Floor Discussion

Paul S. Hanczaryk
Bureau of the Census

Mr. Michael J. Wenzowski of Statistics Canada presented a paper on the advances in automated coding at Statistics Canada. Statistics Canada has used automated coding software in a number of different applications since the original release of the Automated Coding by Text Recognition (ACTR) system in 1986. Automated coding systems produce significantly lower costs and fewer errors than manual coding operations. Mr. Wenzowski discussed a new release of the ACTR system, explaining the details of the system architecture and the coding process. The new software is well suited for both data capture operations (including CATI and CAPI) and high-volume batch applications.

Ms. Betty Perloff of the U.S. Department of Agriculture (USDA) presented the second paper on USDA's experience with computer-assisted food coding in its 1994 nationwide food survey. The Agricultural Research Service (ARS) conducts a survey, the Continuing Survey of Food Intakes by Individuals (CSFII), that measures the kinds and amounts of food eaten by Americans. The ARS developed an automated-coding system, Survey Net, that codes, edits, and manages food consumption data. Survey Net was instrumental in effecting survey efficiencies--the 1994 CSFII was released in less than half the time of the previous surveys.

Mr. Chad E. Russell of the Bureau of the Census, the chair of this session, emphasized the benefits that automated coding systems contribute to the Statistics Canada and USDA surveys. Mr. Russell cited the higher quality of code assignments, the reduction in analytical costs, and the improvements in the overall timeliness of the surveys.

Questions were raised on the quality measurement of the automated system coding. Michael Wenzowski stated that there is not a quality assurance component built into the ACTR system. However, subject-matter staff evaluate the resulting codes to identify systematic errors. Clearly, the quality of the resulting codes is essential to the success of any automated coding operation. Ms. Perloff explained that ten percent of the automated coding work is recoded by supervisory staff to test the reasonableness of the codes. Further, predetermined data checks are used to identify coding errors. Overall, there is less than a one percent error rate in the automated food coding system.

Someone raised the point that automated coding systems generally assign codes based on predominance in the event of combined activities. For example, an activity that included 70 percent babysitting services and 30 percent tutoring services would be coded to its primary activity, that is, babysitting services. Mr. Wenzowski reiterated that the ACTR system would assign the babysitting activity in every case. However, an enhancement could be developed to apportion combined activities properly at some later point. Ms. Perloff indicated that partial coding is critical in the CSFII program because every specific ingredient is measured.