Journal of Official Statistics
Vol. 8. No. 3, 1992, pp. 389–400

Use of a Score Function to Prioritize and Limit Recontacts in Editing Business Surveys[1]

Michel Latouche and Jean-Marie Berthelot[2]

Abstract: In order to reduce survey costs, large statistical agencies intend to concentrate respondent follow-ups to suspicious units that may have an important effect on the estimates. This paper first presents a summary of the overall strategy for recontact and follow-up in economic surveys. Then, it discusses and compares three score functions used to classify suspicious units according to their potential effect on the estimates. Finally, a trial application of a score function to the Canadian Annual Retail Trade Survey is presented.

Key words: Data collection and capture; data quality; follow-up; recontact; respondent burden; survey costs.

1. Introduction

Over the past few years, many large statistical agencies have devoted much effort to reducing respondent burden, increasing the efficiency of their production process and improving the quality of their end products. To respond most effectively to these objectives, a rethinking of some survey operations was necessary. This led to the creation of working groups or committees in these organizations. For example, the United Nations' Development Programme Statistical Computing Project has established the Data Editing Joint Group (1982); the United States' Federal Committee on Statistical Methodology has formed a subcommittee on data editing and another on data collection (Federal Committee on Statistical Methodology 1990a and 1990b); and, Statistics Canada has established the Generalized Survey Function Development project (Colledge 1987).

One of the functions which was addressed, in rethinking the production process at Statistics Canada, is the data collection and data capture operations. This function consists of all activities required in the acquisition of survey information, its validation and conversion to a form suitable for subsequent automated processing (Generalized Survey Function Development Team 1989).

[1] An earlier version of this paper was presented at the International Conference on Measurement Errors in Surveys, Tucson, Arizona, USA, November 12, 1990.

[2] The authors are Senior Methodologists, Business Surveys Methods Division, Statistics Canada, 11th floor, Section A, R.H. Coats Building, Ottawa, Ontario, Canada, K1A OT6. FAX (613) 951-1462.

Acknowledgements: We would like to acknowledge support from the Generalized Survey Development Project of Statistics Canada, which allowed us to initiate this work. We would like to thank the following persons from Statistics Canada: P. Whitridge, S. Perron, J. Morabito, and B. Downer from the Business Survey Methods Division; J.-F. Gosselin, L. Boucher, and J. Beauchamp from the Head Office Operations Division; and R. Rasia, M. Rivest, and D. Roeske, from the Industry Division, all of whom made this study possible.

The collection and capture process has always been time and labour intensive. In particular, the most expensive task within it is the *recontact* of the sample unit for error correction of data in the case of partial nonresponse, and the *follow-up* of total nonresponse. Until now, the usual strategy for follow-up and recontact has been to contact every sample unit which either does not return the questionnaire, returns an incomplete questionnaire, or returns a questionnaire with suspicious data as judged by the edit rules. This is very expensive, time consuming and results in a considerable respondent burden. In the United States federal statistical agencies, the median editing costs for economic surveys are reported to be 40% of the total survey cost (Federal Committee on Statistical Methodology 1990a). Furthermore, there is no evidence that such an extreme strategy is necessary to obtain good quality end products. For example, in the World Fertility Survey (Pullum, Harphman, and Ozsever 1986) the editing had no discernible effect on estimates other than to delay their release by about one year.

In the development of an effective recontact and follow-up strategy, we have to minimize the amount of resources used without affecting significantly the overall data quality and timeliness of the survey. Linacre and Trewin (1989) have worked on the general topic of optimizing the allocation of resources to reduce all aspects of nonsampling errors. Our work is more closely related to optimizing the collection process rather than addressing the allocation of resources. The strategy that we are putting forward promotes a selective recontact and follow-up strategy. This strategy is twofold. Firstly, a minimum follow-up effort is performed to ensure that at least the status (in-scope or out of scope) of the sampling unit is obtained for proper imputation at a later stage. Secondly, the recontact resources are concentrated on suspicious units that may have a significant effect on the estimates. In order to ensure the consistency of the data, the remaining respondents and nonrespondents are handled by an automatic edit and imputation system (Kovar, MacMillan and Whitridge 1988). The second step of the strategy is in accordance with Granquist (1987) who stated that consistency should be achieved quickly and practically without ever referring to the physical questionnaires. It is believed that a selective recontact and follow-up strategy reduces operational costs and improves timeliness without affecting the quality of the estimates substantially. To this end, considerable effort has gone into the development of a score function that helps identify the respondents that need to be recontacted for error correction.

This paper presents the results of the research which led to the development of a score function. The concept presented in this paper takes advantage of two characteristics specific to business surveys. First, business surveys collect mainly quantitative data. Second, the variables of interest often have nonuniform or highly skewed distributions so that a relatively small number of units contribute a large percentage of the total estimate. Finally, the score functions presented in this paper assume the availability of historical information from periodic surveys or of administrative data sources.

Section 2 presents the concepts that lead to the creation of three score functions that are described in Section 3. Section 4 presents a comparison study of these functions using a subsample of the 1987 Canadian Annual Retail Trade Survey (CARTS), while Section 5 discusses the results of a trial application of the most suitable function to the CARTS.

2. Score Function

In the context of a business survey, the score function has to include four major elements: the size of the responding unit, the size and number of suspicious data items in the respondent questionnaire, and relative importance of the variables.

The size of the responding units according to a variable of interest is a good indicator of the influence it may have on the estimates. Such an indicator can usually be obtained from the current capture data, from a previous cycle of the survey or from administrative sources.

A measure of the error introduced by the suspicious data is another useful indicator of the influence of the respondent. For a given variable, if the respondents with the largest errors are recontacted, it will be useless at some point to recontact more respondents since the effect of the remaining errors is minor (Greenberg and Petkunas 1987). The size of the errors can be approximated using trend or discrepancy between current reported values and historical data. Note that the estimation weight also has to be considered when measuring the size of an error.

Similarly, the number of suspicious values on the questionnaire is another factor for which the score function should account. Priority should go to the recontacts that allow us to verify the largest number of suspicious values.

The fourth element is the relative importance of the collected variables, as determined by the subject matter specialists and methodologists. Some variables may be considered as essential and require more attention than others. In addition, variables that cannot be efficiently and effectively imputed are better candidates for recontact. Finally, the response rate also has to be considered. A low response rate at some disaggregated level may increase the need for recontact at that level in order to improve data quality or insure that enough responses are available when one plans to use donor imputation.

In addition to these four elements, the score function should consider operational aspects. It should be easy to implement, and should be flexible enough to suit different surveys. Also, the formula for the score function must have a simple logical interpretation so that its application can be justified to subject matter specialists.

Another important operational factor is that the use of the score function must not delay the survey process. The inputs to the function therefore should not depend on the flow of responses. They should rather be based on prespecified parameters, possibly based upon data from prior survey cycles, and on data from the current cycle provided only by the respondent being processed.

Lastly, it would be useful to have a function which produces scores that would have similar distributions for the same variable regardless of the level of aggregation. In that way, the parameters in the function could be computed at a higher level of aggregation. In addition, similarity of the distributions would insure that, for the same variable, all strata or other levels of aggregation are treated equally.

3. Description of Three Score Functions

The score function assigns a relative score to each respondent. The function uses as input the characteristics of the respondents that are related to the potential effect they may have on the estimates. The function must be able to combine all factors mentioned above. A score is computed for each variable in the questionnaire, and these are summed up to give a global score to the

respondent. The respondents with the highest score values are considered to be the most influential and they should be recontacted with high priority.

Three score functions have been developed: RATIO, FLAG and DIFF. Each of them places emphasis on different elements or operational constraints. RATIO tries to produce a similar distribution of the score values among publication cell, while FLAG puts the emphasis on simplicity. Finally, DIFF tries to reconcile simplicity and similarity of the distributions.

The first function, called RATIO, is derived from the work of Hidiroglou and Berthelot (1986), and Miller and Carpenter (1982). It is based on the ratio of the current reported value and the final value after processing for the previous cycle.

Suppose that one deals with a periodic survey which collects I variables, and that the population is divided into P cells for publication purposes. Say that there are K_p respondents in cell p.

Let $y_{k,i,t}$ be the value reported by respondent k ($k = 1, 2, \ldots K_p$) within cell p ($p = 1, 2, \ldots, P$) for the variable i ($i = 1, 2, \ldots, I$) at time t;

$y'_{k,i,t-1}$ be the final value for the respondent k for variable i at time $t - 1$;

$r_{k,i,t}$ be an estimate of the error given by

$$r_{k,i,t} = \frac{y_{k,i,t}}{y'_{k,i,t-1}};$$

$MDR_{.,i,t-1}$ be the median of the $r_{k,i,t-1}$ computed at the cell level using data from time $t - 1$ and time $t - 2$.

Unfortunately the multiplication of $r_{k,i,t-1}$ with the importance of the respondent ($y_{k,i,t}$) is not a good discriminator. In order for the score function to have a large value when the error is very large or very small, we apply the following transformation

$$s_{k,i,t} = \begin{cases} \left| \frac{r_{k,i,t}}{MDR_{.,i,t-1}} - 1 \right| \text{ if } r_{k,i,t} > MDR_{k\ldots t-1} \\ \left| 1 - \frac{MDR_{.,i,t-1}}{r_{k,i,t}} \right| \text{ otherwise} \end{cases}$$

Now, $s_{k,i,t}$ can be multiplied by a factor measuring the importance of the respondent, based on the value of the variable as well as the survey weight $w_{k,i,t}$

$$g_{k,i,t} = s_{k,i,t} \times w_{k,i,t} \times (\mathrm{MAX}(y_{k,i,t}, y'_{k,i,t-1}))^U.$$

The $\mathrm{MAX}(y_{k,i,t}, y'_{k,i,t-1})$ is used in order to handle partial nonresponse. The exponent U ($0 \leqslant U \leqslant 1$) provides a control on the importance associated with the magnitude of the data. The parameter U is not very sensitive and the same value can be used for many variables of a survey (Granquist 1990). Based on the results obtained by Lalande (1988) and by Bilocq and Berthelot (1990), a value of 0.5 is used. The relation becomes

$$g_{k,i,t} = s_{k,i,t} \times w_{k,i,t} \times \sqrt{\mathrm{MAX}(y_{k,i,t}, y'_{k,i,t-1})}.$$

Finally, to make the cell to cell distribution more uniform, we use the score function given by

$$f_{k,i,t} = \frac{|g_{k,i,t} - MDG_{.,i,t-1}|}{IRG_{.,i,t-1}}$$

where

$MDG_{.,i,t-1}$ is the median of the $g_{k,i,t-1}$ computed at the cell level using previous cycle data, and

$IRG_{.,i,t-1}$ is the interquartile range of the $g_{k,i,t-1}$.

The global score for the respondent is given by

$$\mathrm{RATIO}_{k,.,t} = \sum_{t=1}^{I} f_{k,i,t} \times z_{k,i,t} \times \hat{v}_{.,i,t}$$

where

$z_{k,i,t}$ =

$$\left\{\begin{array}{l} 0 \text{ if } y_{k,i,t} \text{ is accepted by the editing process} \\ 1 \text{ if } y_{k,i,t} \text{ is labeled as suspicious} \end{array}\right\}$$

$v_{.,i,t}$ is a weight related to the importance of the variable i.

Because of the parameters $MDR_{.,i,t-1}$ and $MDG_{.,i,t-1}$, RATIO also requires data from time $t - 2$. In contrast, the next two functions need only information from time t to $t - 1$.

FLAG, the second score function is the simplest one. It has been developed in an attempt to give prominence to the most important variable of the questionnaire, as specified by the subject matter specialists. Let J be the index of this variable, then FLAG is defined by

$$\text{FLAG}_{k,.,t} = w_{k,J,t} \times \sqrt{\text{MAX}(y_{k,J,t}, y'_{k,J,t-1})} \times \sum_{i=1}^{I} z_{k,i,t} \times v_{.,i,t}.$$

The square root comes from the parameter U of the RATIO score function.

The last score function, DIFF, is a compromise between the complexity of RATIO and the simplicity of FLAG. It emphasizes the absolute discrepancy between the current reported value and the released values of the previous cycle. This difference is weighted by the total $\hat{Y}_{.,i,t-1}$ (at a given level) for that variable from the previous cycle. The result is subsequently multiplied by the error flag $z_{k,i,t}$. The sum over all values gives the score for a record

$$\text{DIFF} = \sum_{i=1}^{I} \times \frac{w_{k,i,t} \times |y_{k,i,t} - y'_{k,i,t-1}| \times z_{k,i,t} \times v_{.,i,t}}{\hat{Y}_{.,i,t-1}}.$$

4. Comparison Study

In order to compare the three score functions, the data collection and data capture process was simulated using data from the 1987 Canadian Annual Retail Trade Survey (CARTS). This survey is a census of retailers that have Total Net Sales and Receipts greater than or equal to 1 million Canadian dollars. The survey collects twelve financial variables such as sales and inventory. The population is classified into 18 trade groups and 12 provinces or territories. Estimates are produced for every cell made from combinations of trade group and province/territory.

For the purpose of the study, the raw data from 2054 establishment questionnaires containing 12 continuous variables were recaptured. These questionnaires were from the food products sector and the motor vehicle equipment/manufacturing industry in addition to all questionnaires from the provinces of Prince Edward Island and Alberta. Of these records, 196 were removed as they were considered to be out of scope. Since it was the first attempt, and we did not have any idea about how the score functions would perform, it was decided to simplify the set up by assuming that the twelve variables had the same importance weight ($v_{.,i,t} = 1$).

As well as the 1987 reported data, the final data from 1985–87 were also used. The parameters $MDR_{.,i,t-1}$, $MDG_{.,i,t-1}$, $IRG_{.,i,t-1}$ and $\hat{Y}_{.,i,t-1}$ were derived using 1985–86 data. The 1987 final data, as released by subject matter review, were used as a control comparison to indicate the efficiency of the score functions. After a responding unit has been identified as being suspicious (at least one suspicious datum point detected by the editing system), a recontact is possible. All units with suspicious data were recontacted in cells with less than 10 sample units. In addition, follow-up was performed for all total nonresponses and potentially out of scope units (i.e., units reporting less than one million dollars total net sales and

receipts). For the remainder of this paper, follow-ups and recontacts within small cells will be referred as to automatic follow-ups. With those automatic follow-ups aside, the remaining units in error may be recontacted according to their score function values.

Follow-ups and recontacts were simulated by replacing all the 1987 reported data of a questionnaire flagged for recontact by the corresponding 1987 released data. This was based on the assumption that one recontact could correct all of the errors in a given questionnaire. It was recognized that this assumption may be somewhat unrealistic but was necessary for the simulation.

A measure estimating bias due to response error was defined for comparing the three score functions. For any given number of recontacts, an estimate $\hat{Y}_{.,i,87}$ of the total for the full sample was computed using 1987 reported values for non-recontacted units, and 1987 released values for units that were recontacted or automatically followed up. Comparison of the three functions was essentially based on the behaviour of the absolute relative discrepancy between $\hat{Y}_{.,i,87}$ and the total $\hat{Y}'_{.,i,87}$ that was released in 1987. This absolute relative discrepancy constituted the "absolute pseudo-bias"

$$\text{absolute pseudo-bias} = \frac{|\hat{Y}_{.,i,87} - \hat{Y}'_{.,i,87}|}{\hat{Y}'_{.,i,87}} \times 100.$$

In Figure 1, the absolute pseudo-bias is plotted versus the number of recontacts. On the x axis, the recontacts are ordered from the most to the least influential according to the value of a particular score function (high score values being more influential). The algorithm used for this ordering sorted units within a cell as well as according to their overall score value. As a result, when a score function performs well, selecting a fraction of the ordered sequence of units would ensure that approximately the same percentage of units were recontacted in every cell.

For instance, Figures 1A, 1B and 1C show the relation between the absolute pseudo-bias for the variables GROSS COMMISSION, OTHER OPERATING REVENUE, and TOTAL NET SALES AND RECEIPTS. Two different patterns occurred; for the less frequently reported variables such as GROSS COMMISSION and OTHER OPERATING REVENUE, the graphs revealed an expected pattern: the pseudo-bias decreased as the number of recontacts increased. For these variables, the DIFF score function seems better than the two others. The slope generated by DIFF has the largest magnitude indicating the fastest decrease of the pseudo-bias as the number of recontacts increases.

The most frequently reported variables such as TOTAL NET SALES AND RECEIPTS (Figure 1c) gave results that were more difficult to analyze. For these variables, the use of only automatic follow-ups leads to a small pseudo-bias. This pattern occurs because the most important variables are used to identify total non-response. Consequently, all the functions gave good results, and the shapes of the curves were irregular. When one let the score functions handle the automatic follow-ups as shown by Figure 1d, similar curves as in the case of less reported variables were obtained, but FLAG and RATIO surpassed DIFF. However, most of the first 300 recontacts generated by FLAG and RATIO were total nonrespondents. With these results and the requirement of simplicity of explanations, it was decided to keep automatic follow-ups and to use the DIFF score function for the recontacts.

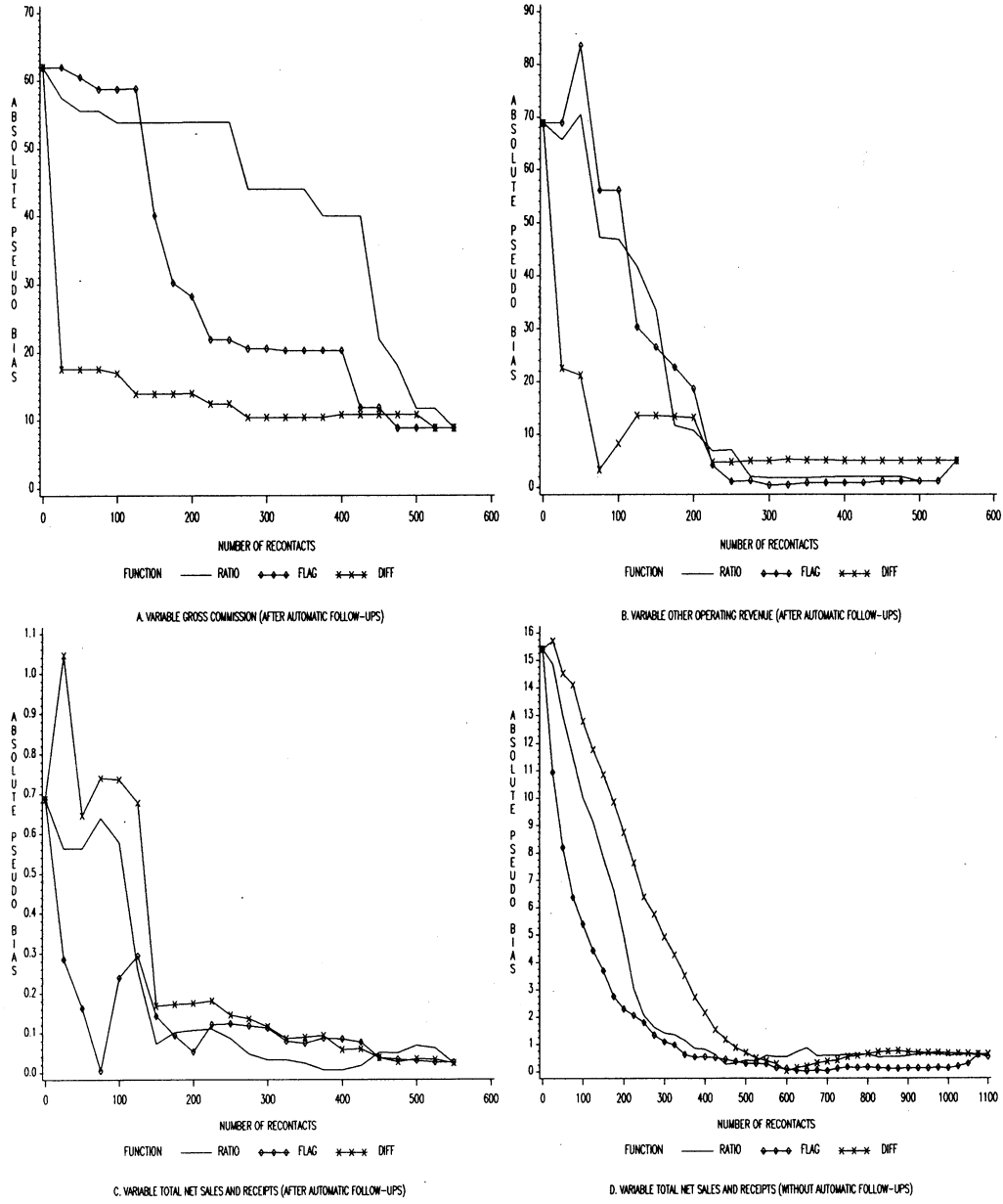

Fig. 1. Absolute pseudo-bias (in percent) versus number of recontacts (ordered by score function)

5. Trial Application

In a regular production environment, a score function critical value would be used to identify units to be recontacted. This critical value would be computed using previous cycles of the survey, in a way to approximately achieve a predetermined recontact rate. In this manner, the records could be processed as soon as they are captured, so no delay is imposed on the survey

process. The erroneous records with a score function larger than the critical value would be recontacted, while the other ones would be imputed.

In order to evaluate the effect of the score function concept in a production environment, different recontact rates using DIFF were applied to the same sample described in Section 4. In this trial application, no critical score function values were computed; instead, proportions of erroneous records were used. The recontact percentages considered were 0, 17, 34, 50 and 100. Zero percent of recontacts was considered to show the results when only the automatic follow-ups were performed. One hundred percent was considered to show the results when all errors detected were recontacted, as is currently done for most surveys. The 17%, 34% and 50% rates were chosen as possible alternatives to either 0% or 100% rate. The important premise here is that the cost of extra recontacts may be unnecessary if there is only a small potential gain in the accuracy of the estimates. In this trial application, the data were not imputed; records that would normally be imputed according to the strategy remained as captured.

With the out of scope records removed, there were 1858 records remaining in the sample. A total of 967 records failed at least one of the edit rules. Of these 967 records, 397 satisfied the automatic follow-up criteria. From these automatic follow-ups, 245 were total nonresponse, 20 were erroneous units in small cells, and 132 potential out of scope records remained in the sample after follow-up and correction were made.

The five different recontact rates for the remaining 570 suspicious units determined the rest of the recontacts. When these rates were converted to actual numbers of recontacts, they corresponded to 0 (0%), 97 (17%), 194 (34%), 285 (50%) and 570 (100%). This recontact rate was approximately the same for all combinations of trade groups and provinces contained in the sample.

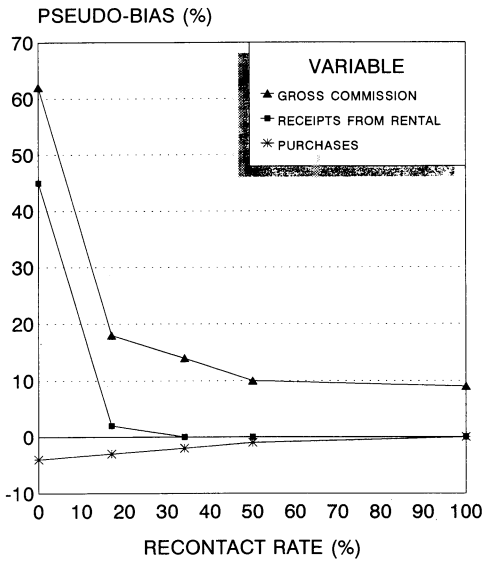

Fig. 2. Pseudo-bias versus recontact rates for three different variables

Figure 2 illustrates the overall pseudo-bias for three different variables: GROSS COMMISSION, RECEIPTS FROM RENTAL, and PURCHASES. For all recontact rates, the results were encouraging with initial pseudo-bias of lower magnitude for the frequently reported variables such as PURCHASES.

The critical result from Figure 2 is the following. The reduction in absolute magnitude of the pseudo-bias in going from 0% to 17% of recontacts is much greater than in going from 17% to 34% or any other percentage. With a 17% recontact rate the pseudo-bias for GROSS COMMISSION has already been reduced to 18% and only improves to 14% when applying a 34% rate. These 91 extra recontacts improved the estimate but not substantially. A similar pattern was observed for most variables. As

Table 1. Means and standard errors of the pseudo-bias computed on 47 cells for 5 recontact rates for total net sales and receipts

Recontact rate (%)	Pseudo-bias means	Pseudo-bias standard errors
0	−0.002	0.021
17	−0.003	0.021
34	−0.001	0.009
50	−0.001	0.009
100	−0.002	0.009

expected, the pseudo-bias was the lowest when all errors had been recontacted (100% rate). The pseudo-bias is not always 0 with 100% of the flagged errors recontacted because some additional errors can slip through the edits. Subsequent to data processing, errors may have been identified by subject matter analysis and corrected at that time. That is, some misunderstanding errors can only be found by macroediting after data collection (Granquist 1984). Future budget constraints may prevent this rate of recontact from being a logical alternative.

Further comparison of the five rates was required at a level beneath the global level described above in order to reach appropriate conclusions. As a result, the pseudo-bias was considered for each of the 47 trade group by province combinations (cells). Table 1 shows for each recontact rate, the mean and standard error of the cell pseudo-bias for TOTAL NET SALES AND RECEIPTS.

The means of the cell pseudo-bias do not change significantly from one rate to another. However, one sees a significant decrease in the standard error when changing from 17% to 34% of recontacts. On the other hand, little is gained in going from 34% to 100%.

The results suggested the use of either a 17% recontact rate if aggregate levels are of primary interest or a 34% rate if disaggregate levels are deemed more important. Finally, it was decided to recommend a 34% recontact rate because of production considerations. In a production environment the observed pseudo-bias would be larger than those obtained in this trial application for two reasons. Firstly, recontacts could not all be successful, and secondly, the actual recontact rate could be less than the desired rate because it would be based on a predetermined critical value rather than the recontact rate itself.

Consequently, it was recommended to follow-up all nonrespondents, erroneous records in small cells and potentially out of scope units, as well as to recontact approximately 34% of the remaining erroneous records. The gain in quality of the estimates in going beyond this point seems to be an unnecessary expense. The remaining erroneous records should be corrected by an automatic edit and imputation system.

Table 2 shows the overall pseudo-bias resulting from the recommended 34% recontact rate and the degree to which the variable is reported. Because partial nonresponse was not completely corrected, almost all variables had a negative pseudo-bias. In general, each had a small pseudo-bias except GROSS COMMISSION and OTHER OPERATING REVENUE with 13.91% and −13.2%, respectively. Note that with a 100% recontact rate, the pseudo-bias of these two variables went to 9% and 5%, respectively. In order to improve the pseudo-bias of GROSS COMMISSION and OTHER OPERATING REVENUE, one could increase their importance weight, thus increasing the number of questionnaires having suspicious values for these variables to be recontacted.

For the variable TOTAL CLOSING INVENTORY, the pseudo-bias for the 47 cells varies from −5% to 1%, but almost all

Table 2. Pseudo-bias obtained with 34% recontact rate

Variables	*% Pseudo-bias*	*% Time reported*
Net sales	−0.63	100
Total net sales and receipts	−0.18	100
Opening inventory	−0.87	100
Closing inventory	−0.34	100
Purchases	−2.23	99
Salaries	−0.92	99
Receipts from repairs	−1.43	46
Non-operating revenue	−1.46	22
Other operating revenue	−13.20	15
Receipts from rentals	−0.08	10
Gross commission	13.91	5
Receipts from food services	2.68	2

of them have a pseudo-bias less than 2% which indicates very little variability in the distribution. This result, a pattern common to all variables when excluding a few outlying cells, is very encouraging. About 60% of the cells have zero pseudo-bias, largely due to the fact that all suspicious respondents that are in a small cell are recontacted. Since these recontacts have the reported data replaced by the final released data, they will have a zero pseudo-bias by design.

6. Conclusion

Although the simulation was run for only the Canadian Annual Retail Trade Survey, we believe it is representative of business surveys in general. The results obtained are encouraging. It is clear that recontact of at least some of the most influential units is advantageous with respect to the estimates obtained. Errors were usually investigated through these recontacts. More specifically, for a sample of 967 questionnaires with erroneous or missing data (52% of the full sample), a strategy with approximately 40% automatic follow-ups, 20% recontacts based on the score function, and 40% uncorrected brought the overall simulated estimates within 0.1% of the production estimates for the frequently reported variables. For smaller production cells level, the standard error of this discrepancy was not significantly reduced when more than one-third of the remaining errors were recontacted. For infrequently reported variables, quality could be improved by increasing their importance weights.

In a production environment, it is not always possible to correct all errors by follow-up and recontact. Some respondents always refuse to cooperate, or cannot be reached. Also, a single recontact is not sufficient to correct all errors. In addition, the use of a predetermined critical value to set recontacts would lead only to an approximate recontact rate. Therefore, it is preferable to use a conservative recontact rate as was recommended in Section 5.

The study has shown that recontacting a limited number of units can achieve about the same level of data quality as full recontact. While the pseudo-bias measure used here has limitations as a measure of response bias, it reflects what is achievable by the recontact and editing procedures of the survey. If this approach is used in production, the collection process becomes faster and valuable resources can be saved or reallocated to other purposes without significantly affecting the data quality.

The score function can be used in a number of different ways besides the one presented here. It can be applied without automatic follow-ups. Although, the automatic follow-up approach is good from the frame update point of view, selection of nonrespondents using a score function

would allow a major decrease in follow-up effort, while still providing good estimates. However, such an approach could cause bias in the estimates. The score function could also be applied to both follow-ups and recontacts to define a priority order. This order determines how often one should attempt to reach the respondent.

As a quality assurance tool, different recontact rates could be applied according to the score function class values. Future work would have to be done to define stratification of the scores and allocation of recontact rates.

Future investigation is also needed to determine the effectiveness of a critical value as a decision rule. It would also be interesting to know how the response rate can be included in the score function. This factor not only improves the score function for business surveys, but also helps to develop a score function for social surveys. Finally, the score function has to be modified in order to be able to handle qualitative variables.

7. References

Bilocq, F. and Berthelot, J.-M. (1990). Analysis on Grouping of Variables and on Detection of Questionable Units. Statistics Canada Working Paper No. BSMD-90-005E/F.

Colledge, M. (1987). The Business Survey Redesign Project – Implementation of a New Strategy at Statistics Canada. Proceedings of the 1987 Annual Research Conference, U.S. Bureau of the Census, 550–576.

Data Editing Joint Group (1982). Glossary of Terms. Working Paper Version 2, July 1982.

Federal Committee on Statistical Methodology (1990a). Data Editing in Federal Statistical Agencies. Statistical Policy Working Paper 18.

Federal Committee on Statistical Methodology (1990b). Computer Assisted Survey Information Collection. Statistical Policy Working Paper 19.

Generalized Survey Function Development Team (1989). Methodological and Operational Concepts in the Collection and Capture Module. Technical Report, Statistics Canada.

Granquist, L. (1984). On the Role of Editing. Statistisk tidskrift, 22, 105–118.

Granquist, L. (1987). On the Need for Generalized Numeric and Imputation Systems. Statistics Sweden, CES Seminar 23 R. December 1987.

Granquist, L. (1990). A Review of Some Macro-Editing Methods for Rationalizing the Editing Process. Proceedings of Statistics Canada Symposium 1990.

Greenberg, B. and Petkunas, T. (1987). An Evaluation of Edit and Imputation Procedures Used in the 1982 Economic Censuses in Business Division. In the 1982 Economic Censuses and Census of Governments Evaluation Studies, Washington, D.C.: U.S. Department of Commerce, 85–98.

Hidiroglou, M.A. and Berthelot, J.-M. (1986). Statistical Editing and Imputation for Periodic Business Surveys. Survey Methodology, 12, 73–83.

Kovar, J.G., MacMillan, J.H., and Whitridge, P. (1988). Overview and Strategy for the Generalized Edit and Imputation System. Statistics Canada, Methodology Branch Working Paper No. BSMD-88-007E.

Lalande, D. (1988). Système de détection des données aberrantes-SIO Documentation-MICRO. Technical Report, Statistics Canada.

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.

Miller, K. and Carpenter, R. (1982). The Feasibility Study on Error Detection Methods in the Edit System of the Full Civil Aviation Project. Internal Report, Statistics Canada.

Pullum, T.W., Harphman, 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 February 1991
Revised April 1992