DISCUSSION

Nash J. Monsour[1], Bureau of the Census
Washington, D.C. 20233

First, I would like to thank all the authors for four interesting papers that show that new survey developments are continuing at Statistics Canada and the Bureau of Labor Statistics. Hopefully, these improvements can be rapidly implemented in the agencies' surveys and censuses. Possibly the fact that there is little free market competition in the government survey area slows innovation somewhat. In the discussion that follows I shall try to relate these new survey developments to quality improvement.

In Bilocq and Berthelot's paper, "An Editing Scheme Based on Multivariate Data Analysis", we are given an excellent way of gathering variables for later editing in survey analysis. This is a good first step to the involved editing and imputation routines that are used in establishment surveys which collect data on many items. They use the SAS VARCLUS procedure to cluster related variables for editing and imputation so that only variables in the same cluster are used for editing against each other and for imputation of missing items. This eliminates many unnecessary edits.

It seems that the VARCLUS procedure has a good use as a check on the clustering of related variables done by survey practitioners based on their knowledge of the logical relations among the variables. Bilocq and Berthelot's examples seem to confirm such a use. This gives the survey analysts a mathematical way of checking and possibly correcting their intuition.

There are some variable gathering problems that occur when editing or imputing survey data for which Bilocq and Berthelot's methods can be expanded to handle. These are the instances where either none or only one of the variables in a cluster are reported by a survey respondent in the current survey period, and no information is available from previous periods for the cluster. In either instance how do we edit the data, and in the former instance how do we impute? In either situation we need to gather more variables for editing purposes. Possibly the VARCLUS routine could be adapted to choose the variables that are the next most closely associated with a given cluster. The above mentioned situations can often occur in a complex survey, e.g., when a respondent leaves one section of a survey questionnaire blank.

In "Address Register Research at Statistics Canada", Drew, Royce, and van Baaren show fairly conclusively that the use of an address register should reduce the undercoverage of the 1991 Census of Population in Canada. It is reasonable that the post-list method, which is more labor intensive than the pre-list method is the more accurate. If the register is successfully used in 1991, it may be used for other surveys. This will have the effect of improving the address register, as more different types of surveys and censuses use and update it, and of simultaneously making it cheaper to use by spreading the cost of maintaining it. Here, quality and productivity improvement go hand in hand.

There are some observations I would like to make about the creation of an all inclusive address register for Canada which seems to be the ultimate goal of Drew, Royce and van Baaren's research and a goal which seems laudable. First, such a complete register with names, address, telephone numbers, etc., formed from other available lists has connotations of "big brother". I think the public can be adequately protected by laws protecting the confidentiality of individuals and businesses. Moreover, having a centralized statistical system, as in Canada, confines the use of the list to one agency which offers further protection. Another observation is that having a centralized statistical agency makes it easier to form an address register in the first place. This is because administrative lists should be easier to obtain from other government agencies. Also maintaining all the country's statistical surveys and censuses yields other lists and more occasions to update the register. These occasions, as in the quinquennial census of population, can lead to actual field checking and updating of the address register. Furthermore, a centralized agency can not only standardize the addresses on the register itself, but attempt to have others, especially government agencies, to begin using a standard format for addresses.

In "Controlling Response Error in an Establishment Survey", Ponikowski and Meily have shown us a fairly simple way of identifying response errors in a periodic survey and removing them from the survey in later periods. Though their CATI record check method is applied to the Current Employment Statistics (CES) Survey of the Bureau of Labor Statistics (BLS), it is generaly applicable to other establishment surveys. Let's hope that the method is fully implemented in the CES.

Ponikowski and Meily's method, if implemented in a survey, is an excellent example of continuous improvement in reducing response errors. Though most establishment surveys have a system of edits that can identify large response and other errors and protect against the publication of distorted data, these systems cannot protect against many minor and/or consistent forms of response error. Many of us who are familiar with panel surveys are also

familiar with the old saying and/or practice about correcting persistent errors: if the errors are not too large, then don't correct them since they don't make any difference to the survey estimates; if the errors are large enough, then you can't correct them since correction will distort period to period changes in the survey estimates. Therefore, we are faced with a dilemma in which we usually do nothing. The authors have given us a way out of the dilemma. By continually applying Ponikowski and Meily's CATI record check to detect and correct survey response errors on a rotating basis through the life of the panel, we can improve the survey estimates or at least prevent their deterioration. I believe that more survey resources need to be put into the interface between the respondent or would-be respondent and the survey agency.

Several years ago Dr. Deming pointed out to me, at one of his quality courses, that as far as quality was concerned the collection of the data was the most important part of survey work. Being a mathematical statistican, I had always thought that survey design and estimation were the most important part of the survey. Only a few months ago, Martin Wilk put me down another peg when he said that mathematical statisticians' contributions to surveys were overrated. Also he thought the mathematical statisticians main contributions were introducing logic to the survey process not probability sample design and estimation. When Joe Garrett, the chair of this session, worked for me years ago, he also pointed this out to me saying he thought our work on the Census Bureau's business surveys was more that of a logician rather than a researcher or methodologist. It's time to apply our logical skills to improve the quality of the survey-respondent interface. Ponikowski and Meily have shown us a way to start improving.

Now turning to the last paper, "Developing a Cost Model of Alternative Data Collection Methods: Mail, CATI, and TDE", Clayton and Harrell have given us a good study of trying to obtain the best quality data for the buck in the CES survey. The paper is insightful in that they are looking ahead to the future cost of data collection.

I believe they should take a broader view of quality. They should investigate response error more. Maybe their study could be combined with Ponikowski and Meily's. This might change which methods of data collection are most cost effective given that you want low response error or that you want a mechanism to control response error. Also, Clayton and Harrell mention the use of voice recognition in data collection. I recommend further investigation of this, though it may be some years before voice recognition is ready for general survey use. Another method of data collection, that is currently in use by the Energy Information Administration (EIA) for its petroleum surveys, is computerized self administered questionnaires. Respondents who choose to use this method are sent a diskette to use on their PC. The respondent enters the data into a user friendly program, the data are edited, and when they are acceptable, the respondent dials up one of EIA's computers and transmits the clean data. This method may be worth considering for larger CES respondents. Note that self administered questionnaires as well as touchtone data entry are less expensive since the respondents enter the data directly into the computer database freeing the collection agency from this task. On another point, I believe that in the future, good data collection will mean incorporating editing as part of the collection process.

Now for some general comments on survey quality. As I have previously mentioned, no aspect of the survey process is more important than data collection when it comes to the quality of the final publication product of the survey. I believe that applying one of the newer product improvement techniques, quality function deployment, to the survey process would show that more resources should be devoted to the data collection function of a survey. It seems that a significant part of the later stages of the survey process consists of repairing problems that occur in data collection, problems like incorrect data or no data, problems that could be prevented by better data collection.

Another quality problem in repeated surveys is that, at times, more emphasis is put on consistency that on truth. I believe that this emphasis is a stumbling block when it comes to survey quality improvement. The over emphasis on consistency is just a more general way of looking at the problem of not correcting errors that I mentioned above. I am not saying consistency is unimportant, it is important. It seems that the over emphasis on consistency leads to the same problems that using non-probability samples or not sampling the entire population lead to: changes may occur in the population that are not accurately detected. One example of reliance on an artificial consistency is freezing the industrial codes of establishments in a survey until a new sample is drawn or a new census is taken. It is hard to update the industrial coding of establishments, but not having the correct codes will mask real changes in the population. But if we always came close to the truth in repeated surveys, then supposed inconsistencies would be meaningful changes which are what the surveys were designed to measure.

I have a final thought on survey data. Some have likened the data collected in a survey to the incoming material in manufacturing. This is a good analogy, but the survey data, I believe, is more important to the survey than the incoming product to the manufacturing process for some existential reasons. First, the survey data is irreplaceable,

whereas, in manufacturing, other materials can be substituted. Substitiution by imputation, proxy response, or reweighting always hurts in a survey. Secondly, in a survey the data always exists. It is always there for us to tabulate and analyze in many different ways. The incoming material, however, is changed into another thing in the manufacturing process. The fact that the data we collect theoretically can last forever, should give us pause and lead us to concentrate more effort on improving the quality of our data collection. Our countries deserve no less, we can do no more.

---

1 The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.