2009 EUROPEAN ESTABLISHMENT STATISTICS WORKSHOP

DATA EDITING – THE EDEN PROJECT

Introduction

Business survey data are collected and edited to identify and correct suspect data values where necessary. Editing is the most resource intensive part of business surveys in the UK accounting for approximately 35 to 40 per cent of the total survey costs. Due to this there has been a desire to improve the effectiveness of our editing processes.

In many statistics offices the approach to editing has been piecemeal. Surveys have different procedures and even within a particular survey changes are made over time on an ad hoc basis. This means that some edit checks can be redundant or do not add value. Conversely results analysts generate queries because key movements or anomalies are missed at the initial editing stage.

The Eden project aims to introduce a new customer focused editing strategy for business surveys. The new strategy will ensure that data sets are fit for purpose. It brings a pro-active approach to survey work and managing efficiencies. It will replace the reactive approach, which historically has reduced samples or extended validation gates and resulted in receiving poor customer feedback.

The Eden project aims to take a radical approach to surveys. It is based on the Australian experience which delivered efficiencies and improved quality with limited resources and using existing systems. Like the Office for National Statistics (ONS) in the UK, the Australian Bureau of Statistics (ABS) was facing demands for more timely, coherent and relevant data. They recognised that in order to meet the demands, a new editing strategy was needed; a strategy that uses new and more efficient editing techniques such as significance editing and a better balance between micro and macro editing. Like ONS, they were faced with a number of operational constraints as well as the task of changing the culture underpinning their approach to editing. Their solution has been to develop an Editing Guide for staff responsible for directing, managing and directly editing statistical or administrative collections.

This paper describes the early UK work on taking forward this new approach. Two surveys, a short period one, the monthly retail sales survey and a structural survey the annual business inquiry which collects financial variables were chosen. Progress on the former that shows the approach and impact of the introduction of selective editing is described paper by Silva (2009).

The work looks at all the stages of editing starting from micro editing which detects anomalies through the analysis of unit record data and the subsequent resolution of those anomalies. This includes initial editing, selective editing and automatic editing and imputation. It also embraces macro editing that detects anomalies of aggregate data and prepares and clears data for release.

Business Objectives

1. Understand user needs and priorities
The strategy is heavily focused on customer needs and priorities. They are part of the process from the very initial stages and they define where the focus should be of resources and quality.

2. Improve quality of outputs

Editing efforts are concentrated on the highest priority outputs that fit the user needs whilst maintaining quality to a professional standard. The strategy will ensure that we fully understand the customer needs before editing starts. Resources are concentrated where it makes difference to customers and can show a measurable improvement

3. Increase job satisfaction for staff

Staff are clear on the goals of the editing strategy and have better understanding of the data and what is important to their customers. They will be able to take pride in producing fit for purpose outputs and concentrate their efforts on where it really makes a difference. A closer relationship with their customers gives a better understanding of the purpose of the outputs and how their work contributes to making a higher quality product.

4. Introduce an evaluation process

The editing strategy is reviewed and evaluated once a year to ensure that data are fit for purpose and to identify improvements for the next round. This creates an ethos of continuous improvement learning lessons from experience.

5. Reduce resources needed for processing business surveys

We will only focus on anomalies which will have the most impact on the quality of the final outputs. This can be achieved through better understanding of where and how the data is going to be used.

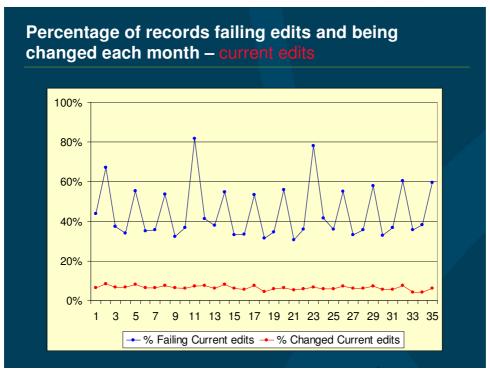## 6. Reduce the administrative burden on businesses

The process will reduce the contact with providers. For most providers there should be no need for contact to clarify data. If contact is necessary it should be as soon as feasible after the data is collected.
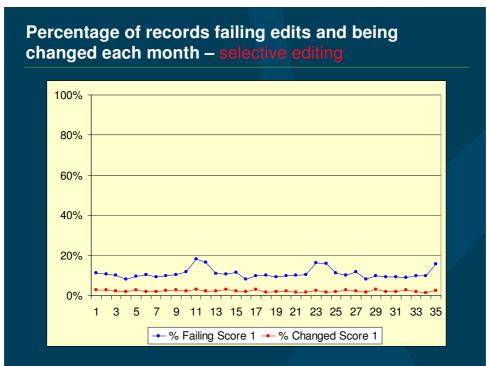
7. Spread best practice and create common vocabulary.

In many areas there are already examples of good practice which the Eden Project team can transfer to other areas as they take on other surveys. Additionally, as the project is rolled out across surveys a common vocabulary for editing and validation can be developed.

Potential of the Eden Approach

The potential of this approach can be demonstrated by the work of Silva on the retail sales survey. At present the micro editing approach has a high failure rate of records of around 40 per cent which rises to 60 per cent on quarter months when employment information is also collected. However only around 10 per cent of records are changed in the editing process. By working with users to establish their priorities and introducing selective editing techniques for turnover the failure rate has been substantially reduced.

Percentage of records failing edits and being changed each month – current edits

Percentage of records failing edits and being changed each month – selective editing

Development of an editing strategy for the Annual Business inquiry

The annual business inquiry is a large structural business survey collecting a wide range of financial information. Results are produced at a low level of detail with there being over 400 industries at four digit level of the economy. The survey collects a large number of financial variables – a standard questionnaire asks for 45 data items – with the longest forms having over a hundred questions. Efficiency measures over recent years have forced the survey team to make difficult decisions which included

reducing the sample by ten thousand and widening some of the gates used in editing. These changes were made quickly with little involvement form users and were not well received. A standard approach to looking at efficiencies in editing has been taken as far as it can by removing any redundant checks and widening the gates as far as possible. Nonetheless the costs of running this survey are the largest of any within ONS and efficiency pressures and weaknesses of the process remain. The Eden approach is being taken forward as a radical way forward to improve the operation of the survey working closer with users.

The starting point was a visit from Eden Brinkley and Lynne Bismire from ABS who spent two weeks in ONS in November 2007. Using the Annual Business Inquiry Part 2 (ABI/2) as an example, they carried out workshops and seminars demonstrating how their approach can benefit ONS and how it can be adopted to meet ONS needs. This training involved a wide cross section of those involved in the survey from staff involved in data collection and validation, producing results, national accounts analysts and support staff working in methodology and IT.

A formal two person project was set up in the spring of 2009 with a dedicated project manager/statistician with administrative support to work together to take forward this work. The main aims over the first year was to engage with users to gain a better understanding of their needs and priorities and to educate them and the ABI staff on the process to get buy-in for the new strategy.

The ABI User group which meets every six months provided a helpful vehicle to inform users promote the work and answer their concerns and questions. The project received a very positive response as users felt that their views were being listened to and that they had an opportunity to influence this development from the outset. Individual meetings were held with users to establish how they used the data and to establish their needs and priorities discussing and challenging their requirements. This was a lengthy process with a diverse and large number of key customers which included national and regional accounts in ONS and a number of government departments. In general it required a number of meetings with each to come up with an agreed set of priorities.

The process has helped prioritise the editing process. The priority for industrial detail is at a higher level of the economy which equates to the level of the supply/use tables which equates to a little over a hundred groups. These aggregates at two to three digit level of the economy will save much detailed checking. There is the possibility of adding a few lower level aggregates should key policy needs arise. We have similarly been able to focus on a smaller number of variables where five key ones have been identified – gross value added at basic prices, net capital expenditure, turnover, total purchases and total employment costs.

We are planning in the autumn of 2009 to build on the work done on the RSI to develop selective editing at the macro editing stage for these five variables. As this development is breaking new ground we will be conducting a pilot exercise. We have taken a copy of the datafiles prior to the commencement of the validation and editing. The provisional set of results will be processed in the usual way and published before the end of 2009. Once that results process is complete the original datafiles will be edited on the new basis for two sectors. We have chosen the motor trades and catering industries as these are two areas where it is relatively easy to account for the change in the industrial classification which we are handling at the same time. The lessons leant from this work are intended to give confidence to a wider roll out in 2010 on the 2009 survey which we hope will cover most industries. Thereafter in the annual

reviews of the survey processing the process will be refined to meet any changed user needs and lessons learnt.

Our vision ahead
ONS aims to build on the work for these two surveys and extend the roll out to other business surveys in future years. This will take forward the seven key priorities for business change

1. Regular liaison with users as the editing plans are reviewed each year focusing on their needs and priorities
2. Improved quality for users as editing is focused on where it makes a difference for them
3. Staff will have greater job satisfaction as they are closer to users and can see how their work makes a difference
4. There will be an annual evaluation which builds on learning lessons and takes account of changing user needs
5. We expect to save resources tailoring the checking to the most important anomalies
6. The burden on companies will be lower as this will result in less recontact.
7. As the project is rolled out through the central team there will be a more generic approach and best practice will be shared

Ole Black
Head of Economic Surveys Division
Office for National Statistics
UK

References
ABS (2007) The Editing Guide: validating data from businesses. Canberra – Australian Bureau of Statitics

Silva P L N investigating selective editing ideas towards improving data editing in the UK retail sales Inquiry Internal report for ONS

Silva P L N investigating selective editing ideas towards improving data editing in the UK retail sales Inquiry – paper presented to EESW09