

**UNITED NATIONS  
ECONOMIC COMMISSION FOR EUROPE**

**CONFERENCE OF EUROPEAN STATISTICIANS**

**Work Session on Statistical Data Editing**  
(Paris, France, 28-30 April 2014)

Topic (i): Selective editing / macro editing

**Maintenance of Selective Editing in ONS Business Surveys**

Prepared by Daniel Lewis, Office for National Statistics (UK)

**I. Introduction**

1. The Office for National Statistics (ONS) uses selective editing on a number of business surveys to increase efficiency and to focus the editing effort where it improves output quality. These range from monthly surveys with only a few variables to annual surveys with very many variables. Selective editing follows up only those businesses which would have a material impact on published results if their data were in error. Businesses are scored, based on the suspicion of their return and their impact on aggregate estimates. Any business with a score above a pre-determined threshold is edited. It is important that the thresholds are regularly reviewed to ensure that the quality of estimates is maintained. This is difficult to do in practice, since the process inevitably results in businesses whose data are never verified.

2. When many of the ONS business surveys moved to using selective editing, it was accepted in principle that part of the savings realised should be used to regularly review the process by sampling some of the businesses that pass selective editing, to check that the thresholds are still appropriate and re-set them when necessary. However, at that point no formal process or methodology was created for carrying out these reviews.

3. This paper discusses work to develop methods and a strategy for carrying out regular maintenance reviews of each of the ONS surveys using selective editing. The maintenance strategy also includes a process for dealing with any unforeseen data issues arising from selective editing and for dealing with cultural issues in implementing selective editing on surveys.

**II. Background to selective editing in ONS**

4. Selective editing was originally introduced for a small number of ONS business surveys in 2001. A more recent initiative (see Skentelbery et al, 2011) led to the implementation of a more efficient selective editing method for six short term business surveys as well as implementation of bespoke selective editing systems for two annual business surveys which did not suit the other approach.

5. The method for short term business surveys involves calculating the following item score for each key variable for each unit:

$$score_{ij}^t = 100 \times \frac{a_i^t |z_{ij}^t - \hat{y}_{ij}^t|}{\hat{T}_{jd}^{t-1}} \quad (1)$$

$a_i^t$  is the sample design weight for variable  $j$ , unit  $i$  at time  $t$

$z_{ij}^t$  is the unedited variable  $j$  value for unit  $i$  at time  $t$

$\hat{y}_{ij}^t$  is a predicted variable  $j$  value for unit  $i$  at time  $t$

$\hat{T}_{jd}^{t-1}$  is the previous period's total variable  $j$  estimate for domain  $d$ .

6. The key variables are the most important variables for the survey outputs and were identified following consultation with users. The predicted variable is generally the previous period value whenever it is available, or else is based on the relationship between survey variables and business register variables (turnover or employment).

7. The mean of the item scores is calculated to produce a unit score. The system also allows for the use of the maximum score as a unit score, but this option has generally not performed as well in testing.

8. As data arrive for each business, the unit score is calculated and compared against a pre-defined domain threshold. These thresholds are derived through analysis of around three years of previous survey data and are designed to minimise the number of edit failures whilst keeping the domain estimates that would be calculated from the selectively edited micro data within 1% of the final estimate from the previous traditional editing approach. Any business with a unit score above the threshold is re-contacted and asked to confirm all of the values they returned.

9. The method described above was originally introduced for the Retail Sales Inquiry and Monthly Business Survey and has since been implemented also on the Quarterly Business Survey, Quarterly Public Sector Employment Surveys, Vacancies Survey, and Monthly Wages and Salaries Survey.

10. This standard ONS selective editing method works well for surveys with a small number of key variables and good quality predicted variables. The previous period value is available for a large proportion of businesses in the surveys listed above thanks to a high sample overlap between successive periods. ONS has some surveys that do not meet these criteria and so a different approach is required.

11. The Annual Business Survey (ABS) is the survey ONS uses for Structural Business Statistics. It is a complex survey with a large number of detailed financial variables. There is relatively high sample rotation from period to period, meaning that the previous period value is only available for a minority of businesses. In order to implement selective editing in ABS, ONS investigated the use of the Selekt tool developed by Statistics Sweden (Norberg and Arvidson, 2008).

12. Selekt is more appropriate for ABS because of its flexibility in making use of the available data. The scoring function in Selekt is split into three parts: suspicion, impact and importance. The suspicion of a unit's response can be measured in a variety of ways including use of traditional edit rules and by statistical comparison of survey variables (or ratios of survey variables) within homogeneous groups. The potential impact on estimates if the data were in error is estimated using the survey-weighted difference between the returned value and a predicted value. The importance weight allows for flexibility in attaching more importance to particular variables or domains and also includes a term to standardise the score (using either an estimated domain total or standard error).

13. ONS has implemented Selekt for each of the seven sectors that make up ABS. In each case the relative importance of each variable and domain classification was determined through user consultation. These priorities were reflected through differential use of importance weights. A number of useful traditional edit rules were retained to help identify suspicious responses. In addition to this, test variables were set for many variables (and ratios of variables). An iterative approach was used with multiple periods of previous survey data to choose the most effective range of test variables and to set the various parameters in the Selekt system. As before, the aim was to minimise edit failures whilst keeping published estimates within an acceptable level of quality.

14. ONS has also developed a bespoke system for selective editing on the Prodcum survey. The large number of products, including some which are made by only a small number of businesses, makes it

impossible to effectively use the standard ONS selective editing method. The key aim for Prodcom selective editing was to reduce the amount of editing effort in the most efficient way. The selective editing method retains all of the traditional edit rules and uses a score similar to that described in formula 1 to prioritise failing businesses for re-contact.

15. Selective editing has also been in use on the Annual Survey of Hours and Earnings since 2002. The method calculates scores similar to those described in formula 1. Similarly to Prodcom the scores are only calculated for businesses that fail one of the traditional edit rules used in the survey.

16. For each of the surveys described above, selective editing has led to a significant reduction in the micro editing effort required with so far no apparent detrimental effect on the quality of published outputs. However, there is a strong possibility that as time goes on and the underlying business survey population evolves, the methods and parameters selected for those surveys will become inefficient. This could lead to either an increase in the number of edit failures or an unacceptable decrease in the quality of resulting outputs. For this reason, it is important that a maintenance programme is put in place to regularly review the performance of selective editing. Section III of this paper outlines the potential issues that can hinder the success of selective editing in the long term. Section IV describes a process developed for maintaining selective editing in ONS business surveys.

### **III. Issues with the successful long term implementation of selective editing**

#### **A. Threshold review**

17. The key challenge in maintaining a successful selective editing system in the long term is ensuring that thresholds and parameters are kept up to date. Periodically re-calculating the thresholds ensures that the editing remains efficient. This would ideally involve repeating the analysis originally used when setting up selective editing, to balance edit failures with quality of outputs using more recent survey data. However, by implementing selective editing in the first place there are naturally many businesses whose data may have been in error but, because they would not impact on published outputs, were not re-contacted. The lack of any confirmed values for businesses that pass selective editing means that it is not possible to review the performance of thresholds in the same way using only the data regularly collected and edited when producing survey results.

18. There are two options for reviewing selective editing thresholds for a survey that has been using the method for a period of time. Firstly, it might be possible to make some model assumptions about the data and their error structure. From these assumptions, a selective editing method may be developed that does not rely on comparisons with fully edited data to set thresholds. For example, see the methods in Di Zio and Guarnera (2013) and Arbues, Revilla and Salgado (2013).

19. If we do not feel confident enough in the strength of our modelling assumptions, the alternative option is to take a sample of businesses that passed selective editing and re-contact them to confirm their data. In practice, this should mean contacting some businesses whose data were in error but were not deemed by the current thresholds to be important enough to be re-contacted. It may also mean contacting some businesses whose data showed no signs of being in error at all. Clearly this would have to be done in a careful way, with well thought out instructions for querying the data for those businesses.

20. Of these two options, the one being considered by ONS is to re-contact a sample of businesses that passed selective editing. The method to select this sample and to use it to re-set thresholds is still to be finalised. However, an initial study was carried out in 2012 to evaluate the performance of selective editing thresholds in the Retail Sales Inquiry (RSI).

21. RSI has an average monthly sample size of about 3500 businesses with around 900 failing selective editing each month. For the RSI study, it was agreed that the editing team had resource to re-contact an additional 600 businesses as a one-off exercise. Selective editing thresholds are set for each of the 24 smallest published domains in RSI. The 600 businesses were shared amongst these domains using the formula below:

$$\frac{\sqrt{n_{resp} - n_{fail}}}{\sum_{\text{Domains}} \sqrt{n_{resp} - n_{fail}}} \quad (2)$$

Where  $n_{resp}$  is the number of responders in the particular domain and  $n_{fail}$  is the number failing selective editing. A constraint was applied to ensure that each domain had a minimum sample size of 20.

22. Following the allocation of sample size to each domain, the next step was to decide how to choose which businesses are sampled. In doing this, there is a balance between using a sampling method which can be easily used to make inference about the rest of the population and ensuring that the samples contain enough of the businesses whose score is below but close to the threshold to be able to fine tune parameters. For the RSI study, three approaches were considered:

- a) Select a random sample within each domain.
- b) Select a stratified random sample within each domain, using the unit score as a stratification variable.
- c) Split each domain into two strata based on unit score, with full enumeration of the stratum containing the highest scores and a random sample drawn from the other stratum.

For approaches (b) and (c), it is necessary to decide which proportion of businesses should go in each stratum. This will vary by domain since each domain has a different threshold and range of scores.

23. The relative performance of the three approaches outlined above was analysed using RSI data from the periods before selective editing was introduced. Using these data, it was possible to mimic the situation if selective editing had already been in use for those periods, using the method and thresholds implemented later. At the same time, both unedited and fully edited data were available to evaluate the performance of each sampling method.

24. Two measures of ‘‘pseudo-bias’’ were calculated to estimate the impact of selectively editing the RSI data using a particular set of thresholds. The first measure was a full sample estimate of the amount of bias introduced to the survey estimate for each domain. This is exactly the measure of quality used when determining selective editing thresholds in the first instance and is calculated as:

$$\frac{\sum_{\text{full sample}} w_i |y_{i,sel} - y_{i,trad}|}{\sum_{\text{full sample}} w_i y_{i,trad}} \quad (3)$$

Where  $y_{i,sel}$  is the edited value for variable  $y$ , business  $i$  after selective editing and  $y_{i,trad}$  is the edited value after traditional editing.

25. The second measure estimates the pseudo-bias using only a sub-sample of businesses, selected using each of the three approaches described above. The performance of each sampling method could then be evaluated by comparing the bias estimates from the sub-sample with those from the full sample. The sub-sample bias measure was estimated as follows:

$$\frac{\frac{n_{pass}}{n_{sub}} \sum_{\text{sub-sample}} w_i |y_{i,sel} - y_{i,trad}|}{\sum_{\text{sub-sample}} w_i y_{i,trad}} \quad (4)$$

Where  $n_{pass}$  is the number of businesses passing selective editing and  $n_{sub}$  is the number of businesses in the sub-sample.

26. These bias measure comparisons were calculated for 35 months of RSI data, from February 2007 to December 2009. For approaches (b) and (c), a range of different proportions were tested for splitting the sample amongst strata. The results of the study showed that approach (a), selecting a random sample within each domain, was by far the most accurate method for estimating the pseudo-bias.

27. Approach (a) was therefore used in the study to estimate whether the quality of RSI data had deteriorated since the introduction of selective editing. The pseudo-bias was estimated to be smaller than 1% for each output domain, so no further action was required. Note that using approach (a) does not guarantee that we select enough units close enough to the thresholds to be able to adjust the thresholds when the bias does exceed 1% (or if we discover the bias is much smaller than 1% for every domain, suggesting that we are editing too many businesses). More development is required to identify a method which will both accurately estimate the bias and allow us to re-set thresholds.

28. The discussion here has focused on reviewing thresholds for surveys using the standard ONS selective editing method. For ABS, any review would also require updating the various parameters and choice of test variables in the Selekt system. This would potentially require a different sampling method. Note that Norberg and Arvidson (2008) discuss a method for selecting units to review Selekt parameters. They propose using Probability Proportional to Size (PPS) sampling to select a sub-sample of units below the selective editing threshold with probabilities proportional to the score. Their suggestion is then to take a Poisson sample without replacement.

## **B. Unforeseen data issues**

29. When selective editing is implemented on a survey, it naturally brings a change to the way data are processed and edited. In some cases, unforeseen data problems can emerge that the selective editing system does not cope properly with.

30. One example of an unforeseen issue involves the use of specific edits in RSI to identify when responses do not refer to the correct period. These edits were removed when selective editing was implemented, using the logic that any important deviation in the response data from what is expected should be identified by selective editing. Specific analysis prior to the introduction of selective editing suggested that removing the date edits would not impact on the quality of RSI outputs. However, the analysis at that time focused on businesses that returned data for periods of the wrong length. It has since become apparent that some businesses return data for the correct length of period, but for the wrong time of year – for example returning March data when the survey asked for April. Because this outcome was not planned for, the only way to resolve these issues has been to manually identify and re-contact every business exhibiting this behaviour. This results in a less efficient editing process.

31. It should be possible to deal with data issues, such as the one described above, by specifying additional (often automated) procedures as part of the overall editing process. However, it is important to have a mechanism for the affected survey area to be able to raise these issues as soon as they are discovered and, wherever possible, for the issues to be resolved in a timely way. Dealing with this type of data issue is an important part of the maintenance of a selective editing system.

32. Changes to surveys can also make it necessary to review selective editing. For example, a change in the nature or number of questions on a survey or to the industries covered can make it necessary to re-specify the system. Whenever such changes are planned, it is important that the impact on selective editing is considered sooner rather than later.

## **C. Cultural issues**

33. In most cases, one of the key aims of introducing selective editing on a survey is to reduce the amount of resource spent on the editing process. To achieve this, it is not sufficient to simply provide the scoring method and associated thresholds. It is important that the people working with the survey data understand and support the principles of selective editing. In practice, this means changing the culture away from the traditional attitude of examining every record containing a suspicious value, just in case it might be an error.

34. The change in culture should ideally be brought about at the same time as the selective editing method is being developed and implemented. In ONS, some success was achieved in doing this for the initial surveys that used the now standard ONS method. It was, of course, important to involve the survey managers in the process right from the beginning, to ensure they supported the initiative. However, in addition to this, regular workshops were organised for everyone involved with the survey. This included both the editing and results teams, but also some of the internal users of survey results. The workshops described the plans for selective editing but also gave the opportunity for everyone to make their own suggestions and raise any concerns. This approach was ultimately effective in changing the culture.

35. The importance of involving all staff affected by the introduction of selective editing is underlined when considering surveys where this has not happened. After the successful implementation of the standard ONS method in a number of surveys, the move of subsequent surveys to selective editing happened much more quickly. In one case, the survey team did not fully trust the method and continued to carry out a large amount of additional micro editing, thus losing most of the efficiency gains that selective editing had introduced. When this was realised, a workshop was organised to discuss the issues. It transpired that the key problem was a lack of understanding of the selective editing method and the impact it had on published estimates. Once the survey team realised this, they were happy to stop using the additional micro editing checks.

36. The example above shows the importance of embedding an appropriate culture of selective editing. Without spending time to describe the method to staff involved, there is a danger that the improvements in editing efficiency will not be realised.

#### **IV. A process for maintaining selective editing in ONS**

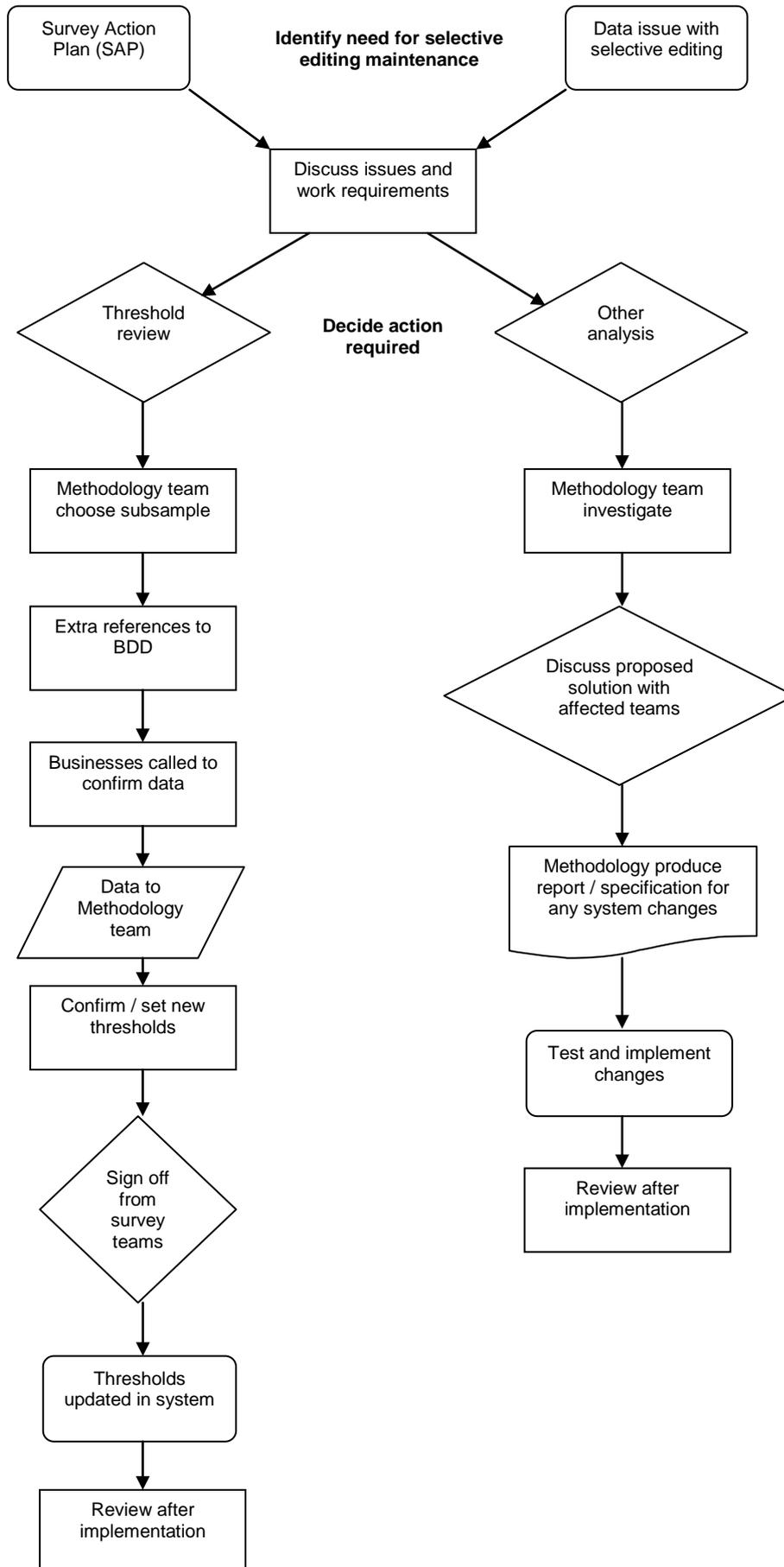
37. Taking account of the need to regularly review thresholds, to deal with data issues and to ensure an appropriate change in culture, ONS has begun to develop a continuous process for maintaining selective editing. Without a dedicated process, it can be difficult to secure the resource required for re-contacting extra businesses, testing thresholds and dealing with data issues.

38. It has been agreed within ONS that any annual survey using selective editing should be reviewed every five years, and that the monthly and quarterly surveys should be reviewed every three years. This review should include testing that thresholds are still valid and re-setting them if necessary. Any data and cultural issues identified should also be addressed. All reviews will involve staff from all affected areas, including editing and results teams, methodology and the IT teams maintaining systems.

39. A pilot review will be carried out in 2014 to finalise methods for sampling businesses that pass selective editing and for re-setting thresholds when appropriate. Following this, a rolling programme of reviews will be established to fit in every survey and ensure an even spread of resource across the cycle. Additional development work will be required to determine the best way for reviewing thresholds and parameters for surveys using the Selekt system.

40. It may be the case that problems are identified with selective editing in surveys before they are due to be reviewed. It is important that there is a vehicle for these issues to be raised and addressed. ONS is currently setting up a wider annual programme of Survey Action Plans, to review surveys and determine if there are any important issues that need to be dealt with. As part of this programme, each survey using selective editing will be asked about the continuing effectiveness of their editing process. A resource will be made available for dealing with any problems that are encountered. Figure 1 shows a flow chart for the selective editing maintenance process.

**Figure 1: Flow chart of selective editing maintenance process**



## V. Conclusions

41. This paper has discussed the need to formally maintain selective editing systems used in business surveys. When selective editing is first introduced, it offers the opportunity for efficient micro editing, focused on ensuring the quality of survey outputs. However, without regular review the thresholds can become out of date, potentially leading to an inefficient process and low quality outputs.

42. Because selective editing involves re-contacting only a subset of businesses with suspicious values, it is necessary to develop a strategy for reviewing and re-setting thresholds. In ONS it is planned to do this by sampling some of the businesses that would normally pass selective editing and re-contacting them to confirm their data.

43. Selective editing performance can also be compromised by unforeseen data issues and a failure to properly embed the culture of selective editing in the statistics office. To address all of these issues in a structured way, ONS has set up a formal maintenance process for surveys using selective editing. The process will be piloted in 2014 and fully implemented following that.

## VI. References

Arbues, I., Revilla, P. and Salgado, D., 2013, *An Optimization Approach to Selective Editing*, Journal of Official Statistics, Vol. 29, No. 4, pp. 489-510.

Di Zio, M. and Guarnera, U., 2013, *A Contamination Model for Selective Editing*, Journal of Official Statistics, Vol. 29, No. 4, pp. 539-555.

Norberg, A., & Arvidson, G., 2008, *New Tools for Statistical Data Editing*, UNECE Work Session on Statistical Data Editing, Vienna, April 2008.

Skentelbery, R., Finselbach, H. and Dobbins, C., 2011, *Improving the efficiency of editing for ONS business surveys*, UNECE Work Session on Statistical Data Editing, Ljubljana, May 2011.