WP.20
ENGLISH ONLY

UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing
(Ljubljana, Slovenia, 9-11 May 2011)

Topic (iii): Macro Editing Methods

Improving the efficiency of editing for ONS business surveys

Invited Paper

Prepared by Rachel Skentelbery, Hannah Finselbach & Claire Dobbins, Office for National Statistics (UK)[1]

I. Introduction

1. The Office for National Statistics (ONS) is currently running a stream of work known as 'The Eden Project'. The aim of the project is to focus the editing process on meeting customer needs for data quality, whilst making the most efficient use of resources and reducing respondent burden, where possible. One of the main ways this will be achieved is through increased use of selective editing.

2. Two ONS short-term business surveys (Retail Sales Inquiry [RSI] and Monthly Business Survey [MBS]) have already been through the Eden Project and a new selective editing methodology has been implemented (Hooper & Lewis, 2010). The current work is focused on investigating ways of implementing selective editing into the Annual Business Survey (the ONS equivalent to the Structural Business Survey). The Annual Business Survey (ABS) presents more challenges for selective editing than short term surveys, such as RSI, because of the large number of variables and the fact that a high proportion of the businesses are in the sample for the first time and so do not have previous values. The approach we currently have for the ABS is a piece of selective editing software called SELEKT, developed by Statistics Sweden. Through collaborative work with Southampton University and Statistics Sweden the SELEKT tool has been applied to the Production & Construction (PC) sector of the ABS.

3. This paper gives an overview of the Eden Project and briefly describes the methods used for the short term business surveys. It focuses on the approach that ONS are investigating to use the Statistics Sweden selective editing tool, SELEKT, in combination with the ONS tool, Snowdon-X for the Annual Business Survey.

II. Overview of the Eden Project

4. The Eden Project aims to develop and implement an efficient editing strategy for ONS business surveys. The project is considering both micro and macro editing and focuses on the customer needs for quality. Traditionally, ONS has used an approach which tries to detect and correct all errors in survey data by using a series of standard and user defined edit rules. If a record fails an edit rule then, generally, the respondent is re-contacted to confirm or correct the return. However, this can result in a high cost and

[1] The author acknowledges the valuable work of Emma Hooper and Daniel Lewis of the Office for National Statistics and Pedro De Silva of Southampton University, for their work on selective editing of RSI and MBS. Also Pedro De Silva, and Anders Norberg of Statistics Sweden for their valuable input in applying SELEKT to Annual Business Survey Data.

resource can be wasted checking information that may not lead to a change. Also, if a change does occur it may not impact the final output.

5. The editing strategy under 'The Eden Project' involves utilising selective editing techniques to try to minimise re-contact with respondents whilst maintaining data quality. This paper will focus on the application of selective editing to micro data and use examples for the RSI, MBS and ABS.

III. Methodology for Short Term Business Surveys

6. ONS has applied a 'new' selective editing methodology to both the Retail Sales Inquiry (RSI) and the Monthly Business Survey (MBS). The 'old' selective editing approach, for MBS, involved a sequence of automatic editing; edit rules, and finally the selective editing score calculation. If any of the failing responses had missing values or the contributor was new to the survey the return was checked. For all other edit rule failures the return would pass through the selective editing score calculation. Those returns with a score above a specified threshold would then be manually edited.

7. The 'new' approach identified key variables and domains, after consultation with users, and calculates an item score for each of these variables. For RSI, the key variables are turnover and employment and the key domain is the lowest published output group. For MBS, the key variables are turnover, employment, new orders and export turnover; and the key domain is input-output groups. The scores compare returned values with predicted values, where the predicted values are generally estimated using past data, a pseudo-imputed value, or other available information related to the variable, for example, from administrative sources. An item score is calculated for each identified key variable for each unit as:

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

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.

8. The items scores are then combined into a unit score using the Minkowski distance function (Hedlin, 2008).

$$u_i^t = \left(\sum_{j=1}^{p} (score_{ij}^t)^{\lambda} \right)^{\lambda^{-1}} \qquad (2)$$

p is the number of item scores for unit i

$\lambda = 1$ is equivalent to sum of the item scores (or, similarly, the mean)

$\lambda = 2$ is equivalent to the Euclidean distance

Large values of λ would be roughly equivalent to the maximum item score.

9. Limitations with the ONS processing system meant that implementation was restricted to using the mean or maximum of scores, however evidence showed that this was feasible. For both RSI & MBS the mean of scores was used. This unit score is then compared against a pre-specified threshold to decide

whether a record should be manually edited. In setting the thresholds a number of quality indicators were examined, such as bias, change rates and savings (Hooper & Lewis, 2010).

10. It is important to note that a key difference between this method and the old ONS editing approach is that the entire questionnaire fails selective editing rather than certain variables. Therefore if a record fails selective editing the entire record is checked manually.

11. For both the RSI and MBS the Absolute Relative Bias (ARB) was kept below 1%, for all key domains.

$$\hat{ARB}^t_{jd} = \sum_{i \in s^t} w^t_{ij} \times | z^t_{ij} - y^t_{ij} | \times I\left(u^t_i < c_d\right) \Big/ \hat{T}^t_{jd} \qquad (3)$$

s^t is the sample at time t

w^t_{ij} is the estimation weight for variable j, unit i at time t

y^t_{ij} is the edited value for variable j, for unit i at time t

I is equal to 1 if the unit score for unit i at time t is less than threshold c for domain d, else it is equal to 0.

$\hat{T}^t_{jd}$ is the current period's total variable j estimate for domain d.

The relative savings were calculated as:

$$Savings^t_d = \frac{trad^t_d - select^t_d}{trad^t_d} \qquad (4)$$

$trad^t_d$ is the number of units failing at least one traditional edit rule at time t in domain d

$select^t_d$ is the number of units with a unit score above the threshold at time t in domain d.

12. Since Hooper & Lewis, 2010, published their paper some fine-tuning of parameters has been undertaken on the MBS which has had a slight impact on the savings. The estimated savings for RSI and MBS can be seen in Table 1 below. The table shows different savings for non-end, and end of quarter months. These differ as employment is only collected on end of quarter months, employment months were more difficult to find savings for, especially for MBS. One of the reasons for this is that the specified rules in MBS were working quite well for employment before the 'new' methodology was used.

Table 1 – Estimated savings (%) for RSI and MBS

	Non-end of quarter months (employment not collect)	End of quarter months (employment collected)
RSI	63.9	41.2
MBS	46.7	3.2

13. The 'new' selective editing approach was implemented in RSI in July 2010 and MBS in August 2010. There were no major issues during implementation and the savings have been realised. The next step for both surveys is to investigate the macro editing to see if it can be introduced to compliment the micro editing. This work will occur during 2011. Also, in 2011, a study will be undertaken to reset the thresholds for both surveys. In order to achieve this, a sub-sample of those records which pass selective editing will be selected and manually edited. This will enable us to investigate whether the errors which are having an impact are captured using the existing thresholds.

14. For more detailed information on the methodology for RSI and MBS selective editing refer to papers by Hooper & Lewis, 2010 and Silva, 2009.

IV. Use of SELEKT for the Annual Business Survey

A. Background to Annual Business Survey (ABS)

15. The ABS, formerly the Annual Business Inquiry (ABI/2), is the UK equivalent of the European Structural Business Survey (SBS). It is a large, complex, expensive survey which samples UK businesses and is split into seven industrial sectors – Catering, Motor Trades, Production & Construction, Property, Retail, Services, and Wholesale. The ABS covers approximately two thirds of the UK economy and collects information from UK businesses on a number of financial variables. Each of the different sectors can be viewed as a survey in its own right with survey specific questionnaires and edit rules. There are different 'form-types' which are used to collect detailed (long form) and less detailed (short-form) information in the different sectors.

B. Current ABS editing method

16. The ABS currently uses a series of both standard and user-defined edit rules to identify errors and suspicious values in the data. The edit rules vary depending on the sector that they are referring to. For example, in the Production and Construction sector there were approximately 50 edit rules applied to the survey responses in 2007.

17. The rules are a mixture of the usual types of traditional edits. They test for errors such as consistency and a few also refer to expected changes over time. Another rule that is applied to all ABS sectors is to flag the record if there are any comments written in the comments box in the questionnaire. This is to ensure that any comment made by a respondent is checked by someone in ONS. However this does create a lot of edit 'failures' when nothing on the form will be changed.

18. In 2005, a change was made to the editing strategy for ABS, some of the smaller respondents were excluded from manual editing, these being businesses with less than £250k turnover. Therefore a prerequisite was put into the system before the edits were run that filtered out those businesses with returned turnover less than £250k. The change was brought in as an efficiency saving because previous testing using selective editing could not be implemented effectively. However, many of these smaller respondents were then picked up in the macro editing run by the results teams which resulted in manual editing. Therefore, these businesses were not excluded from the investigation into using SELEKT for ABS, this means that savings are more difficult to achieve.

C. Use of Snowdon-X to evaluate current edit rules

19. Snowdon-X is a generic tool, developed by ONS and Southampton University, to assess the impact of changing the edits in any business survey (Silva et al, 2008). The tool outputs a number of quality measures such as estimates of bias resulting from missing errors, edit failure rates, hit rates and savings. The tool can be used to understand how current edits are working within the survey and also the impact on quality of any changes to the edit rules.

20. For testing purposes data from 2007 was initially used. Before the selective editing tool, SELEKT, was applied to the ABS the traditional rules were tested using Snowdon-X. SELEKT uses these rules to calculate a measure of suspicion (see paragraph 23). As mentioned, the PC sector had 50 traditional edits, in 2007, being applied to the returns. This was resulting in an overall failure rate of 70.2%. On investigation, it appeared that some of the rules could be refined without impacting on data quality. The list of rules was reduced and revised without an impact on the quality of the estimates of key variables. For more information see Lewis, 2011. Once the new set of rules had been decided, SELEKT could be implemented.

D. Evaluating the efficiency of SELEKT in the Annual Business Survey

21. SELEKT is a tool for selective editing, developed by Statistics Sweden (Norberg & Arvidson, 2008). The tool is very flexible and seems to be an effective methodology to apply to the challenges of the ABS. The ABS is more complicated than many business surveys for a number of reasons. Firstly, there are a large number of variables, which we have overcome by selecting key variables in consultation with users. Also, the rotation of businesses out of the sample is relatively high. Most businesses taking part in the survey in a particular year will not have taken part in the previous year; therefore we do not have previous values to use as predicted values.

22. SELEKT is deemed to be appropriate for the ABS because of its flexibility. SELEKT uses a general scoring function which is calculated for each unit, variable and domain. The scoring function can be split into three parts: suspicion, impact (on estimates) and importance.

23. The suspicion of a unit (for a particular variable) relates either to the traditional edit rules for the survey or test variables. It is very important for the suspicion measure to ensure that the edit rules that are included are correctly specified, that is the reason Snowdon-X was used before the application of SELEKT. The traditional edit rules result in a constant suspicion, which lies between 0 and 1, if all edit rules pass then the suspicion is equal to 0. In the application of SELEKT to the PC sector all hard edits are given a suspicion value equal to 1. These are rules such as balance edits which pick up when components do not sum to a total. It is also possible to use traditional query edits (or soft edits) and fix the suspicion to equal the hit rate of that edit, which is calculated in Snowdon-X. Test variables compare returned values with other variables related to the variable being tested. An example of a test variable for the PC sector is:

$$\frac{\text{Sales of Merchant Goods}_i}{\text{Total Turnover}_i}$$

24. The test variable is usually based on the ratio or difference between the variable value being tested and another variable value, or a corresponding value from the register or previous survey round. Upper and lower limits are calculated for the test variable, either using time series data or using cross sectional analysis, where quantiles can be calculated within pre-specified homogenous edit groups. This alleviates the problem of having little past data to calculate predicted values. A suspicion value, between 0 and 1, is then calculated for each test variable. Each survey variable can have more than one test variable, and the maximum suspicion is used.

25. The impact measure is similar to the score used for the RSI and MBS (see equation 1) and is a measurement of the potential impact on the estimate if the value is in error. It is calculated as:

$$Potimp_{j,k} = w_k(y_{j,k} - \tilde{y}_{j,k}) \quad \text{(Norberg et al, 2010)} \qquad (5)$$

where $y_{j,k}$ is the unedited value of variable y_j for unit k

$\tilde{y}_{j,k}$ is the predicted value of variable y_j for unit k

and w_k is the sampling weight for unit k

26. The anticipated impact is then the product of the suspicion measure and the potential impact.

$$Antimp_{j,k} = Susp_{j,k}.Potimp_{j,k} \quad \text{(Norberg et al, 2010)} \qquad (6)$$

27. Moving on to importance, the importance weight allows the user to attach greater importance to particular variables and domain classifications.

28. In order to calculate the local score the anticipated impact (equation 6) is multiplied by the importance weight. The scores are then globalised, step by step, over domains, variables and observations to the respondent unit.

29. The rest of this paper will concentrate on one sector of the ABS, Production & Construction (PC). The key variables, identified by users for the PC sector of the ABS were:

- Gross value added at basic price (GVA);
- Net capital expenditure;
- Total purchases of goods, materials and services;
- Total employment costs;
- Turnover; and
- Total taxes

30. The first five variables in the list are considered key variables for all sectors of ABS; however total taxes was included as a secondary variable for the PC sector as it is seen as important to users. These key variables are given a variable importance weight of 1. Other variables are given an importance weight of 0.1, so that their contribution to the local score is less.

31. SELEKT allows us to consider the final output for several domain classifications, for example if a respondent contributed to more than one publication table. For the ABS PC sector the key domain was 3 digit Standard Industrial Classification (SIC). Therefore this was the only classification, and there was no need for classification importance weights.

32. The importance measure is also used to standardise the score. A parameter is available which gives flexibility in the score function to compute the impact in absolute terms or relevant to the size of the domain. In this application of SELEKT for the PC sector we have used the Absolute Relative Bias (as for RSI and MBS), which uses the total as the denominator (see equation 3), although the standard error can also be used.

33. For more details of the SELEKT method, see Norberg and Arvidson, 2008.

V. Results of SELEKT Testing

34. An initial investigation and testing of parameter options was carried out on the PC sector of ABS. From this initial research the best potential savings, in terms of the number of questionnaires being manually edited was **22%,** with all domains having an absolute relative bias of less than 1%. This saving is based on the following scenario:

- Comments box comments are not flagged for editing – these can either be checked only for the questionnaires that fail selective editing, or additional comments can be checked by the results team when they are investigating the ‘story’ of the final estimates. Another option is to remove the comments box completely, as has been done for RSI and MBS, however this option needs to be investigated further.
- All fatal, or hard edits, e.g. consistency checks are applied to all units and units that fail these rules are flagged as an error, regardless of their impact on results.
- All query, or soft edits have been replaced by continuous suspicion measures, which will be used in conjunction with the impact and importance measures to score the unit against a threshold. The continuous suspicion measures have been applied to all units, not just those over £250k, as was the case in the previous ABS editing.
- Some 3 digit SIC domains needed to be combined due to small sample sizes.

VI. Limitations & Next Steps

35. This analysis is still work in progress and will need further investigation and refinements before final estimates of savings are produced. As a result of this study we are going to continue to analyse the results on both the PC sector and also the Catering sector (CA) of ABS. There are a number of parameter options still to consider and some sensitivity analysis will be run in order to investigate whether more savings can be achieved.

36. The current savings estimate is only based upon the long form questionnaires, and needs to be extended to include the short forms. Re-introducing the validation to all units (i.e. all businesses with less than £250k turnover) may have more of an impact on the short forms, if there are more of these type of businesses in the short form sample. Also, this savings estimate is only based upon the 2007 data at present, and needs to be tested on more recent data to determine if the parameters and thresholds are robust.

37. Once parameters have been set, and finalised , for the PC and CA sectors, the other sectors will be tested with a view to implementing SELEKT into the ONS system in 2012.

38. Finally, ONS are also investigating the use of administrative data for editing and imputation purposes (Lewis, 2011). Over the next few years, as more administrative data becomes available for ONS' use there is a possibility of utilising this for editing purposes within SELEKT.

V. References

Hedlin, D., 2008, *Local and global score functions in selective editing*, Conference of European Statisticians, Work session on Statistical Data Editing.

Hooper, E. & Lewis, L., 2010, *The application of selective editing to the ONS Monthly Business Survey*, European Conference on Quality in Official Statistics, Helsinki, 4-6 May, 2010.

Lewis, D., 2011, *Evaluating the benefits of using VAT data to improve the efficiency of editing in a multivariate annual business survey*, UNECE Work Session on Statistical Data Editing, Ljubljana, 9-11 May 2011.

Norberg, A., & Arvidson, G., 2008, *New Tools for Statistical Data Editing*, UNECE Work Session on Statistical Data Editing, Vienna, 21-23 April 2008.

Norberg, A.., Adolfsson, C., Arvidson, G., Gidlund, P., & Nordberg, L., 2010, *A General Methodology for Selective Data Editing*, Statistics Sweden, not published.

Silva, P.L.N., 2009, *Investigating selective editing ideas towards improving editing in the UK Retail Sales Inquiry*, European Establishment Statistics Workshop.

Silva P.L.N., Bucknall R., Zong P. and Al-Hamad A., 2008, *A generic tool to assess impact of changing edit rules in a business survey – an application to the UK Annual Business Inquiry Part 2*, UNECE Work Session on Statistical Data Editing, Vienna, 21-23 April 2008.