Journal of Official Statistics
Vol. 10. No. 4, 1994, pp. 437–447

Significance Editing in the Australian Survey of Average Weekly Earnings

David Lawrence and Chris McDavitt[1]

Abstract: Editing survey returns is a substantial component of the cost of running sample surveys. This paper presents a method for reducing the resources allocated to editing. The method is based on estimating the effect on the survey estimates of resolving the edit queries for each unit. This method has been introduced to the Australian Survey of Average Weekly Earnings. As a result the number of respondents contacted to resolve edit queries has been more than halved without affecting the quality of the final statistics. This has reduced the total respondent burden and also the overall costs of conducting the survey. It has also enabled an earlier publication of the survey results. This paper gives a description of the technique, and also presents results from an evaluation study that was conducted prior to its implementation.

Key words: Data quality; editing; respondent burden; respondent re-contact; survey costs.

1. Introduction

In recent times methodologists in government statistical agencies have looked critically at the overall survey process instead of focussing mainly on the sample design as has been the case in the past. With, in general, only small gains to be made from exploring the more sophisticated sample design options, effort has been concentrated on quantifying and reducing non-sampling errors, improving the efficiency of survey processing and decreasing respondent burden. In Australia, Linacre and Trewin (1989) looked at the broad overall question of optimising the allocation of resources across the survey process in order to control non-sampling errors, to the extent possible. One survey operation, the editing of survey returns, is the focus of this paper.

Not so long ago the goal of editing was generally seen as identifying and correcting all (or as many as possible) of the respondent errors that could be found in the survey returns. Only recently have questions been raised about the value of editing and how much editing should be done. Pullum, Harpham, and Ozsever (1986) reported that in the World Fertility Survey the editing had little noticeable effect on the estimates or the analysis of the data other than delaying the release of the results by about one year. Latouche and Berthelot (1992) looked at scoring individual survey

[1] Statistical Services Section, Australian Bureau of Statistics, GPO Box K881, Perth, Western Australia 6001.
Acknowledgement: The authors wish to acknowledge Farwell (1991) upon which this work was based.
The opinions expressed in this paper are those of the authors and do not necessarily reflect the views or policies of the Australian Bureau of Statistics.

returns based on their potential effects on the estimates. They showed that in a trial application in the Canadian Annual Retail Trade Survey, respondent re-contacts could be restricted to units with higher scores with only minor effects on the survey aggregates.

This paper looks at a method of estimating the effect of resolving edit queries on the final survey estimates. In a sense this can be viewed as setting edit tolerances dynamically by examining the magnitude of the error that would result from not resolving a certain set of edit queries, rather than, say, choosing the tolerance for an edit to achieve a pre-determined failure rate. In this way editing resources can be reduced by ignoring queries of trivial value and resources can be concentrated on ensuring the more important edit queries are properly resolved.

This process, which we have called *significance editing*, has been applied in the Australian Survey of Average Weekly Earnings (ASAWE). An evaluation study was performed prior to its introduction. For this evaluation study both the original raw (unedited) survey returns and the final survey data were captured and stored for two successive collection periods. A *score function* (to estimate the effect of resolving the edit queries) was developed and the effect of the significance editing strategy was evaluated by replacing the final survey data with the original raw data where edits had been deemed by the strategy not to be worth pursuing. Estimates were then formed for various amounts of respondent re-contact to evaluate the effect of the strategy.

This paper describes the significance editing process in the context of the ASAWE, and presents results from the evaluation study. Section 2 describes the ASAWE and the traditional editing approach. Section 3 describes the significance editing strategy. Section 4 reports results from the evaluation study and Section 5 discusses the production implementation of the strategy in the ASAWE.

Although this paper is written in the context of the ASAWE, the concepts can readily be extended to a wide range of business surveys collecting quantitative data. Currently work is in progress within the Australian Bureau of Statistics to implement this editing methodology in other business surveys.

2. The Australian Survey of Average Weekly Earnings and Traditional Editing

The ASAWE is a quarterly survey comprising a sample of approximately 5,000 business enterprises drawn from a population of approximately 600,000 enterprises. The survey, which is mail based, covers all sectors of the Australian economy except for agriculture. Overlapping samples are employed with quarterly sample rotation between 5% and 10%. The primary aim of the survey is to measure changes in average earnings for Australian wage and salary earners. Information is collected on total employment and total wages paid in the selected firms, split by sex and full-time/part-time employment status. Estimates of total wages and total employment in each category are formed using ratio estimation (with total employment as maintained on the register of businesses used as the benchmark variable), and estimates of average earnings are calculated from these totals. Only the average earnings estimates are published. Change in average earnings is estimated as the simple difference in two quarterly average earnings figures. The population is classified into eight states or territories, twenty broad industrial groups and two sectors (private and public). The survey has been designed to meet standard error constraints at the state level. Finer level figures are considered to be a by-product. Because of the relatively small sample size any non-response is vigorously pursued, whether item or unit non-response. Non-response rates are very low (consistently under 2%).

In this paper we use the term *editing* to refer to the process by which survey returns are checked for possible errors. This definition includes re-contacting respondents where suspect data are identified and the resulting amendment of data where errors are found. The editing process is divided into input and output phases. The input editing phase involves checking individual survey returns in isolation from other responses in the current processing cycle. The output phase involves identifying possible errors by examining the full set of survey returns for the current cycle as a whole. The editing process involves generating edit queries (which identify suspect responses) according to a set of edit rules and then re-contacting the respondent to either confirm or amend the reported data.

This paper concentrates on the input editing phase. In the ASAWE a series of *fatal* edits is applied to all survey returns. These are triggered if the data that are entered into the computer are logically inconsistent or any of the data items are not reported (e.g., a non-zero wage figure is reported, but there is no reported employment). After applying fatal edits the survey returns are split into two streams: units new to the survey (between 5% and 10% of units each cycle) and continuing units.

For new units a series of non-fatal *query* edits are applied based on expectations of average earnings for units in similar industry groups. All edit queries are pursued by telephone.

The continuing units are subject to *historical query* edits based on previously reported figures. Our editing model for continuing units looks for changes relative to previously reported figures.

A typical historical query edit used in the ASAWE has the form

$$\left|\frac{z_{ti}}{x_{ti}} - g\frac{z_{t-1,i}}{x_{t-1,i}}\right| < \alpha \qquad (1)$$

where z_{ti} and x_{ti} are the current reported values for the numerator (wages) and denominator (employment) variables, respectively, for the ith unit. Similarly, $z_{t-1,i}$ and $x_{t-1,i}$ are the previous quarter's values for the ith unit.

Here, α is an edit tolerance which may be a fixed number or a percentage of $z_{t-1,i}/x_{t-1,i}$. In general the edit tolerance α is set both for a minimum absolute change and a minimum percentage change relative to the previous figures.

A growth factor, g, is applied to the average earnings figure in the previous quarter. Essentially our edit model assumes there will be little change in wage rates between quarters. Historically we have observed a quarterly growth rate in estimates of average earnings of around 2%. This is equivalent to using a growth factor of 1.02.

3. Overview of the Significance Editing Process

Historically, all edit queries generated in the input editing phase, as described above, have been pursued. The significance editing process is concerned with estimating the effect on the survey estimates that would result from resolving an edit query. Because we do not know what the amended figures will be, we impute an expected amended figure, using the same assumptions as in our traditional edit model. For continuing units this means applying a growth rate to the figures supplied in the previous quarter. We then calculate the change to the estimates that would occur if the reported figures were replaced by the expected amended figures for that unit. A score function is used to estimate this change in the survey estimates. So that the editing process will proceed without delay the score function is required to be independent of information supplied in the current cycle by any units other than the unit under consideration.

The ASAWE estimates are ratios, of the form $\hat{Z}/\hat{X}$ where $\hat{Z} = \sum w_i z_i$ is an estimate of total earnings and $\hat{X} = \sum w_i x_i$ is an estimate of total employment. It is straightforward to show that if the original reported figures for unit i are replaced by the expected amended figures the change in the estimate of average earnings is given by

$$S_i = \frac{w_i(\hat{X}\delta z_i - \hat{Z}\delta x_i)}{\hat{X}(\hat{X} + w_i \delta x_i)} \tag{2}$$

where z_i and x_i are the responses for the numerator and denominator variables, respectively, for the ith unit (here wages and employment, respectively).

Here, δz_i and δx_i are the differences between the reported values and the imputed values for variables z and x, respectively, for the ith unit.

The ith unit's estimation weight is denoted by w_i and $\hat{Z}$ and $\hat{X}$ are the population estimates of numerator (wages) and denominator (employment), respectively.

The S_i represents the change in the estimate of average earnings that would result from replacing the figures on the current survey return for the ith unit, with the expected amended figures for that unit.

We use equation (2) as our score function where S_i is the score for the ith unit.

In practice the estimates $\hat{Z}$ and $\hat{X}$ in equation (2) are not known at the editing stage, as they are the final survey estimates. In our trial we found the score function was not sensitive to small changes in their values, and as such we were able to model these estimates readily from previous figures (great accuracy is not needed).

Estimators of different functional forms would produce different score functions. For instance, an estimate of a total $\hat{X} = \sum w_i x_i$ would suggest a score function of the form $w_i \delta x_i$.

We applied the significance editing procedure only to continuing units with non-fatal edit queries. The fatal edits occur infrequently and they are regarded as sufficiently important to be resolved in all cases. For new units we felt there is value in contacting the respondents by telephone once to ensure understanding of data item definitions and to clarify procedures for completing the form. As our editing model for continuing units looks for changes relative to previously reported figures, it is important that the initial figures are carefully checked, otherwise errors can go undetected for long periods. Thus we allow a high rate of respondent re-contact for new units.

For continuing units the score function is used to determine which units are subject to re-contact during editing. Where the respondent is not re-contacted the original survey data are accepted without alteration.

There are several aspects to the implementation of the process. The first is the method of imputing the expected amended figures. Our traditional edit model has the underlying assumption that there will be little change in average earnings rates between quarters. We use the same assumption to produce the expected amended figures for use in the score function (2). The expected amended figures are formed by applying a growth factor g to the previous quarter's figures. In the ASAWE there is a very strong correlation between the earnings figures and the corresponding employment figures, so we calculate our growth factor on the basis of changes in the average earnings figures. Of course separate growth factors could be used for numerator and denominator variables if desired.

Individual unit scores can be calculated to reflect the effect of resolving the edit queries on the estimate at various levels of disaggregation. For instance, each unit's score could reflect the expected effect of resolving the edit queries on the estimate for the appropriate state or territory, or the national estimate. Each unit could be expected to have a larger effect at a finer level of disaggregation, than on the broader estimation levels to which it contributes. One aspect of the evaluation trial was to study the effects of calculating unit scores based on differing estimation levels.

In order to apply the significance editing technique to reduce the amount of respondent re-contact and total editing load, it is valuable to combine the scores associated with individual variables to produce one global score for the survey return. In practice the respondent would then be re-contacted only if the global score exceeded a predetermined cut-off value. Otherwise the figures would be accepted without query. The ranking of units by their global score effectively gives a priority order to units for edit resolution.

The unit score for a particular variable is called the local score. It may not be necessary to calculate these local scores for all survey variables. For example, when strong correlations exist between variables (e.g., individual components and a total) scoring all variables may be redundant. We chose five of the nine published ASAWE variables as variables of interest for calculating local scores.

Next we wanted a method of combining the local scores for each selected variable of interest to give one global score for the survey return. All the published estimates in the ASAWE are in the same units (Australian dollars) and have similar relative standard errors, so we chose to calculate the global score as the maximum of the absolute values of the local scores. Thus a significant query for any variable will always be pursued. Alternatively, global scores could be formed by using functions of weighted local scores if, say, some variables were considered to be more important than others, or required greater accuracy.

Local scores would need to be transformed into a single unit of measurement if different variables are measured in different units. For instance, if it was sought to combine local scores for a variable measuring total employment and a variable measuring total wages paid, it would be appropriate to multiply the local score for employment by an average earnings figure to produce a score comparable to that for total wages.

4. Significance Editing Trial

A trial was undertaken to test various factors such as the choice of variables of interest, use of growth factors, the effectiveness of our method for imputing the expected amended values, and to allow the setting of cut-off score values. To do this, survey data were captured for two successive collection periods (May and August 1991) both in the raw form (as reported on the survey returns) and final form (after all editing had been performed). The significance editing strategy was applied by *scoring* each record on the raw data file that had been flagged for query action (excluding fatal errors and the small

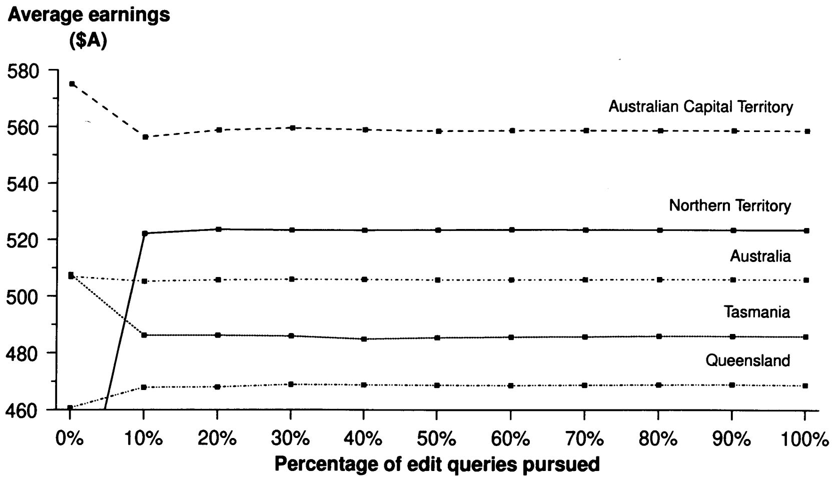

Fig. 1. Average ordinary time earnings for full-time adult females by percentage of edit queries pursued. Selected states and Australia

number of new units – which we recommended would always be contacted). The units were then ranked in descending order of their global score, and for each decile of the score a composite file was formed. For units with scores above the decile value, figures for the composite file were taken from the final data file (i.e., edited data values were used); otherwise figures were taken from the raw data file. Figures for new units and units failing fatal edits were automatically taken from the final data file. Estimates were then formed from each of the composite files. The zero percent decile represented the results if only the fatal errors and new units were pursued. The one hundred percent decile represented the final published results and was used for comparative purposes.

The trial was used to fine tune elements of the process. For example, one cut-off value could be set for the survey, or separate cut-off values could be set for each state, or for each industry. That is, the unit scores could be calculated based on their effects at differing estimation levels. We settled on calculating the effect of editing on estimates at the state level. This is the finest estimation level routinely published (although figures at finer levels of disaggregation are available on request) and the survey is designed to achieve design standard error constraints specified at the state level. Different relative standard errors are achieved in each state, and we chose to match the magnitude of error induced by the significance editing procedure proportionately to the level of sampling accuracy obtained in each state. We found that calculating unit scores at a finer level of disaggregation implicitly controlled the effects on broader level estimates.

Figure 1 shows the results for point estimates for one variable – average ordinary time earnings for full time adult females, at the state level. Latouche and Berthelot (1992) define, logically, the *absolute pseudo-bias* for an estimate at the qth percentile as

$$\text{Absolute pseudo-bias} = \left| \frac{\hat{Y}_q - \hat{Y}_{100\%}}{\hat{Y}_{100\%}} \right| \tag{3}$$

where $\hat{Y}_q$ represents an estimate, of any form, formed from the composite file for the qth decile. Here $\hat{Y}_q = \hat{Z}_q / \hat{X}_q$ for point

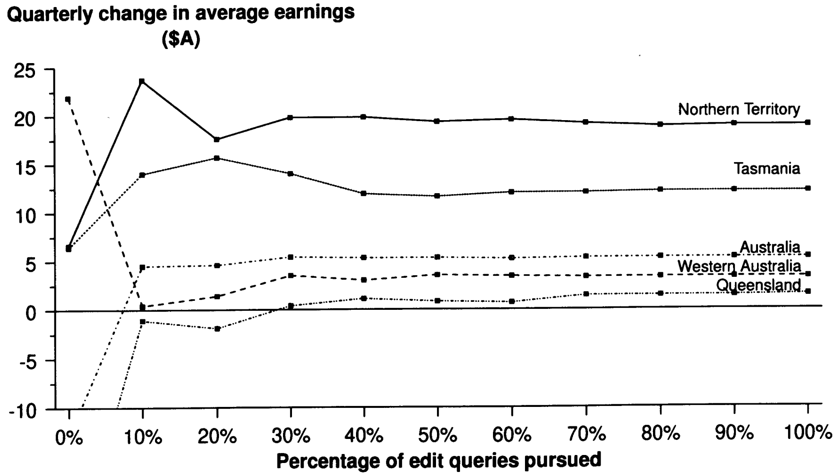

Fig. 2. Quarterly change in average gross earnings for full-time adult males by percentage of edit queries pursued. Selected states and Australia

estimates or $\hat{Y}_q = \hat{Z}_{qc}/\hat{X}_{qc} - \hat{Z}_{xp}/\hat{X}_{qp}$ for estimates of quarterly change. The $\hat{Y}_{100\%}$ represents the final published result formed when all editing is performed.

In this case, it can be seen from Figure 1 that the absolute pseudo-bias is negligible from the 10% decile onwards. Similar results for the absolute pseudo-biases were obtained for all variables in both quarters tested.

We used data from two quarters to evaluate the effects on estimates of quarterly change. These are the more important results of the survey, and are used as economic indicators. Figure 2 shows the estimates of quarterly change obtained by using the composite files obtained in both quarters for each of the deciles for one variable – average weekly earnings for full-time adult males. We found that for this variable, and similarly for each of the other variables, the absolute pseudo-bias does not become entirely negligible until the 40% decile.

For estimates of quarterly change, which can often be quite small, the absolute pseudo-bias can be misleading. We define the *relative pseudo-bias*, to compare the pseudo-bias with the standard error, rather than the magnitude of the estimate, as

$$\text{Relative pseudo-bias} = \left| \frac{\hat{Y}_q - \hat{Y}_{100\%}}{SE(\hat{Y}_{100\%})} \right| \tag{4}$$

where $SE(\hat{Y}_{100\%})$ represents the standard error associated with the estimate $\hat{Y}_{100\%}$.

The relative pseudo-bias has the advantage of comparing one measure of the non-sampling error associated with this editing strategy with a measure of one other source of error inherent in the survey.

The relative pseudo-bias for estimates of quarterly change at the finer industry level for selected deciles is shown in Table 1 for one variable – average ordinary time earnings for full-time adult males. Once again, similar results were observed for the pseudo-biases from each of the other variables. Again by the 40% level of editing the relative pseudo-bias has fallen to small proportions. It can be seen that at the 40% level of editing the largest relative pseudo-bias was 0.17. That the differences in the estimates are so small at this level of editing suggests that a large percentage of

Table 1. Relative pseudo-bias for estimates of quarterly change. Average ordinary time earnings for full-time adult males by industry

Industry	Percentage of edit queries pursued					
	0%	10%	20%	40%	60%	80%
Mining	8.7096	0.2643	0.1271	0.0031	0.0645	0.0000
Food, beverages & tobacco	7.2928	0.7485	0.4066	0.0915	0.0465	0.0176
Textiles, clothing & footwear	0.1408	0.5795	0.3110	0.1042	0.0171	0.0171
Paper & printing	0.2866	0.4862	0.4862	0.0943	0.0000	0.0000
Basic chemicals & fuel	2.9224	0.1052	0.1641	0.0136	0.0190	0.0000
Basic metals	1.1913	0.1616	0.1616	0.1670	0.0381	0.0082
Fabricated metal	2.4264	0.3903	0.1170	0.0308	0.0002	0.0003
Motor vehicles & transport equipment	34.8250	0.4541	0.4125	0.0982	0.0429	0.0006
Miscellaneous manufacturing	0.4731	0.2139	0.1920	0.0918	0.0100	0.0000
Electricity, gas & water	8.2644	0.3087	0.3737	0.0521	0.0339	0.0037
Construction	0.2370	0.0843	0.1357	0.0258	0.0029	0.0003
Wholesale trade	0.0541	0.2809	0.1494	0.0045	0.0015	0.0012
Retail trade	1.1132	0.0375	0.0145	0.0106	0.0051	0.0145
Transport & storage	1.9052	0.2131	0.0087	0.0278	0.0117	0.0002
Communication	178.9765	20.3588	0.1400	0.1400	0.1400	0.0764
Finance, property & business services	3.8104	0.0570	0.0307	0.0028	0.0001	0.0007
Public administration & defence	2.8416	0.2828	0.0868	0.0282	0.0203	0.0000
Community services	4.5340	0.1865	0.1161	0.0239	0.0079	0.0039
Recreational & personal services	0.2923	0.1599	0.0001	0.0166	0.0046	0.0007

the edit queries could have been ignored without affecting the inferences that would be made from the survey results.

We also evaluated the effect on estimated standard errors for both point estimates and estimates of quarterly change. If respondent errors occurred at random it is possible that their cumulative effect could cancel out, but their presence (when edit queries go unaddressed) could increase the variability of the data. For point estimates, standard errors calculated from the composite files were within one percentage point of the final (100%) figure from the 20% decile onwards, at both state and industry levels. For estimates of quarterly change, standard errors were within one percentage point of the final values from the 40% decile onwards at the state level and within five percentage points of the final values at the finer industry level.

The sensitivity of the procedure to the growth factor used for the imputation of the expected amended figures was tested by using excessive growth rates – of $\pm 20\%$ in a single quarter. In the ASAWE series quarterly changes in average earnings have rarely exceeded 2%. Estimates were calculated in the same way as above from newly formed composite files for each decile using new expected amended values. The results were barely different from those obtained previously. Table 2 shows the extent of the changes in individual unit rankings (for the global scores) when a growth rate of -20% was applied. This is equivalent to a growth factor of 0.8. In Table 2 it can be seen that only

Table 2. Changes in individual unit rankings May quarter, 1991 when data is scored using a growth factor of 0.8 for imputation

	Percentage of units								Aust (%)
	State								
	NSW (%)	WA (%)	SA (%)	Vic (%)	NT (%)	Tas (%)	Qld (%)	ACT (%)	
Change in percentage ranking									
No change	3.84	12.68	11.95	4.97	6.29	8.26	4.63	9.90	6.34
< 0.5% change	12.05	10.24	0	10.69	0	0	5.79	0	7.90
0.5%–1% change	10.89	5.85	11.41	6.46	14.17	11.57	6.56	10.81	9.36
1%–2% change	13.46	10.73	12.50	13.18	14.96	9.09	15.05	14.41	13.15
2%–5% change	30.00	29.26	29.34	25.87	28.34	34.71	25.48	26.12	28.55
5%–10% change	17.30	17.56	15.76	25.12	18.11	19.83	22.39	18.91	19.50
10%–20% change	10.64	11.70	15.76	9.95	14.96	11.57	15.05	14.41	12.06
20%–50% change	1.79	1.95	3.26	3.48	3.14	4.95	5.01	5.40	3.06
50%–100% change	0	0	0	0.24	0	0	0	0	0.04
Total records	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

15% of records changed ranking by more than 10% with this excessively small growth rate, and this would only flow through to a change in the estimate for a given decile where units cross the decile boundary. The procedure was also found to be insensitive to small changes in the modelled values of the estimates $\hat{Z}$ and $\hat{X}$.

One point of concern was the possible effect upon the output editing load (where possible data errors are detected by examining the data in aggregate). The various output editing processes used in the ASAWE were simulated on the 40% decile composite file for both study quarters. We found that in each quarter only a half-dozen survey returns were subject to scrutiny at the output editing stage where they would have been passed over at the input editing stage had the 40% decile been used to set the cut-off value for input editing. By estimating the effect upon estimates of resolving edit queries, the significance editing process takes in some of the concepts used in our normal output editing processes, so this result is not entirely surprising.

One important question is whether this process would have a cumulatively deleterious effect over time. If a set of edit queries was passed over in one quarter this could suggest the data used in producing the following quarter's expected amended values for editing would be less accurate. To investigate this we created composite files for each decile for our second quarter of study data by taking previous values for use in generation of the expected amended values from the 40% composite file from the first study quarter. Once again we produced estimates from each composite file and found no noticeable differences from the previous figures. In fact we found that only 4% of units changed their ranking by two percentage points or more, when compared to the rankings formed by taking imputed values from the final survey data for the first study quarter.

One factor contributing to this result was the relative inefficiency of the old editing system. We found only one out of every three units raising edit queries had an amendment made to any data items following query action. Our score function turned out to be a good predictor of which edit queries were likely to result in an amendment to data. As such at the 40% level the significance editing strategy was suppressing many edit queries that would turn out to be unfruitful.

5. Implementation of Significance Editing in the Australian Survey of Average Weekly Earnings

For the production implementation, five of the nine survey variables were chosen as variables of interest for scoring. Scores were calculated representing effects on state level estimates and cut-off values were chosen to match the 40% decile from the trial. This produced different cut-off values in different states roughly proportional to the design relative standard errors. Expressing the relative pseudo-bias as a percentage, we found the highest value (at the 40% decile) was 30% (for an industry level estimate of quarterly change). At state level the highest relative pseudo-bias on the estimates of quarterly change was 20%. For point estimates none of the relative pseudo-biases exceeded 20%. In general the pseudo-biases were much lower than these extremes.

The choice of the 40% decile for setting the cut-off values was radical in terms of the amount of editing resources saved (over 50% of the total input editing load was cut), but conservative in its effects given that the additional non-sampling errors likely to be incurred by the significance editing procedure compared to the traditional procedure were much less than the survey standard errors. Significance editing may, in fact, improve the data quality overall because of the way it provides a priority ordering for edit failures, which ensures the most important queries receive full attention. If there is any bias in the traditional editing procedure it is towards the model assumption of little change in figures between quarters. Now that fewer units are being edited it is possible that there has been a decrease in this bias.

The significance editing strategy was first implemented in the ASAWE in the third quarter of 1992. At the time of writing, data from eight successive quarters for the ASAWE have been collected and processed using the strategy. In that time the number of edit queries pursued each quarter has been very close to the 40% of the number generated by the edit rules as recommended from the trial (without adjustment of the cut-off values) indicating the stability of the data in the ASAWE series. There has been no noticeable change in the figures in the series nor have there been any adverse effects encountered during the processing. Resource savings of between 3 and 4 staff years have been achieved (from the approximately 18 staff years used to run the survey).

6. Conclusion

The pursuit of edit queries is a way of addressing one of the many sources of error inherent in a sample survey. Traditionally, many have viewed the goal of editing as the production of a complete and error free survey data file. This goal is unrealistic and often undesirable. In the ASAWE we know, for instance, that there is a limit on the accuracy of the survey returns provided by some respondents due to limitations in their record keeping. With significance editing we have set out to ensure the errors incurred due to accepting selected survey returns without query were kept smaller than the errors coming from other sources – both sampling and non-sampling errors.

Instead of setting fixed edit tolerances to achieve a pre-determined failure rate, say, edits are triggered according to the likely effect on the survey estimates of resolving the edit queries. In this way editing resources can be reduced by ignoring queries of trivial value. The technique reduces respondent burden by reducing the amount of respondent re-contact required. It also conveniently provides a priority ordering for edit queries so that the most important queries can be addressed most rigorously. It has allowed the publication of results sooner after the survey reference date than was previously possible, and it has freed up resources to be directed towards other important quality issues associated with the survey. These improvements have been achieved without adversely affecting the quality of the final results.

The principles implemented here for the ASAWE should be extendable to a wide range of sample surveys. Any survey collecting quantitative data would be a candidate. Where the distribution of the population is skewed, as is the case in most business surveys, the score function should be an effective discriminator. Regular ongoing surveys with overlapping samples collecting stable data items, like the ASAWE, allow for simple imputation of the expected amended values. Even if these conditions are not met, the significance editing technique can still be applied. The basic principle underlying the technique is to take an editing model and use it to evaluate the effect on the survey estimates of changing individual survey returns to fit the model. This can be applied to a wide range of editing models. An empirical study, along the lines we have performed for the ASAWE, can be used to identify how useful the editing is in a particular survey, and to answer the question of how much editing should be done. There is little value in editing to resolve queries that have trivial effects, and we suggest that it is appropriate to allow a small residual error component from leaving some editing queries unresolved. The permissible size of this component should be determined in comparison with the errors, both sampling and non-sampling, from other sources.

7. References

Farwell, K. (1991). Fundamentals and Suggestions for Ranking and Scoring Edited Unit Responses. Unpublished technical report, Australian Bureau of Statistics.

Latouche, M. and Berthelot, J.M. (1992). Use of a Score Function to Prioritize and Limit Recontacts in Editing Business Surveys. Journal of Official Statistics, 8, 389–400.

Linacre, S.J. and Trewin, D.J. (1989). Evaluation of Errors and Appropriate Resource Allocation in Economic Collections. Proceedings of the 1989 Annual Research Conference, U.S. Bureau of the Census, 197–210.

Pullum, T.W., Harpham, T., and Ozsever, N. (1986). The Machine Editing of Large Sample Surveys: The Experience of the World Fertility Survey. International Statistical Review, 54, 311–326.

Received December 1993
Revised November 1994