected are grouped into core and modules. Besides multi-stage sampling, it is expected that some IFS data collections will be based on multi-phase sampling. There are at least two fundamental differences between multi-stage and multi-phase sampling:

* In multi-stage sampling, the units of selection are generally different at different stages, forming some kind of hierarchy. Most frequently, this hierarchy is determined by different levels of spatial units (e.g. enumeration areas, holdings). In multi-phase sampling, on the other hand, selection units are the same at each phase.
* In multi-stage sampling, information is collected only from the units selected at the last stage of sampling (from holdings). In multi-phase sampling, on the other hand, information is collected after each phase, and information collected in the previous phase(s) of sampling is used in the later procedures. There are two main ways in which information from the previous phase can be used:
  + Information is used for the later phase of the sampling procedure (e.g. for stratification). In IFS data collections, some core variables collected with the core sample (e.g. agricultural area utilised, livestock) can be used for the stratification necessary for the selection of the modules' sub-sample(s).
  + Information is used as auxiliary information (e.g. for ratio estimators) in the estimation procedure. In IFS data collections, the estimates of the totals of main variables in the larger core sample can be used in the procedure for calibrating the modules' smaller samples.

Two-phase sampling can be simplified to two-stage sampling, if the independence condition is fulfilled. Independence basically means that that the information collected for the core sample is not used in selecting the module sub-sample. If, for example, the module sub-sample is selected at the same time as the core sample, and the data for both the core sample and the module sub-sample are collected in parallel, the independence condition is met. In such a case, the theory is straightforward, since we are dealing with two independent sampling mechanisms and the inclusion probabilities are (as in the case of two-stage sampling) products of two unconditional probabilities. It is very well known that unit non-response is a case of two-phase sampling, where the sample is selected in the first phase and the respondent ‘sample’ is self-selected in the second phase. In practice, the Missing Completely at Random or the Missing at Random response mechanism assumption is accepted. This means that independent sampling in the second phase (more explicitly, for generating the response mechanism) is assumed, and the re-weighting formulae in fact assume direct sampling of respondents. Under the Not Missing at Random assumption, the second phase sampling (the response) depends on some variable(s) measured in the first phase (the sample). In this case, the simple re-weighting approach cannot be taken and a more complicated (usually model-based) approach must be employed.

**4. Conclusions**

This document has presented the main elements that improve and streamline the quality assessment of farm structure surveys. The main drivers for improving quality, which constitute the main components of the FSS quality framework, are: the quality rating system, the harmonisation and adaptation of variance estimation methods and software tools to all national sampling designs and to the requirements of the proposal for a Framework Regulation on Integrated Farm Statistics, the standard quality reporting and the validation checklist. Quality is checked at the source (e.g. through standard quality reports describing the applied methods) and at the final stage – where data are checked for reliability on the basis of the quality rating system.

**5. References**

Eurostat (2013), Handbook on precision requirements and variance estimation for ESS household surveys. Available at: <https://publications.europa.eu/en/publication-detail/-/publication/f51ba4a6-6587-414d-93d3-a8aaff3106e2> (Accessed: 8 May 2018).

Eurostat (2014), Continuous improvement of EU-SILC quality: standard error estimation and new quality reporting system, Q2014 conference.

Eurostat (2015), Variance estimation in the Labour Force Survey, Internal document for item 4.1 of the agenda, Working Group Labour Market Statistics, Luxembourg.

Eurostat (2017), Description of EU-SILC target variables.

Eurostat (2015-2018), Various Deliverables to the Specific Contract ‘Methodological support in the context of the modernisation in the area of statistics on agriculture and fisheries’, under the Eurostat Framework Contract on Methodology.

Proposal for a Regulation of the European Parliament and of the Council on integrated farm statistics and repealing Regulations (EC) No 1166/2008 and (EU) No 1337/2011, COM(2016) 786 final. Available at: [https://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/COM-2016-786-F1-EN-MAIN.PDF](https://ec.europa.eu/transparency/regdoc/rep/1/2016/EN/COM-2016-786-F1-EN-MAIN.PDF%20%20)  (Accessed: 8 May 2018).

Regulation (EC) No 1166/2008 of the European Parliament and of the Council of 19 November 2008 on farm structure surveys and the survey on agricultural production methods and repealing Council Regulation (EEC) No 571/88 (Text with EEA relevance). Available at: [https://publications.europa.eu/en/publication-detail/-/publication/2e21e197-0847-11e4-a7d0-01aa75ed71a1/language-en](https://publications.europa.eu/en/publication-detail/-/publication/2e21e197-0847-11e4-a7d0-01aa75ed71a1/language-en%20%20%20%20) (Accessed: 8 May 2018).

Statistics Canada, Farm Environmental Management Survey (FEMS). Available at: <http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5044> (Accessed: 8 May 2018).

**ANNEX**

**Table 1. Distribution by quality level of the data broken down by NUTS2 region (table *ef\_m\_farmleg*)**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Quality rating** | **Number of farms** | **Utilised agricultural area (hectare)** | **Farm area excluding special agricultural production areas (hectare)** | **Number of farms with livestock** | **Livestock units (LSU)** | **Standard output (euro)** | **Labour force directly employed (annual working unit)** | **Number of farms whose household consumes more than 50% of the final production** | **Total** | **% in Total** |
| **reliable** | 225 | 225 | 225 | 225 | 225 | 225 | 217 | 223 | 1790 | 99.7 |
| **use with caution** | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 4 | 0.2 |
| **unreliable** | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 0.1 |
| **Total** | **225** | **225** | **225** | **225** | **225** | **225** | **223** | **223** | **1796** | **100.0** |

**Table 2. Distribution by quality level** **of data broken down by economic size and NUTS 2 regions (table *ef\_m\_farmleg*)**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Quality rating** | **Number of farms** | **Utilised agricultural area (hectare)** | **Farm area excluding special agricultural production areas (hectare)** | **Number of farms with livestock** | **Livestock units (LSU)** | **Standard output (euro)** | **Labour force directly employed (annual working unit)** | **Number of farms whose household consumes more than 50% of the final production** | **Total** | **% in Total** |
| **reliable** | 2371 | 2178 | 2149 | 2295 | 2155 | 2286 | 2012 | 2354 | 17800 | 94.4 |
| **use with caution** | 0 | 83 | 94 | 0 | 102 | 41 | 140 | 1 | 461 | 2.4 |
| **unreliable** | 0 | 110 | 128 | 0 | 114 | 44 | 207 | 0 | 603 | 3.2 |
| **Total** | **2371** | **2371** | **2371** | **2295** | **2371** | **2371** | **2359** | **2355** | **18864** | **100.0** |

**Table 3. Distribution by quality level** **of data broken down by agricultural area, type of farm and NUTS 2 regions (table *ef\_m\_farmleg*)\***

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Quality rating** | **Number of farms** | **Utilised agricultural area (hectare)** | **Farm area excluding special agricultural production areas (hectare)** | **Number of farms with livestock** | **Livestock units (LSU)** | **Standard output (euro)** | **Labour force directly employed (annual working unit)** | **Number of farms whose household consumes more than 50% of the final production** | **Total** | **% in Total** |
| **reliable** | 43718 | 26497 | 25665 | 38552 | 29102 | 26689 | 18580 | 40701 | 249504 | 73.5 |
| **use with caution** | 0 | 4134 | 4240 | 45 | 3151 | 4207 | 4818 | 90 | 20685 | 6.1 |
| **unreliable** | 0 | 13087 | 13813 | 26 | 11465 | 12822 | 17776 | 81 | 69070 | 20.4 |
| **Total** | **43718** | **43718** | **43718** | **38623** | **43718** | **43718** | **41174** | **40872** | **339259** | **100.0** |

\* Data exclude France, for which the type of farm is not yet available.

**Table 4*.* Distribution by quality level** **of data broken down by agricultural area, age and sex of manager and NUTS 2 regions (table *ef\_m\_farmlang*)**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Quality rating** | **Number of farms** | **Utilised agricultural area (hectare)** | **Farm area excluding special agricultural production areas (hectare)** | **Number of farms with livestock** | **Livestock units (LSU)** | **Standard output (euro)** | **Labour force directly employed (annual working unit)** | **Number of farms whose household consumes more than 50% of the final production** | **Total** | **% in Total** |
| **reliable** | 19340 | 10890 | 9682 | 16341 | 10239 | 9617 | 9757 | 17022 | 102888 | 67.0 |
| **use with caution** | 173 | 2543 | 2733 | 1349 | 2205 | 2212 | 2728 | 1467 | 15410 | 10.0 |
| **unreliable** | 110 | 6190 | 7208 | 1030 | 7179 | 5452 | 7138 | 1070 | 35377 | 23.0 |
| **Total** | **19623** | **19623** | **19623** | **18720** | **19623** | **17281** | **19623** | **19559** | **153675** | **100.0** |

1. The methodological developments presented in this paper were made possible with the support of Rudi Seljak, the Secretary of the Statistical Office of the Republic of Slovenia. [↑](#footnote-ref-1)
2. For example, for the ‘proportion of holdings that applied for support according to a certain rural development programme’, the ‘opposite’ is ‘the proportion of holdings that did not apply for support according to that rural development programme'. [↑](#footnote-ref-2)
3. For example, for the count ‘number of holdings that applied for support according to a certain rural development programme’, the ‘opposite’ count is the ‘number of holdings that did not apply for support according to that rural development programme’. [↑](#footnote-ref-3)
4. Bulgaria, Czech Republic, Denmark, Germany, Estonia, Ireland, Greece, Spain, France, Croatia, Cyprus, Latvia, Lithuania, Malta, Austria, Poland, Portugal, Romania, Slovenia, Slovak Republic, Finland, Sweden, United Kingdom. [↑](#footnote-ref-4)
5. Self-representing PSUs are PSUs selected with a probability of one. For variance estimation purposes, each of them is considered a stratum rather than a PSU. Their SSUs become PSUs. [↑](#footnote-ref-5)