

RESULTS FROM THE CURRENT POPULATION SURVEY CATI PHASE-IN PROJECT

Harland H. Shoemaker, Jr.
U.S. Bureau of the Census
Washington, DC 20233

KEY WORDS: CAI, data quality

INTRODUCTION

My paper presents results from the Current Population Survey CATI phase-in project. The paper is a compilation of results from numerous analyses of the data done over the past few years.

An abbreviated version of the paper is given here, containing only our most important findings. The reference section contains a complete listing of documents containing analytical results from the CATI phase-in project. First, I'll briefly describe the Current Population Survey and the sample design for the CATI phase-in study.

The CPS - Purpose and Sample Design

The Current Population Survey is a monthly survey of approximately 60,000 households, sponsored by the Bureau of Labor Statistics. It provides data on employment, unemployment, and other labor force information about the adult civilian noninstitutional population of the U.S. Sample households are selected in all 50 states and the District of Columbia. Each month's sample is composed of eight panels that rotate on a schedule of 4 months in, 8 months out, then 4 months in.

Before January 1989 all CPS interviews were conducted using paper and pencil. In the first and fifth month interview, personal visits are required. In remaining months, field representatives are encouraged to telephone respondents from their homes.

Purpose and Design of CATI Phase-in

The purpose of the CATI phase-in project was to measure the effect of using centralized, computer-assisted telephone interviewing in the CPS. We wanted to look at differences in the

estimates themselves, and also compare nonresponse rates and response variance. In January 1989, a small portion of the live CPS sample was sent to be interviewed from a CATI facility in Hagerstown, MD. Initially, a maximum of 2500 cases was sent to this facility each month. We allowed this limit to increase gradually as the operation became more routine and we had some idea of the effect on the CPS estimates. In May 1992, a second CATI facility opened in Tucson, AZ. Currently about 6 percent of the total CPS sample is sent to a CATI facility for interview.

We restricted CATI-eligibility to CPS self-representing primary sampling units (PSUs) having multiple interviewers and where we had difficulty hiring and retaining field representatives. These are mainly large metropolitan areas. For ease of administration, some of these CATI-eligible PSUs were divided into smaller areas called "subPSUs." Division was usually based on geography. Sample housing units within each subPSU were assigned to "random groups." Each random group contained about 16 housing units.

Each month, based on their needs, regional offices would determine the number of random groups to be sent to the CATI facility from each CATI-eligible subPSU. The random groups sent to CATI became the "Test" group. The remaining random groups in the subPSU were the "Control" group. In December 1992, there were about 10,000 housing units in the Test group and about 13,000 in the Control group. If a subPSU became 100-percent CATI, we dropped it from our analysis.

As in the CPS, sample housing units in both the Test and Control groups were interviewed in person using paper and pencil during their first

and fifth interviews. For the remaining months in sample, the majority of Test group units were sent to the CATI facility, while the majority of Control group units were interviewed by telephone from field representative's homes.

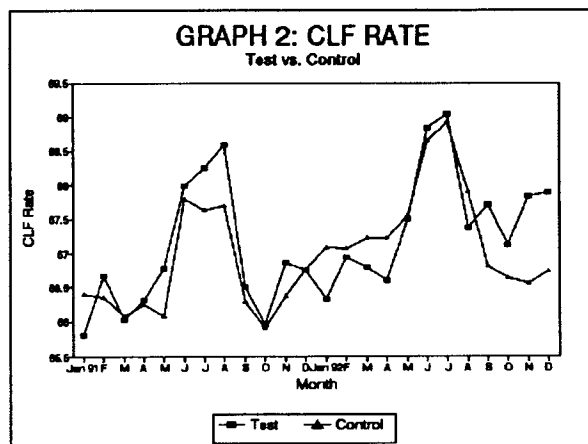
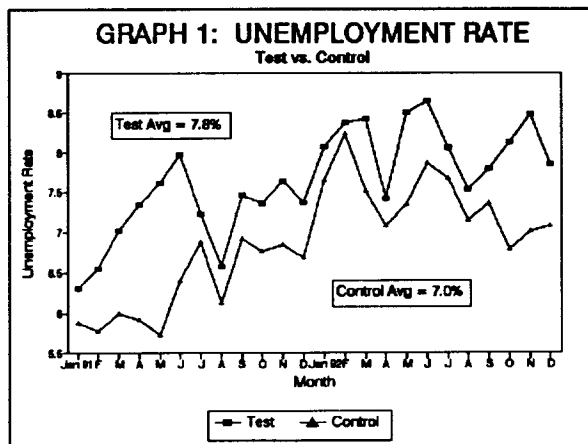
RESULTS

The following results are based on CATI phase-in data from January 1991 through December 1992. It's difficult to interpret these results because we had different subPSUs in sample at different times and, over time, the distribution of Test and Control random groups changed within these subPSUs. These areas represent only themselves and are not nationally representative. The weighting procedure included only the baseweight (the inverse of the selection probability) and a factor which adjusted for the size of the Test and Control groups within each subPSU.

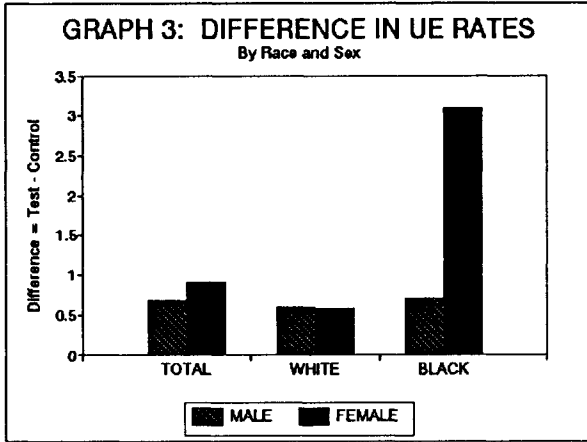
Labor Force Estimates

The labor force estimate most affected was the unemployment (UE) rate. Based on the 24 months of data we examined, the unemployment rate for the Test group was about 0.8 percentage points higher than the rate for the Control group (7.8 vs 7.0 percent, $t=3.39$). Even though the differences were not always statistically significant, the Test group had a higher UE rate for each of the 24 months (Graph 1). This result is consistent with previous CATI research [9].

The Test group's higher UE rate is mainly due to its higher estimate of the number of unemployed persons. The estimates of civilian labor force are about the same for Test and Control (Graph 2).



Although the overall difference in the UE rate was about 0.8 percentage points, it varied considerably for some subgroups. For black females, the difference was about 3.0 percentage points ($t=2.65$). For other race-sex subgroups, the differences were between 0.6 and 0.7 percentage points (Graph 3).



Since first and fifth-month interviews were done using pencil and paper for both Test and Control, we looked at months-in-sample (MIS) 1 and 5 and other MIS separately. As expected, most of the Test and Control labor force estimates did not differ significantly for MIS 1 and 5. They did differ significantly for other the MIS, where the modes of interview differed.

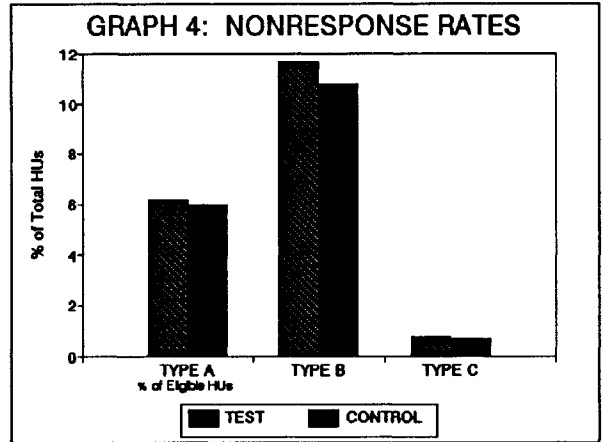
Nonresponse Rates

We compared item nonresponse rates in previous CATI research, so we focused mainly on unit nonresponse rates in the CATI phase-in study. We examined three types of unit nonresponse.

Type A nonresponses are units eligible for the survey but not able to be interviewed. Respondent refusals or the respondent's temporary absence are two examples of this type of nonresponse. Type B nonresponses are vacant units or units occupied by persons ineligible for interview. Examples are units under construction or units unfit for occupancy. Type C nonresponses are units which have been demolished, converted permanently to storage or business use, or units found in sample by mistake.

We found that the Test group had a significantly higher Type B rate than the Control group (11.7% vs. 10.8%, $t=3.14$), due mostly to a higher percentage of Type B-vacant units.

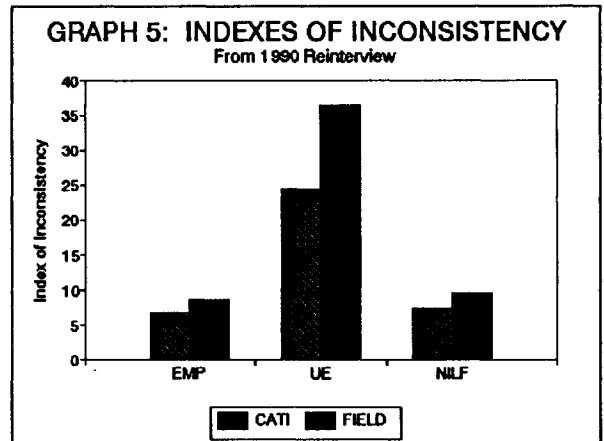
The overall Type A and C rates are similar (Graph 4).



Response Variance

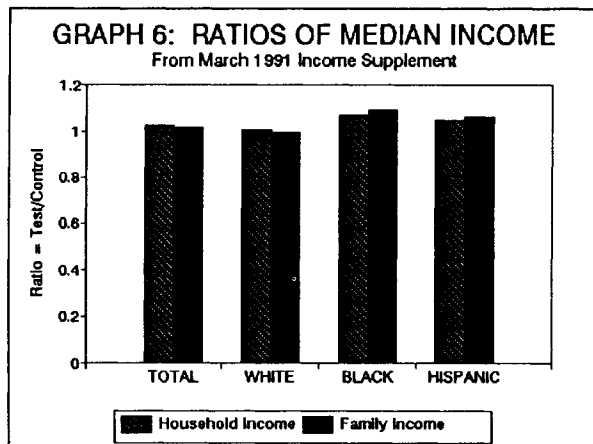
The Index of Inconsistency is a measure of response variance. It is the ratio of simple response variance to total variance for a category of a question. Historically, employed (EMP) and not-in-labor force (NILF) have low indexes of inconsistency, while unemployed has a moderate index.

Reinterview data from 1990 show that the indexes of inconsistency were significantly lower for CATI cases than for cases interviewed in the field ($\alpha \leq 0.10$). The index for UE was 36.5 for the field and 24.5 for the CATI group. The differences in the indices for EMP and NILF were smaller (Graph 5).



Income Estimates

Based on data from the March 1991 CPS income supplement, we found no significant differences in estimates of median household or family income for total persons, or by race or ethnicity (Graph 6). However there was some evidence that estimates from the Test group were higher for some subpopulations. The full version of my paper contains more details.



CURRENT RESEARCH AND CONCLUSION

In July 1992, we began testing a revised, fully automated version of the CPS questionnaire. Interviews are conducted by field representatives either in person or by telephone using laptop computers, or by telephone from a CATI facility. This test is called the CATI/CAPI overlap sample, and consists of about 15,000 households per month. One objective of this study is to separately estimate the effects of automation, centralization, and the new questionnaire. The CPS will use these new data collection methods exclusively beginning in January 1994.

The CATI phase-in study found significant differences between estimates from the Test and Control group, the most notable being the higher unemployment rate. What's causing the differences? To be perfectly honest, we can't

really say. There are two major differences between the CATI mode of interview and the paper-and-pencil mode: the automated questionnaire and centralized interviewing. The design of the phase-in study did not allow us to examine these effects separately. We hope that the analysis of data from the CATI/CAPI overlap sample will shed additional light on this.

REFERENCES

- [1] Adams, David E. (1992), "Effect of Q22A on Estimates of the Unemployment Rate from CPS CATI Phase-In Data." Internal memorandum, U.S. Bureau of the Census.
- [2] Bushery, John M., et. al. (1987), "Using CATI in the Current Population Survey," Proceedings of the Section on Survey Research, American Statistical Association.
- [3] Cahoon, Lawrence S. (1990, 1991, 1992, 1993), "Results of CPS CATI Phase-in Analysis." Series of internal memoranda (published quarterly), U.S. Bureau of the Census.
- [4] Cahoon, Lawrence S., Gregory Russell (1991), "Effects of CATI on Estimates from Demographic Surveys Conducted by the Bureau of the Census." Paper presented at the Joint Census Advisory Committee Meeting of the American Marketing Association and the Committee on Population Statistics.
- [5] Conover, W.J. (1971), Practical Nonparametric Statistics, J. Wiley and Sons, New York, pp.121-124.
- [6] Hess, Jennifer (1993), "Results of the MLS Interviewer Knowledge of Concepts Test." Memorandum for the CPS Questionnaire Evaluation Steering Committee, U.S. Bureau of the Census.

- [7] Ludington, Paul (1991), "CPS CATI Phase-In Analysis Tables." Internal memorandum, U.S. Bureau of the Census.
- [8] Mullin, Paul (1991), "Mode Differences Between CATI and Control Groups, PAPI with CATI-C Versus PAPI with PATI." Internal document, U.S. Bureau of Labor Statistics.
- [9] Shoemaker, Harland, et. al. (1989), "Evaluation of the Use of CATI in the Current Population Survey," Proceedings of the Section on Survey Research Methods, American Statistical Association.
- [10] Shoemaker, Harland (1992), "Investigation of the Effect of Question 19 on Labor Force Estimates." Internal memorandum, U.S. Bureau of the Census.
- [11] _____ (1992), "A Look at Month-in-Sample Bias Indexes for Unemployed." Internal memorandum, U.S. Bureau of the Census.
- [12] _____ (1993), "CLF Month-in-Sample Bias Indexes for CATI Phase-in." Internal memorandum, U.S. Bureau of the Census.
- [13] Tupek, Alan (1991), "Minutes of April 10, 1991 Meeting of the BLS/Census CATI Effects Workgroup." Internal BLS/Census document, U.S. Bureau of Labor Statistics.
- [14] Waite, Preston J. (1991), "1989 and 1990 Current Population Survey CPS/CATI Reinterview Response Variance Report." Internal memorandum, U.S. Bureau of the Census.
- [15] Welniak, Ed (1992), "Income Comparison Between CATI and NonCATI Control Cases from the March 1991 CPS." Internal memorandum, U.S. Bureau of the Census.
- [16] Wolter, Kirk M. (1985), Introduction to Variance Estimation, Springer-Verlag, New York.
- [17] U.S. Bureau of the Census (1978), The Current Population Survey: Design and Methodology, (Technical Paper 40), U.S. Government Printing Office, Washington, D.C., pp. 83-85.