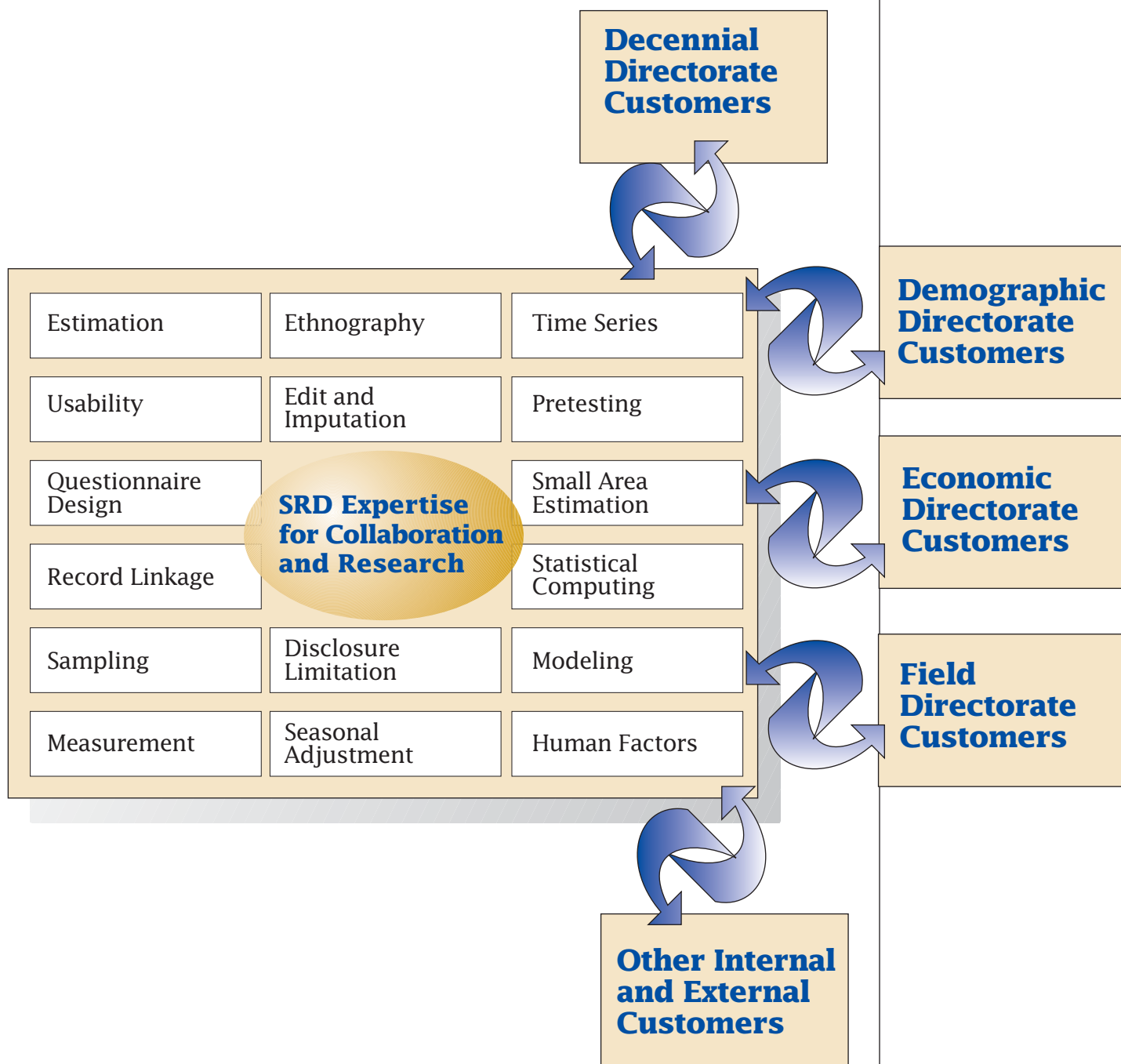


Methodology and Standards Directorate
**Annual Report of the
 Statistical Research Division**
Fiscal Year 2003





U.S. Census Bureau
Statistical Research Division
Federal Building 4
Washington, DC 20233
301-763-1702

We help the Census Bureau improve its processes and products. For fiscal year 2003, this report is an accounting of for whom we did what, why, when, and how.

Statistical Research Division

Highlights of What We Did...

As a technical resource for the Census Bureau, each researcher and technical member of our division is asked to do three things: *collaboration/consulting*, *research*, and *professional activities and development*. We serve as members on teams for a variety of projects or subprojects.

Highlights of a selected sampling of the many activities and results in which Statistical Research Division staff members made contributions during FY 2003 follow, and more details are provided within subsequent pages of this report:

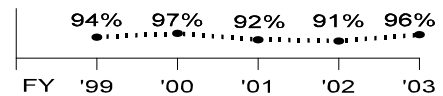
- tested, evaluated, and ensured successful operation of the disclosure limitation techniques for the Advanced Query System of American FactFinder; performed reidentification studies to determine if Census Bureau microdata files are at risk of disclosure.
- developed versions of X-12-ARIMA/SEATS that incorporated new features and code changes to 2003 releases of SEATS, including the computation of spectral diagnostics for trend, seasonal, and cycle components.
- completed development and comparisons between an exact and approximate small area estimation method that accounts for unknown within small area sampling variances.
- completed a study, based on the 1996 panel, of weighting and “model-based” alternatives for compensating for person-level longitudinal nonresponse.
- developed alternative procedures for estimating low-valued exports from the U.S. and compared them to current methods theoretically and through simulation.
- awarded seventeen task orders to 10 different contractors with a value of over 13 million dollars for research in support of efforts by program directorates through the R&D 2007 Contracts Program.
- completed new production for generating ratio edits for economic data; completed first phase of work comparing American Community Survey (ACS) “If-Then-Else,” “NIM,” and “DISCRETE” edit/imputation methodologies with ACS data.
- completed the analytical imputation procedures for the 2001 Residential Finance Survey.
- completed usability evaluations of 2 portable devices, 6 Web sites, 2 electronic forms, an Interactive Voice Recognition form, and 1 data access application; completed accessibility evaluations of several products from 3 divisions.
- conducted or coordinated thirty-eight (38) pretesting activities across the decennial, demographic, and economic areas under the Office of Management and Budget (OMB) Generic Clearance.
- developed the core curriculum and production design of the Refusal Aversion Training included in the June, 2003, National Health Interview Survey re-engineering project.
- completed research examining the sociolinguistic implications of translating cognitive interviews into Chinese.
- conducted cognitive research on a number of revised SIPP modules including: 2003 SIPP Welfare Reform Module, 2003 Reciprocity History Topical Module; and 2003 Employment History Topical Module.
- assisted Demographic Directorate staff in the review of several more SIPP modules including: Fertility History Module, Migration History Module, and the Work-Related Expenses and Child Support Paid Module.
- produced the state variance and covariance estimates using the VPLX code (revised) for the expanded Current Population Survey supplement samples for income years 2000 and 2001; these direct replicate variance estimates were smoothed and used in the Small Area Income and Poverty Estimates (SAIPE) state model for the final state SAIPE estimates of income year 2000, and the preliminary state SAIPE estimates for income year 2001.

How Did We Do...

For a fifth year, we received feedback from our sponsors. Near the end of fiscal year 2003, our efforts on fifty of our program (Decennial, Demographic, Economic, External) sponsored projects/subprojects with substantial activity and progress and sponsor feedback (Appendix A) were measured by use of a Project Performance Measurement Questionnaire (Appendix B). Responses to all fifty questionnaires provided the following results (The graph associated with each measure shows the performance measure over the last five fiscal years):

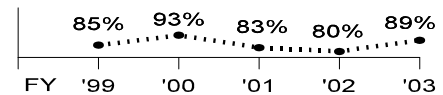
Measure 1. Overall, Work Met Expectations

Percent of FY2003 Program Sponsored Projects/Subprojects where sponsors reported that overall work met their expectations (agree or strongly agree) (48 out of 50) 96%



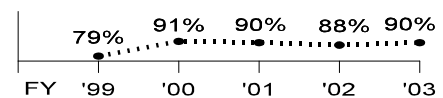
Measure 2. Established Major Deadlines Met

Percent of FY2003 Program Sponsored Projects/Subprojects where sponsors reported that all established major deadlines were met (39 out of 44 responses) 89%



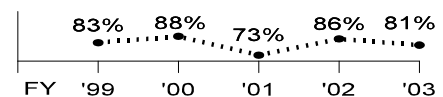
Measure 3a. At Least One Improved Method, Techniques Developed, Solution, or New Insight

Percent of FY2003 Program Sponsored Projects/Subprojects reporting at least one improved method, techniques developed, solution, or new insight (43 out of 48 responses) 90%



Measure 3b. Plans for Implementation

Of these FY2003 Program Sponsored Projects/Subprojects reporting at least one improved method, techniques developed, solution, or new insight, the percent with plans for implementation (35 out of 43 responses) 81%



Measure 4. Predict Cost Efficiencies

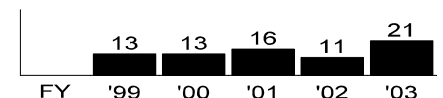
Number of FY2003 Program Sponsored Projects/Subprojects reporting at least one "predicted cost efficiency" 9



From Section 3 of this ANNUAL REPORT, we also have:

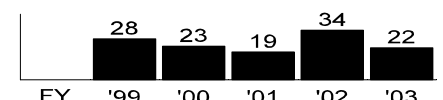
Measure 5. Journal Articles, Publications

Number of peer review journal publications documenting research that appeared (13) or were accepted (8) in FY2003 21



Measure 6. Proceedings, Publications

Number of proceedings publications documenting research that appeared in FY2003 22



Each completed questionnaire and associated details are shared with appropriate staff to help improve our future efforts.

TABLE OF CONTENTS

1. COLLABORATION	1
Decennial Directorate	1
1.1 Project 5010 - Special Purpose Tests	
1.2 Project 5201 - Content and Forms Development	
1.3 Project 5202 - Questionnaire and Public Use Forms Design and Development	
1.4 Project 5604 - Special Populations	
1.5 Project 6121 - Statistical Requirements	
1.6 Project 6367 - Census 2000 Evaluation Projects	
Demographic Directorate	6
1.7 Project 0906 - Current Population Survey	
1.8 Project 1461 - SIPP 2000 Methods Panel	
1.9 Project 1465 - Survey of Income and Program Participation (SIPP) Research	
1.10 Project 7455 - American Housing Survey	
1.11 Project TBA - Evaluation of the National Long Term Care Survey	
1.12 Project 5385 - Continuous Measurement/American Community Survey	
1.13 Project TBA - Research on Item Imputation for the 2001 RFS	
1.14 Project 7165 - NCES Poverty Statistics	
1.15 Project 7167- Cognitive Testing of the Schools and Staffing Survey Teacher Questionnaire	
1.16 Project TBA - Analysis and Forecasting of Demographic Time Series	
Economic Directorate	12
1.17 Project 3420051 - Disclosure Limitation Methods	
1.18 Project 3420052 - Time Series Research	
1.19 Project 3320054 - Implicit Ratio Edit Generation	
Methodology and Standards Directorate	14
1.20 Project 7676 - Statistical Consulting/Postal Rate Commission	
1.21 Project 7566091 - National Park Service Grant Program	
1.22 Project 8863 - National Institute of Standards and Technology -Bayesian Statistical Methodology	
Census Bureau	15
1.23 Project 0251 - Program Division Overhead	
2. RESEARCH	16
2.1 Project 0351 - General Research and Support - Census Bureau	
2.2 Project 1871 - General Research - Methodology and Standards Directorate	
3. PUBLICATIONS	34
3.1 Journal Articles, Publications	
3.2 Books/Book Chapters	
3.3 Proceedings Papers	
3.4 Statistical Research Division Research Reports	
3.5 Statistical Research Division Studies	
3.6 Other Reports	
4. TALKS AND PRESENTATIONS	38
5. STATISTICAL RESEARCH DIVISION SEMINAR SERIES	42
6. PERSONNEL ITEMS	44
6.1 Honors/Awards/Special Recognition	
6.2 Significant Service to Profession	
6.3 Personnel Notes	

APPENDIX A

APPENDIX B

1. COLLABORATION

1.1 - 1.4 DECENNIAL TOPICS

(Decennial Projects 5010, 5201, 5202, 5604)

A. Alternative Questionnaire Experiment (AQE2000)

The objectives of AQE2000 are to continue efforts to develop a user-friendly mail package that can be accurately and efficiently completed by respondents. The AQE2000 includes a test of: 1) the 1990 versus 2000 race/Hispanic origin questions on the short form; 2) the presentation of the residency rules on the short form; and 3) the long form's branching instructions. This experiment was implemented in Census 2000.

During FY2003, the final report of the Residence Rules component of the Alternative Questionnaire Experiment, "An Experiment to Improve Coverage Through Revised Roster Instructions" was issued. The synthesis report, "Synthesis Report: Census 2000 Alternative Questionnaire Experiment" was also issued.

Staff: Eleanor Gerber (x34890), Manuel de la Puente, Aref Dajani, Mary Ann Scaggs, Yves Thibaudeau

B. Response Mode & Incentive Experiment (RM&IE)

This project explores the use of response incentive and alternative modes of data collection for the Census short form.

During FY2003, corrections recommended from the Senior Executive Review last quarter were made to the final report by Westat, and the report was accepted by the Planning, Research, and Evaluation Division (PRED). PRED contracted with Research Triangle Institute to write a consolidated report about the RM&IE. This report provided an overview of Westat's report along with results about a second mailing to initial non-respondents and why mail respondents from the Internet panels used mail instead of the Internet.

Work began on new analysis about demographic characteristics of household Person 1. Staff worked with the Disclosure Review Board to approve this use of the RM&IE data, and will consult with them prior to release of the data.

Staff: Larry Malakhoff (x33688)

C. 2010 Census Residence Rule Working Group

The 2010 Residence Rule Working Group is comprised of colleagues from eleven divisions within the Census Bureau. The purposes of the group are to revise the residence rules and improve their presentation in order to improve coverage. The group will design successive tests leading up to the 2010 Census.

During FY2003, staff participated in virtually every meeting of the inter-divisional Residence Rule Working

Group and the sub-group on revising the residence rules themselves, led by staff from Population Division. We wrote the contract for the Cognitive Test of Residence Rules and Coverage, and reviewed all of the materials produced by Westat, from proposal through questionnaire materials. We reviewed successive drafts of Westat's final report and suggested that Westat add a critical section on the performance of the person count box approach to rostering, which the Decennial Management Division Task Manager approved. We reviewed successive drafts of the proposed 2004 Coverage Research Follow-up study plan and questionnaires, as well as Mobile Computing Device questionnaires to be used in the Nonresponse Follow-up Operation in 2004. Staff working group members recommended that behavior coding of the 2004 Census Site Test forms include a component on question 1 and the residence rules; this was approved and staff is now managing the 2004 census form behavior coding project.

Staff working group members suggested we invite statistical agency colleagues from other countries participating in the Census Bureau's External Visitors' Program to discuss alternative residence rules and de jure/de facto enumeration methods with our working group. This idea was accepted and we were able to talk with high-level counterparts from the United Kingdom, Australia, and New Zealand, and later with Canadians. Staff working group members participated actively in brainstorming issues concerned with the meaning of "usual residence," ("most of the time" vs. "enduring ties") of where people think they belong; de facto/de jure enumeration methods; and whether we should add a reference period to question one on the census form in 2005 to be more clear about what we mean by living or staying somewhere "most of the time."

Staff: Laurie Schwede (x32611), Eleanor Gerber

D. 2010 Coverage Planning Research and Development Group

This overall interdivisional working group provides input to the Decennial Management Division for planning successive operations, and tests broadly related to coverage research during the decade leading up to the 2010 Census. It is an umbrella group that receives proposals from four subgroups: the within-household coverage, residence rules, imputation, and unduplication working groups.

During FY2003, staff reviewed and provided extensive comments on the Operational Plan for Coverage Improvement Research. Staff participated in discussions on the planning of coverage measurement research for the 2010 Census, suggested several

research proposals, and reviewed proposals for coverage research over the decade. Staff conducted cognitive testing of alternative forms for the 2005 Content Test, and were invited to review and comment on the Coverage Research Follow-up questionnaire to be used in conjunction with the 2004 Census Test.

Staff: Laurie Schwede (x32611), Manuel de la Puente, Eleanor Gerber

E. Behavior Coding Evaluation of the Census Quality Survey

Behavior coding of the interviewer-respondent interactions of two different versions of the Census Quality Survey (CQS) questionnaires are being evaluated to aid in the assessment of the performance of two different versions of the race and ethnicity questions. One hundred and twenty-five taped interviews of each of the two questionnaire versions are being coded.

During FY2003, this project was completed.

Staff: Eleanor Gerber (x34890), Betsy Martin (M&S)

F. Decennial Privacy Research

The purpose of this project is to serve on and assist the work of the Privacy Research Coordinating Committee (now the Privacy Policy and Research Committee), and to conduct research to assess public opinion on privacy-related issues, including the increased use of administrative records to assist census enumeration.

During FY2003, staff participated in the creation of research proposals to conduct research to assess public opinion on privacy-related issues, and to assist with the design of privacy-related policies. Staff drafted a policy/issue paper regarding informed consent. Staff chaired a subgroup charged with closing the gaps between the privacy principle of Informed Consent, and current Census Bureau policies and procedures. Staff participated in a team that designed a Title 13 training program for Census Bureau staff. Staff continued to participate in the work of the Privacy Policy Committee (PPRC).

Staff: Tom Mayer (x34930), Jeff Moore, Ashley Landreth

G. Decennial Questionnaire Design and Pretesting

This project involves participation in the 2010 Census Content Planning Group and content-related subgroups. It also involves consultation and testing on questionnaire content for the 2010 Census and tests leading up to it.

During FY2003, staff attended meetings of the

2010 Census Content Planning Group meetings and conducted cognitive research in two areas related to the content of the census form. We conducted in-depth interviews with experts in the field of assisted living facilities to determine whether and how the tenure question in the Decennial Census might be altered to allow reporting of such facilities, since they do not fall neatly into the owner or renter categories (Hunter 2002). We found that assisted living facilities cover a wide scope of living arrangements, and that the diversity of the arrangements cannot be captured easily in a single (new) response category to the existing tenure question.

We conducted unstructured interviews with staff from group homes, both government-sponsored and private. Interviews were also conducted with respondents from across the country who, in Census 2000, reported a household member over aged 18 as a foster child. The purpose of these interviews was to investigate the implications of adding a new response category to the 2004 Census Test relationship question, to provide a place to report residence in group homes (part of a plan to move group homes from the group quarters universe to the mailout universe of the census), and perhaps to improve the reporting of foster children over the age of 18. We found that the majority of group homes are staffed in shifts, rather than having live-in staff. Person 1 on the census form would likely be a resident, and the relationship of the resident to other residents would be housemate/roommate, in most cases. This suggests that a new category may not identify these homes. There are also many disadvantages to moving small group homes into the mailout/mailback universe from the group quarters universe. They function as group quarters, and some of the questions on the housing unit form do not apply to their situation. Callbacks to Census 2000 respondents indicated that many people reported as foster children were in fact, living in adult foster care situations. Based on these results, the recommendation was made that: 1) small group homes be kept in the group quarters universe with no additional category; and 2) the wording of the "foster child" category be changed to include adults. A secondary recommendation, in the case that the first was not accepted, was to change the wording of the proposed "person in custodial care" category to "nonrelative receiving formal or informal care." The primary recommendations were not accepted for the 2004 Census Test, but remain in consideration for the 2010 Census. The recommended wording was modified for the 2004 Census Test.

We are in the planning stages of a behavior coding operation to be conducted on Non-Response Follow-Up (NRFU) data collected in the 2004 Census test in Queens, NY. A small "sample" of cases conducted in English and Spanish will be behavior coded as part of

the evaluation of the race and Hispanic origin questions, and also as part of a general evaluation of the Mobile Computing Device (MCD) instrument.

Staff: Terry DeMaio (x34894), Jennifer Hunter, Ashley Landreth, Lorraine Randall, Laurie Schwede, Luc Perkins

H. Response Options Strategies Working Group

As a sub-group under the 2010 Census Self-Response Options Research and Development Planning Group, this working group is chartered to define and develop strategies for offering self-response options in support of the 2003 Census Test and later testing efforts. Self-response options include mail, Internet, Interactive Voice Response (IVR). Focusing on the IVR and Internet response options, this group proposes testing strategies to identify the best approach to encouraging citizen response in the 2010 Census.

During FY2003, led by staff from the Decennial Systems and Contracts Management Office, the Response Options Team worked with the software contractor (Z-Tech Corporation) to develop the IVR option and the Internet option for the 2003 National Census Test. Usability staff planned and coordinated usability testing with testing contractors. By using automated tools, and interviewing visually impaired recruits, staff also conducted accessibility testing of the prototype Census 2003 Internet form. Results of usability and accessibility testing were reported to the working group, and change requests were submitted to the software contractor for both the IVR and Internet options.

A second round of usability testing was conducted on the Census 2003 Internet form in collaboration with Dr. Kent Norman of the University of Maryland. Staff prepared a short initial report on the usability issues, including recommendations for resolving them, as well as a final report, including details of the testing methods. Staff participated in planning for the 2004 Overseas Enumeration Test, which will evaluate the feasibility of counting Americans living abroad by means of an optional Internet form in the 2010 Census. Usability and accessibility staff developed plans for testing the most recent versions of the IVR application and the 2004 Internet form. Testing will occur in early FY2004.

Staff: Betty Murphy (x34858), Larry Malakhoff, Beth Nichols, Erica Olmsted, Kent Marquis

I. Census 2000 Gateway Testing

The Census 2000 Gateway is a portal page that has been added to the Census Bureau's Web site to help users find Census 2000 data. The purpose of this project was to evaluate the usability of this new portal.

UserWorks, Inc. conducted all testing using the usability lab's remote-testing capabilities.

During FY2003, following the completion of remote usability testing, staff reviewed and began revising the draft report submitted by UserWorks. Staff also began preparing a briefing to be presented to the Decennial Systems and Contracts Management Office. Staff maintained communications with Gateway team members.

Staff: Betty Murphy (x34858), Erica Olmsted, Kent Marquis, Joyce Farmer

J. Mobile Computing Devices - 2004 Non-Response Follow-Up (NRFU) Usability Evaluation

In the 2004 Census Test, households not responding to the mailed out census questionnaire will be followed-up by enumerators using a hand-held device for displaying questions and recording answers. The mobile computing device instrument is being programmed by the Technologies Management Office (TMO). The Decennial Management Division (DMD) asked the usability laboratory to assist TMO by conducting three iterative usability evaluations of the instrument itself (not other features such as GPS, receipt and transmission of assignments, etc.) and reporting any problems found.

During FY2003, staff met with representatives of the Decennial Census staff and worked out the basic features of the collaboration. The usability lab submitted a project proposal by the third week in April. DMD provided additional funding for the program to allow assistance from contractors. Staff outlined a general testing strategy and drafted a request for quotation from a contractor to assist in the project. Staff developed a testing protocol, user training, and the analysis strategy. Staff reconfigured the laboratory to allow recording of the activity on a hand-held device, to allow the playback of the activity recording so the user can supply a retrospective think-aloud commentary, and enabled the recording of both the playback and the commentary for logging and analysis. Staff collaborated with a contractor on data collection and data analysis. Staff ran 9 individuals through the usability test; all 9 individuals came back two weeks later to participate in the second phase of the testing. Staff began preliminary analysis and began composing the list of usability violations.

Staff: Erica Olmsted (x34971), Kent Marquis

1.5 STATISTICAL REQUIREMENTS (Decennial Project 6121)

A. Decennial Coverage Research

The objective is to provide short-term research and

statistical support to the Decennial Statistical Studies Division (DSSD) in preparation for the implementation and assessment of the Accuracy and Coverage Evaluation Survey (A.C.E.).

During FY2003, staff completed evaluations of the synthetic estimates of adjusted population counts and shares for States, Counties, and Places derived from the Revised A.C.E. The evaluation, based on constructing plausible artificial populations for comparison, incorporated the new stratification used in the revised A.C.E. This project was a joint effort with the Decennial Statistical Studies Division. Under a tight time frame, results were delivered on schedule. These results were included in documentation that will accompany the design and development of the Revised A.C.E. counts.

Staff: Don Malec (x31718), Jerry Maples

B. Research on Accuracy of the Census and the A.C.E. Estimates

This project examines the quality of Census 2000 and A.C.E. estimates. These analyses investigate the biases and random errors in the dual system estimator arising from data collection, data processing, missing data, and model bias. A synthesis of these errors takes the form of a total error model. Using the results of the total error model simulations, loss function analyses compare the accuracy of the census and the A.C.E. estimates. These analyses support the decisions on which census data set to release for three different uses on three different occasions. The analyses add refinements as new evaluation data become available, and are tailored for the specific application.

During FY2003, staff directed the implementation and analysis of the evaluation program for the Revised A.C.E. Revision II estimates of coverage error in Census 2000. Reports and other documentation for 12 individual projects in the evaluation program were completed. Staff co-authored two of the reports and conducted in-depth reviews of the others for approval. Staff assessed the relative quality of the census and the A.C.E. Revision II estimates through the formation of bias-corrected confidence intervals and a loss function analysis. Although most of the known error in the A.C.E. was incorporated in the A.C.E. Revision II estimates, the bias-corrected confidence intervals synthesized the remaining nonsampling errors that could be measured in the evaluation program along with the sampling error. The program included the use of administrative records and a clerical review to evaluate the accuracy of the automated identification of duplicate census enumerations.

Staff contributed to the Technical Assessment of the A.C.E. Revision II estimates of coverage error in Census 2000. The Technical Assessment summarized

the limitations of the estimates and synthesized the results of 12 projects that evaluated different aspects of the estimates. The document was used in the deliberations that decided not to use the A.C.E. Revision II estimates in the Census Bureau's Postcensal Estimates Program. This project has been completed.

Staff: Mary Mulry (x31759)

C. 2010 Census Coverage Measurement Research

2010 Census Coverage Measurement Research conducts the research necessary to develop methodology for evaluating the coverage of the 2010 Census. This includes planning, designing, and conducting the research as well as analyzing and synthesizing the results.

During FY2003, staff began preliminary planning of coverage measurement research for the 2010 Census. The initial step is drafting a document that describes the important problem areas and opportunities for improvement related to census coverage measurement. The preparation of this document requires reviewing the results of the evaluations of A.C.E. Revision II and other evaluations of Census 2000.

Staff: Mary Mulry (x31759)

D. A.C.E. Missing Data Research and Development

The objective of this work is to conduct research to guide the development and assessment of appropriate imputation for missing data in the Accuracy and Coverage Evaluation (A.C.E.).

During FY2003, staff produced a final specification for the noninterview adjustment in the A.C.E. Revision II. Staff ran the noninterview adjustment program on the final A.C.E. Revision II data files and assisted in verification of results. Staff produced a memorandum giving an overview of results from the A.C.E. Revision II noninterview adjustment. Staff provided general support to the Decennial Statistical Studies Division on assorted issues related to handling of missing data system for the A.C.E. Revision II and assisted in the verification of the missing data system for the A.C.E. Revision II. Staff also provided consulting support to the Planning, Research, and Evaluation Division on assorted issues related to the evaluation of missing data procedures in the A.C.E. Revision II.

Staff: Mike Ikeda (x31756), Pat Cantwell

E. Administrative Records Linkage Support

The purpose of this project is to apply the record linkage software to a variety of projects, primarily in the Decennial, but also the Demographic and Economic

Directorates, including administrative lists.

During FY2003, staff made improvements to the computer matching system in support of the Special Places/Group Quarters unduplication project. Staff supported the Special Places/Group Quarters (SP/GQ) unduplication with a matching run that found duplicates in the file and the corresponding match. Staff also developed a new matching technique and algorithms which increased the number of duplicates found in the SP/GQ file.

Staff: Ned Porter (x31798)

F. Transparent File Research

This research involves development of methods to construct a Decennial Census data file in which the effects of sampling and estimation are transparent to the data user. Data files of interest include both short and long form data.

During FY2003, staff programmed a possible refinement to the methodology that uses tract race by tenure controls (instead of just tract race controls) in the elimination step of the donor selection phase. A preliminary examination of the output suggests that the additional controls produce mixed results. "Transparent File Construction for the State of New Jersey in Census 2000" has been released as an SRD Research Report. The report outlines the methodology and results for the New Jersey transparent file. This project has concluded.

Staff: Julie Tsay (x34925), Michael Ikeda

G. Decennial Edit/Imputation Research (See Demographic Project 5385-D.)

H. Decennial Disclosure Limitation Research

The purpose of this research is to develop disclosure limitation methods to be used for Census Bureau publicly available decennial data products. Emphasis will be placed on techniques to implement disclosure limitation at the stage of processing. Disclosure research will be conducted on alternative methods to protect both tabular data and microdata from the decennial census. Methods will be developed, tested, evaluated, and documented. We will also aid in the implementation of the methods.

During FY2003, staff developed production software for finding small subtraction geographies (slivers). The Census Bureau publishes demographic profiles of user defined neighborhoods or a limited number of jurisdictions in the census special tabulations program. Profiles have already been released at the tract level as part of the regular census publication. The software finds the smallest units obtainable by the subtraction of neighborhoods and tracts and checks the

population size. Neighborhood boundaries are changed if slivers are found.

Staff tested a revised version of the Advanced Query System. The confidentiality filters are working correctly.

Staff rewrote the swapping routine on Special Censuses. It is difficult to find partnering households in such small geographic areas, so there needed to be exact control on geography and selection of households for swapping.

Staff worked with Associate Directors and other staff members to develop disclosure limitation techniques to be used to protect Group Quarters in Census 2000 special tabulations and the Advanced Query System.

Staff: Laura Zayatz (x34955), Phil Steel, Paul Massell, Sam Hawala

1.6 CENSUS 2000 EVALUATION PROJECTS (Decennial Project 6367)

A. Census 2000 Ethnographic Studies

During FY2003, and the early part of FY2004, staff concluded research and reported findings from Census 2000 ethnographic studies relating to protecting privacy, complex households, civic engagement of Generation X, mobile populations, social network tracing of highly mobile people, and enumeration methods for colonias or border communities. The specific reports are: (1) Eleanor Gerber, "Privacy Schemas and Data Collection: An Ethnographic Account." *Census 2000 Ethnographic Studies*, February 10, 2003; (2) Laurie Schwede, "Complex Households and Relationships in the Decennial Census and in Ethnographic Studies of Six Race/Ethnic Groups." *Census 2000 Ethnographic Studies*, August 27, 2003; (3) Melinda Crowley, "Generation X Speaks Out on Civic Engagement in the Decennial Census: An Ethnographic Approach," *Census 2000 Ethnographic Studies*, June 17, 2003; (4) Manuel de la Puente, Jenny Hunter, and Matt Salo, "Comparative Ethnographic Research on Mobile Populations," *Census 2000 Evaluation J.3.:Ethnographic Studies*, July 17, 2003; (5) Manuel de la Puente and David Stemper, "The Enumeration of Colonias in Census 2000: Perspectives of Ethnographers and Census Enumerators," *Census 2000 Evaluation J.4: Ethnographic Studies (Draft)*, September 15, 2003; (6) Leslie A. Brownrigg, "Ethnographic Social Network Training of Highly Mobile People," *Census 2000 Evaluation J.2: Ethnographic Studies (Final)*, October 16, 2003.; (7) Manuel de la Puente, "Census 2000 Ethnographic Studies," *Census 2000 Topic Report (Draft)*, September 24, 2003.

Staff: See Above Authors

B. Decennial Unduplication and Coverage Evaluation

A new survey instrument to measure coverage and to resolve possible duplications will be developed in conjunction with the Decennial Statistical Studies and Decennial Management Divisions. The instrument is being developed for use in the 2004 Site Test, with the aim of measuring coverage and unduplicating in the 2010 Census. The survey will place individual and whole household duplicates in their proper residence according to decennial residence rules. It will identify both erroneous enumerations and omissions in the Census.

During FY2003, staff helped to create and revise a draft of a survey instrument that can be used to collect information to correctly enumerate duplicate census enumerations, as well as to evaluate general household coverage. Cognitive testing of this instrument was arranged and monitored.

Staff: Eleanor Gerber (x34890)

1.7 CURRENT POPULATION SURVEY (Demographic Project 0906)

Tobacco Use Supplement to the 2003 Current Population Survey: Smoking Cessation

Develop the questionnaire content of the Tobacco Use Supplement to the 2003 Current Population Survey on Smoking Cessation. Conduct cognitive testing on the instrument that will be administered in both CATI and CAPI modes. Data from this survey supplement provides information for national and state estimates on emerging adult tobacco control issues as part of the National Cancer Institute's Extraordinary Opportunities in Tobacco Research. The specific objective of the research is to learn how well newly-devised questions work to capture data on smoking cessation activity, including stages of readiness to quit, measures of addiction, quitting behavior, cessation methods used, and perceptions about different types of cigarettes. This survey is part of a continuing series of surveys that were originally fielded over the 1990s by the Census Bureau, and will be continuing over the next decade, alternating between a standard or core tobacco use survey and a special topic survey focusing on tobacco control issues.

During FY2003, this project was successfully completed.

Staff: Melinda Crowley (x32726), Kristen Hughes

1.8 SIPP 2000 METHODS PANEL (Demographic Project 1461)

The SIPP Methods Panel is the R&D vehicle for development of a redesigned SIPP instrument for the 2004 SIPP Panel. Through a combination of expert

review, user needs assessment, secondary data analysis, and laboratory research, Methods Panel (MP) staff carry out the research activities necessary to implement the recommendations of the Continuous Instrument Improvement Group (CIIG).

During FY2003, Wave 2 of MP2002, the final Methods Panel field test was completed. We completed a set of several "hand-off" meetings - one for each major content area of the SIPP survey - with one production SIPP staff in all demographic area divisions, to educate them about the new Wave 1 and Wave 2+ core instruments, and to assist in converting the Methods Panel instrument from experimental mode to the production instrument for 2004. Staff also conducted similar (though much more abbreviated) briefings for the Office of Management and Budget's SIPP Advisory Group. We documented final Wave 1 and Wave 2+ instrument procedures, and devoted many hours to testing and refining the instruments, including several very valuable sessions involving SIPP Field Representatives. Staff reviewed and commented on draft training materials prepared by Field Division.

We prepared and presented an invited paper describing the field test process (and some results) at the International Conference on Questionnaire Design, Evaluation, and Testing Methods. We also prepared papers for the American Statistical Association's annual meetings, focusing primarily on a section-by-section overview of the new instrument procedures.

Staff: Jeff Moore (x34975), Anna Chan, Katie Gagne, Julia Klein-Griffiths, Joanne Pascale

1.9 SURVEY OF INCOME AND PROGRAM PARTICIPATION (SIPP) RESEARCH (Demographic Project 1465)

A. Measurement Research on SIPP

The purpose of this project is to design, conduct, analyze, and report on research which addresses measurement error and nonresponse issues in SIPP, and which assists the development of new content areas.

During FY2003, staff conducted cognitive research on a number of revised SIPP modules. We conducted cognitive research on the 2003 SIPP Welfare Reform Module (Hunger and Hughes 2003). We found that: 1) respondents were not aware of some benefits such as job subsidies; 2) respondents were not familiar with the concept of workfare or unpaid job; 3) in the core SIPP instrument, respondents were not familiar with the terms "general assistance" or "general relief," what was meant by short-term cash assistance, and whether they were in fact, receiving Medicaid (since the medical assistance program is known by different names in the states where the interviews were conducted). Questionnaire revisions were recommended to remedy these situations, and many of them were accepted.

We conducted cognitive research on the 2003 Reciprocity History Topical Module (Hughes and Hunter 2003). We found that the instrument seemed to successfully capture aspects of the respondents' program history. One notable problem was the reference to the receipt of various program benefits (TANF/AFDC, Food Stamps and Supplemental Security Income (SSI) during one's adult life. This is inappropriate for respondents who receive benefits in their own name because they have become mothers before they were age 16 (which is SIPP's definition of an adult). Another, more minor problem was language in some questions that was too complex for respondents with low levels of education. Questionnaire revisions were recommended to remedy these situations, and most recommendations were accepted.

We conducted cognitive research on the 2003 Employment History Topical Module (Landreth 2003). Testing showed that the instrument seemed to successfully capture aspects of respondents' entry into the labor force and recent employment history. Recalling the year and month for previous work history was not particularly burdensome for respondents when these events were within the past year or two, and they successfully used a variety of cues to formulate their responses. This burden is increased for respondents whose events are temporarily distant and/or not associated with a salient event, and whose work histories are lengthy and complex. One of the most difficult response tasks was estimating the total time spent out of the labor force for spells of six months or longer, especially for those in the increased response burden categories. Questionnaire revisions were recommended to remedy these situations, and most recommendations were accepted.

Staff: Terry DeMaio (x34894), Anna Chan, Kristen Hughes, Jenny Hunter, Ashley Landreth, Jeffrey Moore, Yuling Pan, Lorraine Randall

B. Continuous Instrument Improvement Group (CIIG)

CIIG serves as a vehicle for systematically reviewing the redesigned SIPP instrument to identify data quality problems, for recommending research to address problems arising from instrument design, and for recommending instrument revisions.

During FY2003, CIIG, assisted by subject-matter experts in the Demographic Surveys Division, Housing and Household Economic Statistics Division, and the Population Division (POP), completed its review of several more SIPP modules: the Wave 2 Fertility History module the Wave 2 Migration History module, and the Wave 3/6/9 Work-Related Expenses and Child Support Paid module. All recommendations were approved by the SIPP Executive Committee (SIPP

Exec). (CIIG also completed its review of the Wave 3/6/9 Medical Expenses and Utilization module, but has yet to formally submit its recommendations to SIPP Exec - we await only an anticipated request from Health and Human Services/Assistant Secretary for Planning and Evaluation concerning the addition of a question on non-custodial parents' contributions to children's health care costs). The recommended improvements for these modules consist primarily of (a) trying to make more extensive use of what is already known from the core interview, to enable asking more appropriately-worded questions, and to avoid some questions altogether; (b) re-wording question to reduce sensitivity, to simplify, or to clarify intent (or some combination thereof); (c) deleting questions that are not analytically necessary; or which duplicate information captured in the core interview; (d) where possible, capturing information about multiple people from one respondent (e.g., families who share the same migration history); and (e) adding screening questions for additional efficiency. CIIG also submitted a set of recommended "immediate" (i.e., 2004 Panel) changes to the Wave 2 Marital History, Work Disability History, and Education and Training History modules. CIIG's economic modules" sub-team, also including subject area experts from Demographic Surveys Division and Housing and Household Economics Statistics Division, began its review of the initial components of the Wave 3/6.9 Assets, Liabilities, and Eligibility modules - specifically the Real Estate, Shelter Costs, Dependent Care, and Vehicles sub-module, the Value of Business sub-module and the Interest-Earning Accounts sub-module.

Staff conducted cognitive research on the redesigned Wave 1 modules concerning Reciprocity History and Employment History. We delivered our recommendations to the Demographic Surveys Division staff at the end of February, and completed final reports in March. Although in each case, the cognitive interviews yielded some suggestions for further refinements to the modules, in the main they served to confirm that CIIG's redesign recommendations are sound, and that the modules should work well in the field.

Staff: Jeff Moore (x34975), Anna Chan, Julia Klein Griffiths, Kristen Hughes, Jenny Hunter, Ashley Landreth, Joanne Pascale

C. Longitudinal Weighting

The objective of this project is to design and conduct research required to assess the effectiveness of weighting alternatives for the SIPP longitudinal estimation.

During FY2003, staff developed and evaluated response propensity and program participation models based on data from the 1996 SIPP panel. In addition, we proceeded with empirical investigations designed to

glean greater insight into the effects of attrition and weighting alternatives on SIPP longitudinal estimation. Results of our studies involving the 1996 panel data were promising, yet not definitive when we compared weighting alternatives considered to the existent data adjustment procedures.

We are currently comparing the research results based on the 1996 panel to our results derived for the 1992-93 panels.

Staff: Leroy Bailey (x34917), Todd Williams, Julie Tsay

D. Adapting Standard Analytical Procedures to the Complex Sampling Structure of SIPP

There has long been a need for establishing practical methods to use when analyzing data from complex surveys in order to reduce the time and effort required to obtain valid inferences by “correct statistical” procedures. These procedures are even more complicated for SIPP as a result of the longitudinal nature of its data. The intent of the project is to develop guidelines for adapting standard analysis methods for use with the complex sampling structure of SIPP. These adaptations may require either completely new analyses or adjustments to standard analyses, such as the use of design effects.

During FY2003, work on this project was concluded.

Staff: Ruben Mera (x34938), Leroy Bailey

E. Quick Turnaround Pretesting for Household Surveys (See Projects 0351 and 1871, SURVEY METHODOLOGY)

1.10 AMERICAN HOUSING SURVEY (Demographic Project 7455)

This project provided questionnaire pretesting assistance for the American Housing Survey (AHS).

During FY2003, staff met with staff from the Demographic Surveys Division (DSD) to discuss proposed changes to the AHS. Housing and Urban Development (HUD) received proposals from data users and analysts. After identifying emerging issues, DSD prepared final content scope. Pretesting will be conducted in fall 2003. Enhancements will be integrated in the first half of 2004.

Staff: Eileen O’Brien (x32695)

1.11 EVALUATION OF THE NATIONAL LONG TERM CARE SURVEY (Demographic Project TBA)

The National Long Term Care Survey is conducted every five years, and is scheduled to be administered in 2004. Revisions to the Survey since its last implementation will be field tested in a pretest. Respondent debriefing questions will be field tested in a pretest in December 2003 - January 2004. The survey

is composed of multiple modules depending on whether the sample person lives at home or at an institution. Respondent debriefing questions will be used to evaluate the effectiveness of new and revised questions in each of the modules.

During FY2003, staff met with the Demographic Surveys Division (DSD) to discuss the survey’s objectives and evaluation needs for the pretest. Draft respondent debriefing questions were reviewed, and comments were provided to the sponsor. Additional information was sought from the sponsor regarding intent of specific questions. Staff forwarded a final set of respondent debriefing questions to DSD. Staff requested to review the interviewer training materials and to have access to the survey instrument in order to test the flow of the respondent debriefing questions.

Staff: Jennifer Rothgeb (x34986)

1.12 CONTINUOUS MEASUREMENT AMERICAN COMMUNITY SURVEY (ACS) (Demographic Project 5385)

A. ACS Language Research: Focus Groups with FRs and SFRs and Cognitive Testing of the CAPI/CATI Spanish Language Instrument

Very little research has been conducted on how persons whose primary language is Spanish (and who speak little or no English) respond to Census Bureau surveys. Moreover, there is no research on the in-person interview process of this population. The aim of this effort is to better understand how persons whose primary language is Spanish (and who speak little or no English) understand and interpret the ACS Spanish and English language CATI/CAPI instrument. A better understanding of this process can provide guidance in improving survey question wording and translation. These alterations can, in turn, lead to improvements in data quality. This research entails the conduct of cognitive interviews with persons whose primary language is Spanish using the ACS Spanish and English language CATI/CAPI instrument. Interviews will be conducted in the metropolitan areas of Washington, D.C., Chicago, Los Angeles, among others where there is a high concentration of Hispanics. This project also involves focus groups with field representatives (FRs) and supervisors of field representatives (SFRs) who interview persons whose primary language is Spanish.

During FY2003, staff prepared and delivered several briefings to the Executive Staff, other Census Bureau staff, and the Census Bureau Advisory Committee. Revisions to the final report are in progress. The final report will be issued during the first quarter of FY2004.

Staff: Manuel de la Puente (x34997), Lorena Carrasco (DMD)

B. ACS Small Area Estimation Research Methods for small area estimates at the tract level

will be developed, evaluated, and contrasted with the aim of recommending a production method. We will evaluate and contrast the assumptions and accuracies inherent in Demographic small area population estimates and ACS estimates.

During FY2003, based on a probability model and proposed method that combines ACS and Internal Revenue Service (IRS) data for estimating county migration, we examined the feasibility of using the proposed method with the current ACS data. Preliminary conclusions are that the current data is not rich enough for a full demographic breakdown (age, race, sex, and Hispanic origin) without imposing additional model restrictions. However, this may not be the case when ACS goes to full implementation. Lastly, we have been investigating how to incorporate the change in the classification of race (multi-racial responses) into the demographic breakdown without overextending the available data. Little data analysis work can be done until we can access the IRS data.

Staff presented an overview of project status and research issues, "Using the ACS to Improve Net Domestic County Migration" for the National Academies' Panel on Research on Future Census Methods.

Staff: Donald Malec (x31718)

C. ACS-Edit/Imputation Research

The purpose of this project is to determine the feasibility of applying and adapting an edit/imputation system created for Census 2000 using the DISCRETE prototype edit system and to-be-developed statistically valid item and unit imputation methods to the American Community Survey (ACS). The edit part of the project is (1) to create valid code and sufficiently fast algorithms for editing and (2) to translate traditional decennial edit rules into the Fellegi-Holt framework in a technically feasible manner. The imputation part is to impute for missing and contradictory data using statistically valid methods.

During FY2003, staff continued work on the comparison study of ACS If-Then-Else, NIM, and DISCRETE edit and imputation systems. For DISCRETE, staff completed imputation results for 3-, 4-, 5-, 6- person households. The imputation results were compared to ACS If-Then-Else and NIM. Updated checking programs determined how many of the imputes satisfied edits. An overview paper was updated with the new imputation results. Staff also completed the first draft of the paper "Preorder and Set Covering in the DISCRETE Edit System" that describes the improvement of the preorder and the set covering algorithm performance using the concept of lexicographic ordering. A major improvement was in the removal of most of the storage requirement of the preorder data structure.

Staff worked on improving imputation for discrete

data. The following gives more details. An imputation system based on discriminant analysis to be embedded in the DISCRETE edit/imputation system was completed for imputing items for households of three to six persons in the American Community Survey. Preliminary measurements of the performance of the imputation system suggest it performs well in terms of reconstructing relationships between household members, relative to the currently used If-Then-Else imputation system, and is competitive with the Canadian system NIM. Professors Little and Schafer suggested more statistically valid ways of comparing this new imputation system with that embedded in NIM and If-Then-Else by deleting household items from a complete file of households, using selected deletion mechanisms including, but not limited to, data missing at random, and reconstructing the files with each system for subsequent evaluation. We are in the process of conducting such performance evaluations, as well as extending the imputation system for households of more than six persons. A complete description of each software component in this new imputation system to be embedded in the DISCRETE edit/imputation system is now available.

This project is completed.

Staff: Bill Winkler (x34729), Bor-Chung Chen, Yves Thibaudeau, Todd Williams

E. Nonresponse Weighting Adjustments for ACS Application

This project develops new methodology for model-based nonresponse weighting adjustment of national survey data such as ACS.

During FY2003, staff proposed to apply methodology for model-based nonresponse weighting adjustment of national survey data such as ACS.

Staff has been assembling data from 2000 decennial files in preparation for demographic/geographic modeling (at tract level) of mail and follow-up response rates, for direct comparison with previously developed models of the same type based upon 1990 decennial data. The ultimate objective is to compare the Decennial Census models, and later the 2000 decennial models, with models for response by different "modes" within the ACS, in order to assess the future usefulness of demographic/geographic weighting adjustments within ACS, based on earlier survey (ACS or Decennial Census) results.

Staff has fitted and compared logistic regression models for Decennial 1990 and 2000 Census mail-response by housing units at tract level. Results include the surprisingly good fit of 1990 models to 2000 data. One ultimate objective is to compare the Decennial Census models, and later the 2000 Decennial models with models for response by different "modes" within the ACS, in order to assess the future usefulness of

demographic/geographic weighting adjustments within ACS based on earlier survey (ACS or Decennial Census) results. Another application of this work is to the Population Division's Targeting Database of Greg Robinson.

The year's progress on this project indicates surprisingly high cross-census predictability of household mail response rates at tract level, and good correspondence between predictions from the 1990 logistic regression model and the 2000 observations. The immediate next step will be further comparison of 1990 and 2000 models for rates of response to follow-up enumerators, and quantification of errors if model-based non response weighting adjustment on an ACS-like survey using the Census 2000 frame were done using the models fitted to Census 1990. The 2000 models will later be used to explore predictions on actual ACS data.

Staff: Eric Slud (x34991)

F. ACS Weighting Simplification Research

The current ACS weighting methodology was designed in 1995 and is composed of a series of 13 steps or adjustments. Staff will investigate the possibility of eliminating or combining steps to simplify the overall process. This will be done by individually eliminating each step in the process to determine the effect its elimination has on the weighting and estimation results.

During FY2003, staff reviewed specifications for weighting of, and calculating estimates from, ACS and supplementary surveys data. An exploratory analysis of the adjustment factors suggested weighting steps to target for initial removal. Evaluation measures and characteristics of persons and households to which they would be applied were developed, and specifications were prepared for their calculation. Software for carrying out the comparisons was written and tested. Staff performed these and additional comparisons for data from the ACS 2000 sites, looking for a factor or factors that can be removed from the weighting with only a minor effect on the estimates. A draft report summarizing the results of this analysis was prepared.

It concludes that the effects of dropping either of the noninterview factors by themselves or in combination with dropping the mode bias factor are very similar and generally small.

Staff: Lynn Weidman (x34902), Michael Ikeda, Julie Tsay

G. Program of Integrated Estimates

Research is planned in several divisions to carry out research to improve the Census Bureau's annual intercensal population estimates under the Program of

Integrated Estimates (PIE). Its goal is to combine information from the Census 2000, data on population characteristics from the ACS, and administrative records to produce improved small area population and housing unit estimates. A division staff member is chairing the working group for this project.

During FY2003, staff organized an inter-divisional PIE Working Group to review current work and issues and arranged its meetings. Staff collaborated in the delineation and summarized the status of seven component PIE projects and two related ACS weighting projects. These summaries were presented for discussion and prioritization at meetings of the ACS Research and Evaluation Steering Committee, inclusion in the ACS Operations Pan, and presented at an advisory committee.

Staff: Lynn Weidman (x34902)

1.13 RESEARCH ON ITEM IMPUTATION FOR THE 2001 RFS (Demographic Project TBA)

The goal of this project is to derive statistical approximations for unreported dollar amounts for items on the Residential Finance Survey (RFS) such as value of house at the inception of the mortgage, mortgage balance, and initial amount of the mortgage when these items are not reported, based on available items.

During FY2003, the approximating formulae were reviewed, and a sample of these formulae were cross-validated using the reported items of the nearest neighbors as proxies for the unreported items. The sponsor subsequently reported that it was using the imputation formulae developed by our division, and that they were having considerable impact in improving the quality of editing and imputation for missing data items. The sponsor also suggested cooperating with staff to write a report or paper to give a complete appreciation of this aspect of the processing of the RFS.

Staff: Yves Thibaudeau (x31706), Richard Levy (HHES), Howard Savage (HHES)

1.14 NCES POVERTY STATISTICS (Demographic Project 7165)

A. Research for Small Area Income and Poverty Estimates (SAIPE)

The purpose of this research is to develop, in collaboration with the Housing and Household Economic Statistics Division (HHES), methods to produce "reliable" income and poverty estimates for small geographic areas and/or small demographic domains (e.g., poor children age 5 to 17 for counties). The methods should also produce realistic measures of the accuracy of the estimates (standard errors). The investigation will include assessment of the value of

various auxiliary data (from administrative records or surveys) in producing the desired estimates. Also included would be an evaluation of the techniques developed, along with documentation of the methodology. (See related Project F under Projects 0351 and 1871 Statistical Methodology).

During FY2003, staff produced the state variance and covariance estimates using the VPLX code (revised) for the expanded Current Population Survey (CPS) supplement samples for income years 2000 and 2001. These direct replicate variance estimates were smoothed and used in the SAIPE state model for the final state SAIPE estimates of income year 2000, and the preliminary state SAIPE estimates for income year 2001. The direct replicate sampling variance from the expanded CPS supplement sample and the traditional March CPS supplement sample for income year 2000 were completed.

Staff conducted an empirical study to compare an alternative Bivariate Bayesian model using SCHIP01 (State Children's Health Insurance Program) and SS01 (2001 Supplemental Survey) - a test survey for the American Community Survey (ACS) data with the current SAIPE model for estimating state poverty ratios for income year 2000 for four age groups (age 0-4, 5-17, 18-64, 65+). The preliminary results showed that the prediction error variances of the state poverty ratio using the bivariate model is smaller than the prediction error variance of the univariate model (equivalent to the current SAIPE model) for most of the 51 states. Also, the direct variance estimate of SS01 is similar, and sometimes is smaller than the variance for the model based estimate for some states. It seems that there is significant potential for ACS data in SAIPE.

Staff finished the development, and enhanced the implementation of a SAS IML PROC routine to fit a variety of parameterizations of bivariate linear models with random effects based on Maximum Likelihood or Restricted Maximum likelihood methods. The computation of the empirical Bayes smoothed estimates and their variance, along with the AICs for model fitting was included in the program. A SAS macro program for bivariate mixed models has been modified to better accommodate various county and state models. The SAS program is capable of handling 15,000 to 20,000 observations, which makes it capable of fitting not just county models, but also models for school district estimates.

Staff has assembled a research database containing the data used in the county level production poverty models, as well as in some alternative models that have been examined. This database (in SAS data format) includes data from all production years (income years) 1993-1999). A code book has been written to document all of the variables.

Staff has begun some preliminary work on

researching for improvement of the poverty estimates at the school level. Currently, we are investigating the relationship between the IRS tax data and Census data. Initial analyses indicate that the IRS tax data will be useful in constructing school district poverty estimates. Several approaches are currently being considered. We plan to develop a prediction model to have more timely poverty estimates at the school district level during non-census years.

Staff: Elizabeth Huang (x34923), Jerry Maples, William Bell (M&S), Robert Fay (M&S), George Train (PRED)

1.15 COGNITIVE TESTING OF THE SCHOOLS AND STAFFING SURVEY TEACHER QUESTIONNAIRE (Demographic Project 7167)

The purpose of this research is to cognitively test changes in layout and question wording in the 2003-04 Schools and Staffing Survey (SASS) Teacher Questionnaire. This is a paper questionnaire. UserWorks, a contractor, is conducting the cognitive interviewing on this project. Census Bureau staff is supervising UserWorks.

During FY2003, interviews were completed with the 30 teachers, and UserWorks presented its findings and recommendations in a written report and oral debriefing with Census Bureau staff, the Education Statistics Services Institute, and the National Center for Education Statistics (NCES). In addition to the written report, individual summary reports of the 30 interviews, and the tape recording of the interview were provided to the Census Bureau by UserWorks. Some of the recommendations were immediately accepted by NCES. When applicable, these recommendations were adopted by other SASS questionnaires prior to their cognitive testing. In order to preserve comparison to other data collections and other estimates, other recommendations, while valid, were not accepted.

This project has been completed.

Staff: Beth Nichols (x31724), UserWorks

1.16 ANALYSIS AND FORECASTING OF DEMOGRAPHIC TIME SERIES (Demographic Project TBA)

This project will employ time series methods to incorporate probabilistic estimates of forecast uncertainty into annual population projections.

During FY2003, staff familiarized themselves with the Kohler-Ortega methodology for adjusting birth order-specific fertility for tempo effects. Staff modeled time series of the Kohler-Ortega parameters (three parameters for each of five birth orders in three race groups), using ARIMA models. Forecasts of the parameters were produced from these models.

Alternative ARIMA models for the parameter series were compared, and results were reviewed with the Population Division staff. Staff also began to investigate handling correlation across races and/or birth orders by modeling and forecasting differences from a baseline race/birth order.

Staff met regularly with Population Division staff to explain results, receive feedback and determine next steps. Staffs discussed data limitations, including short series resulting from race detail, and use of CPS estimates to construct the series.

Staff: Kellie Wills (x31722), William Bell (M&S)

1.17 DISCLOSURE LIMITATION METHODS (Economic Project 3420051)

The purpose of this research is to develop disclosure limitation methods to be used for all Census Bureau publicly available economic data products. Emphasis will be placed on techniques to implement disclosure limitation at the stage of data processing. Methods will be developed, tested, evaluated, and documented. We will also aid in the implementation of the methods.

During FY2003, a formal Interagency Agreement (IAA) between the Bureau of Transportation Statistics (BTS) and the Census Bureau was established. This agreement states that a small number of Census Bureau staff with experience in statistical disclosure control methods will, over the course of the next year, evaluate software that BTS sponsored that implements a perturbation method called controlled tabular adjustment. Staff communicated frequently with BTS' confidentiality officer, in writing this IAA.

Staff met with staff from Field Division to discuss their application of the tabular control methods of rollups and cell suppression to foreign trade tables about to be released to ensure that the requirements of the disclosure review board were satisfied. In addition, they discussed what is involved in transitioning from the traditional (n,k) rule to the newer p% rule. Staff wrote a short, informal report comparing those two rules.

Staff explored the cost of doing simultaneous protection of all the primaries in a given table in a single linear programming problem compared to the traditional method of protecting the primaries sequentially. The latter certainly takes longer, but appears to produce more favorable suppression patterns. However, there may be situations in which the differences between the patterns is not significant, but the speedup is. Staff will need to expand the test program for a specific table to a general program for simultaneous suppression in order to get a good idea of the typical tradeoffs.

The multistage suppression software for fine-tuning

the suppression patterns on large 3 dimensional tables with row groups seems to be working. However, more testing is planned. A three-page manual for running the software, with explanations of each step is available.

Staff: Laura Zayatz (x34955), Paul Massell, Phil Steel, Sam Hawala

1.18 TIME SERIES RESEARCH (Economic Project 3420052)

A. Seasonal Adjustment Support

This is an amalgamation of projects whose composition varies from year to year, but always includes maintenance of the seasonal adjustment and benchmarking software used by the Economic Directorate.

During FY2003, seasonal adjustment and X-12-ARIMA support was provided to: Bank of America, Eviews, Inflow, IEE, NTC Research (UK), Sanyo Management, SAS Institute, Gates Corporation, FAME Information Services, Pacific Bell, Renault, TTX Corporation, Federation of Chilean Industry, Center for Economic Policy Research, Foundation for International Business and Economic Research, Bureau of Labor Statistics, Bureau of Economic Analysis, Federal Reserve Board, Dallas Federal Reserve Board, New Jersey Department of Labor, Pennsylvania Department of Labor and Industry, Australian Bureau of Statistics, Statistics Canada, Statistics Denmark, Statistics Norway, Statistics Netherlands, Government of Taiwan, Statistics Sweden, Instituto Nacional de Estadísticas y Censos (Argentina), Department of Treasury (South Australia), West Moreton Health Service District (Australia), INEGI (Mexico), Statistik Austria, MITI (Japan) Bank of Cyprus, Bank of England, Bank of Mozambique, Reserve Bank of South Africa, Bank of Mexico, Bank of Albania, Bank of Korea, Bank of Spain, European Central Bank, IMF, Universidad Austral (Argentina), Free University (Amsterdam), Humbolt University (Germany), University of Tucuman (Argentina), Washington University, University of West Florida, City University of New York, University of Southern Mississippi, Wake Forest University, South Dakota State University, Waseda University (Japan), University of Connecticut.

Staff: Brian Monsell (x31721), Kellie Wills, David Findley (M&S)

B. X-12-ARIMA Development and Evaluation

The goal of this project is a multi-platform computer program for seasonal adjustment, trend estimation, and calendar effect estimation that goes beyond the adjustment capabilities of the Census X-11 and Statistics Canada X-11-ARIMA programs, and provides more effective diagnostics. This fiscal year's

goals include: (1) finishing a release version of the program for the general public that includes the automatic time series modeling capability of the TRAMO/SEATS seasonal adjustment program and (2) further improvements to the X-12-ARIMA user interface, output and documentation. In coordination and collaboration with the Time Series Methods Staff (TSMS), the staff will provide internal and/or external training in the use of X-12-ARIMA and the associated programs, such as X-12-Graph, when appropriate.

During FY2003, staff continued development of the X-12-ARIMA seasonal adjustment program, adding additional features to (a) improve seasonal adjustment diagnostics for indirect adjustments with significant level shifts at the request of the Bank of England, (b) add additional features to output saved by X-12-ARIMA at the request of the Time Series Methods Staff of the Economic Statistical Methods and Programming Division, (c) produce indirect seasonal factors that are not confounded with calendar effects estimated for the component series, and (d) allow user to choose the series whose yearly totals are used as a target when forcing the yearly totals of the seasonally adjusted series.

Staff produced a Beta version of Version 0.3 of X-12-ARIMA, with documentation of the new automatic model identification procedure based on TRAMO/SEATS.

Staff: Brian Monsell (x31721), Kellie Wills

C. Research on Seasonal Time Series - Modeling and Adjustment Issues

The main goal of this research is to discover new ways in which time series models can be used to improve seasonal and calendar effect adjustments. An important secondary goal is the development or improvement of modeling and adjustment diagnostics. This fiscal year's projects include: 1) collaboration with the Time Series Methods Staff (TSMS) of the Economic Statistical Methods and Programming Division (ESMPD) in the further evaluation of the TRAMO/SEATS model-based seasonal adjustment program; 2) the further development of a version of X-12-ARIMA that calls SEATS so that X-12-ARIMA diagnostics can be used to analyze SEATS adjustments, and also so that, when appropriate, SEATS adjustments can be produced by the Economic Directorate.

During FY2003, staff revised the program (provisionally named X-12-ARIMA/SEATS) that combines X-12-ARIMA with the leading fully model-based seasonal adjustment program SEATS (from the Bank of Spain) several times. The revisions incorporated research on diagnostics for model based seasonal adjustment as well as capabilities needed for model-based seasonal adjustments research by our

division and ESMPD.

Staff conducted several research studies on topics including (a) using shrinkage estimators to improve seasonal adjustments, (b) improving a misestimation diagnostic of SEATS, and developing a model-determined sliding spans stability diagnostic for SEATS, and (c) studying the seasonal adjustment properties of some new time series models that generalize the most widely used time series model. Staff developed software to estimate component models and calculate canonical decompositions for ARIMA models in Ox to facilitate much of this research.

Staff: Donald Martin (x33689), Kellie Wills, John Aston, Richard Gagnon, Brian Monsell, David Findley (M&S)

D. Supporting Documentation and Software for X-12-ARIMA

The purpose of this project is to develop supplementary documentation and supplementary programs for X-12-ARIMA that enable both inexperienced seasonal adjusters and experts to use the program as effectively as their backgrounds permit. This fiscal year's goals include collaborating with the Time Series Methods Staff (TSMS) to develop a new and improved version of the X-12-ARIMA Reference Manual in WordPerfect with an extensive index and other aids not available in the TeX version. It also includes the writing of research articles extensively documenting certain research studies completed in FY 2000.

During FY2003, staff finished an extensive paper, "Frequency Domain Analyses of SEATS and X-11/12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series," demonstrating how and why some fundamental filter diagnostics used by the signal processing community should be applied to model-based seasonal adjustment filters and how the model based approach should be understood, and submitted the paper for journal publication. Staff finished another paper for journal publication: "Computation of Asymmetric Signal Extraction Filters and Mean Square Error for ARIMA Component Models."

Staff produced a version of the X-12-ARIMA Reference manual and the X-12-ARIMA Quick Reference in LaTeX for version 0.3 of the X-12-ARIMA program, converting the documentation from the original TeX format.

Staff: Brian Monsell (x31721), John Aston, Michael Furchtgott, Donald Martin, William Bell (M&S), David Findley (M&S)

1.19 IMPLICIT RATIO EDIT GENERATION (Economic Project 3320054)

Survey data editing using the Fellegi-Holt model of editing requires the complete set of explicit and implicit edits. The purpose of this project is to develop SAS software to generate the implicit edits for a given set of ratio edits. The new software will be based on the Generate Edits methodology currently used in the Census Bureau's Plain Vanilla Ratio Module and the SPEER edit systems. This is a joint project with the Economic Statistical Methods and Programming Division (ESMPD).

During FY2003, staff delivered two separate software, one written in SAS/IML, on in SAS/OR, and a User's Guide along with our recommendations. We recommended that the faster SAS/IML software be used for production processing when time is an issue, and the SAS/OR program be used for parameter development when a more detailed examination of the implied edits is needed. An SRD Research Report was written describing the research and all possible alternative methodologies, and a Study Series Report with a detailed description of the two edit generation programs. We completed adjustments to the software per customers' requests.

This project has been completed.

Toward the end of FY2003, staff met with several individuals from the Economic Statistical Methods and Programming Division, Company Statistics Division, Foreign Trade Division and Service Sector Statistics Division to identify possible projects for development with the Economic Directorate. In this meeting, we discussed the possibility of developing strategies for identifying recontact and follow-up of unusual publication cells for more careful scrutiny for three separate economic surveys.

Staff: Maria Garcia (x31703), Bill Winkler, K.J. Thompson (ESMPD), Roger Goodwin (ESMPD)

1.20 STATISTICAL CONSULTING POSTAL RATE COMMISSION (Methodology and Standards Project 7676)

The work associated with this project will entail the review of testimony, interrogatories, decisions, and other documentation relating to proceedings of the Commission in order to identify major statistical issues and provide relevant consultation. The consultation will include: 1) the briefing of the commissioners and other commission officials on the ramifications and desirable approaches to the identified statistical questions; and 2) the presentation of written summaries of the major findings from all assigned reviews.

During FY2003, staff provided extensive briefings for the Commission's Research Analysis and Planning Staff and conducted detailed technical reviews of survey reports and other documentation relating to an

evaluation of the effects of specific restrictions on international mail. We also reviewed the Commission's procedure for estimating the distribution of carrier routes, costs and revenue from the City Carrier Cost System (CCCS), and provided an assessment of the principal reasons for an apparent discrepancy between the CCCS-based estimates and those derived from administrative cost data.

Staff: Leroy Bailey (x34917)

1.21 NATIONAL PARK SERVICE GRANT PROGRAM (Methodology and Standards Project 7566091)

The National Park Service (NPS) administers the Urban Park and Recreation Recovery Program to provide Federal grants to economically hard pressed communities for the rehabilitation of critically needed recreation area facilities and the development of improved recreation programs. The Census Bureau is developing a model using Census 2000 data to assist the NPS in the selection of grant recipients.

During FY2003, the list of eligible jurisdictions, based on Census 2000 population data, was compiled and submitted to the NPS. The NPS staff reviewed the list and requested that we add back in locations that were eligible in the previous study, but no longer met the selection criteria. Data for the selected jurisdictions have been obtained from the Centabs system and are being formatted for input into SAS programs to obtain standardized scores so that the jurisdictions can be ranked. A final report and the accompanying data file are being prepared.

Staff: Pam Ferrari (x34993), Carol Corby

1.22 NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY - BAYESIAN STATISTICAL METHODOLOGY (Methodology and Standards Project 8863)

The purpose of this project is to provide technical expertise to the Statistical Engineering Division of the National Institute of Standards and Technology towards setting up Bayesian Computation software and developing Bayesian methods of Meta-analysis.

During FY2003, staff provided technical expertise to the Statistical Engineering Division towards setting up Bayesian Computation software and developing Bayesian methods of Meta-analysis. Staff prepared a report on Bayesian partition modeling with an example. We are working on the development of certified data sets for MCMC sampling and developing random effects methods for certifying standard reference food

materials for gas chromatography analysis in the presence of corrupted material. Staff is assisting with Bayesian programming on other projects.

Staff: Don Malec (x31718), Mike Ikeda

**1.23 PROGRAM DIVISION OVERHEAD
(Census Bureau Project 0251)**

A. Division Leadership and Support

This staff provides leadership and support for the overall operation of the division.

Staff: Tommy Wright (x31702), Hazel Beaton, Alice Bell, Maria Cantwell, Pat Cantwell, Robert Creecy, Manuel de la Puente, Barbara Palumbo

B. Computer Support

The Computer Support staff provides computer support with the goal of providing a statistical computing environment that provides researchers powerful tools to develop new methods and permits them to share information easily and accurately. Hardware includes SUN servers, workstations, and PCs on a NOVELL network.

Staff: Chad Russell (x33215), Tom Petkunas, Mohammed Chaudhry

2. RESEARCH

2.1 & 2.2 GENERAL RESEARCH (Census Bureau Project 0351) (Methodology and Standards Project 1871)

Statistical Methodology

A. Disclosure Limitation Methods

The purpose of this research is to develop disclosure limitation methods to be used for all Census Bureau publicly available data products. Emphasis will be placed on techniques to implement disclosure limitation at the stage of processing. Methods will be developed, tested, evaluated, and documented. We will also aid in the implementation of the methods. (Also partly funded under Economic Project 3420051 and Decennial Project 6121-H).

During FY2003, staff updated Disclosure Review Board rules for special tabulations and developed Q&As on disclosure limitation for the Director's Office.

Staff key-worded the methods section for the American Statistical Association's Privacy and Confidentiality website. This makes sure that searchers on specific terms will include certain references. A poster session was presented at the 2003 Joint Statistical Meetings on the ASA P&C committee's new website.

Staff contributed to a paper on confidentiality and data access for the Conference of European Statisticians. This contributed paper was not delivered, but rather it was available to participants to enhance the discussion. The paper supported the discussion on "Access to Microdata - Issues, Organization and Approaches."

Staff helped staff from the Demographic Surveys Division develop an agenda and list of presenters for a National Science Foundation Confidentiality Workshop. Staff acted as rapporteur for the technical session of the workshop.

Staff wrote a summary of past practices for the "Census Confidential" designation. Staff also worked with staff from the Housing and Household Economic Statistics and Demographic Surveys Divisions to develop disclosure limitation procedures for the Residential Finance Survey and the American Time Use Survey.

Staff worked with the LEHD staff to develop new disclosure limitation techniques for the Bureau of Transportation Statistics Origin-Destination data. The LEHD staff is going to attempt to develop high quality masked simulated block level data.

Staff is a member of Nontechnical Options for Reducing the Risk of Disclosure among Demographic Data (NORRDDD - a group under the Privacy

Principles Research Committee). The group drafted its recommendations and sent them out for comments.

Staff published a Federal Register Notice concerning the disclosure review process at the Census Bureau.

A pdf version of the *FCSM Statistical Policy Working Paper #22* (Report on Statistical Disclosure Limitation Methodology, May 1994) is now available on the FCSM website (www.fcsm.gov/working-papers/spwp22.html).

Staff conducted re-identification experiments on Census Bureau public using microdata and public records available on the Internet and communicated results. They found some "special uniques" that are at risk of being identified. Special uniques are records that are unique on a combination of key variable values at high levels of geography. Staff extended the SUDA algorithm (Special Unique Detection Algorithm). The extension to the algorithm identifies all record-level unique group attribute combinations (unique group attribute sets) up to a user-specified maximum size, then grades the "risk" of each record by considering the number and distribution of unique group attribute sets that it contains. Records of respondents belonging to a unique group may become at risk of re-identification. The output of the algorithm helps in reducing this risk.

Staff worked on automating downloads from publicly available databases. This automation involves a complex JAVA program, which allows for searches by multiple months of births, multiple years, or multiple counties. We will then attempt to link the output to publicly released Census Bureau microdata.

Staff: Laura Zayatz (x34955), Philip Steel, Paul Massell, Sam Hawala

B. Research File Development-Decennial Application (See Decennial Project 6121)

C. Small Area Estimation-Demographic Applications (See also Decennial Project 6121 and Demographic Project 5385)

This project reports on general small area research which complements the specific efforts for the decennial and demographic projects.

During FY2003, staff completed a project on small area estimation that includes a model of within small area variance. The aim of the project was to demonstrate the feasibility of using a unit level model for ACS that does not require accepting any design-based estimates of variance as having zero variance (as is assumed for most applications of area-level models). Using ACS data from the 1996 Oregon sample, staff made a direct comparison between unit and aggregate-

level logistic models with the aim of providing evidence of any biases in using an approximate-level model. Staff will provide some model diagnostics based on Bayesian prediction intervals. A manuscript is in final preparation for an SRD Research Report and submission to a journal.

Staff: Don Malec (x34892)

D. Compensation for Nonresponse in Longitudinal Surveys

This project requires an extensive examination of relationships between longitudinal survey nonresponse and potential explanatory variables for a variety of survey items. The research objectives are: to 1) apply the results of this investigation in the development of general analytical models which reflect potential survey errors in estimation and analysis ascribed to longitudinal nonresponse, and 2) identify and advance a well-defined process for selecting and evaluating desirable approaches to nonresponse compensation for longitudinal surveys. See Highlights from Demographic Project 1465 C.

Staff: Leroy Bailey (x34917)

E. Seasonal Adjustment (See Economic Project 3420052)

F. Small Area Estimation Methodology (for SAIPE Applications)

This work generally concerns the relative merits of small area estimation methods based upon linear (Fay-Herriot) models versus General Linear (mixed-effect logistic) models for log-counts of child-poor and log-rates of child poverty at the level of county SAIPE estimates. The methods are compared with respect to simulations (based on existing CPS and IRS predictors) and actual model fit based on historical data. (See related Demographic Project 7165.)

During FY2003, staff prepared a census technical report summarizing simulation studies performed over the past two years comparing analysis methods based on Fay-Herriot and logistic-regression mixed-effect models.

Staff completed the revision of a journal paper on alternative model analyses and model-comparisons on SAIPE county-level estimation of child poverty rates.

Staff worked on re-analyses of Fay-Herriot models and residuals on post-1993 SAIPE data, with emphasis on evaluation of Small Area Estimates versus Census 2000 county child-poverty rates. This work was summarized in a completed conference paper to be presented at the Federal Committee on Statistical Methodology Research Conference.

Staff: Eric Slud (x34991)

G. Sampling and Estimation: Economic Surveys

Estimating Low-valued Exports

In this project, we investigate new procedures to estimate the low-valued component of U.S. exports. Each month the Census Bureau receives documents describing all U.S. exports over \$2500. Exporters are not required to report their "low-value" exports - those valued at or under \$2500. Factors determined in the late 1980s to estimate the low-value component of total exports for each country remain the same even as export patterns have changed.

During FY2003, staff developed a new method to estimate the low-value component that (1) accurately represents what it is currently being exported, and (2) can adapt over time with minimal review to reflect changes in export patterns. In the research, we analyzed data from the Automated Export System. Although most exporters report their data under this system, only about 20% provide details on their low-valued exports. Under our approach, we divide the low-valued component for a country into several pieces according to modes of transport and commodity codes, and derive factors for each piece. Where fewer data are available, information is drawn from the nearby countries. Finally, we combine the factors, weighing the amounts of data available and the variability and predictive ability of the information.

Initial evaluations on actual low-valued exports imply that the new procedure works quite well estimating the low-valued exports from detailed export filers. However, the effort continues to demonstrate how well it works with summary filers, those who provide no information on low-valued exports. Current work includes (a) explaining differences between estimates under the new procedure and what is expected according to past reconciliation studies, and (b) presenting a proposal for implementing the new procedure.

Staff: Pat Cantwell (x34982)

H. ADDITIONAL TOPICS

H.1. Disclosure Avoidance for Microdata

This project investigates various methods for masking microdata. Some approaches may be based on the additive noise methods introduced by Kim (1986). Another current approach for limiting disclosure risk for continuous variables in the public use microdata are top- and bottom-code and rounding. These approaches significantly distort variance and correlation between variables. Alternative approaches or remedies for the weakness will be investigated.

During FY2003, staff developed multiplicative noise models for masking continuous variables. For the

purpose of obtaining the first two moments of the unmasked data from the masked, a normal distribution was assumed for the unmasked data. New models were also developed without distributional assumptions. Staff developed new computational procedures for the mixtures of additive noise approach of Roque. Staff wrote new C++ rank swapping software with far faster algorithms than those of Domingo-Ferrer. Staff applied rank swapping, ordinary additive noise of Kim, and mixtures of additive noise to the Kim-Winkler data using the information-loss and disclosure-risk framework introduced by Domingo-Ferrer and Mateo-Sanz. Further research with different databases and additional information-loss metrics is in progress. Staff completed the following research reports on microdata confidentiality: “Using Simulated Annealing for K-Anonymity,” “Single Ranking Micro-Aggregation and Re-Identification,” and “Bayesian Networks Representations, Generalized Imputation, and Synthetic Microdata Satisfying Analytic Constraints.” Work on this project has been completed.

Staff also reviewed six papers on methods for generating synthetic data satisfying analytic restraints. The papers included those applying standard multiple-imputation models, hybrid multiple-imputation models in which only a subset of the data are generated, and maximum entropy models. Staff delivered a new version of the microdata re-identification software to users in our division and at Cornell University. Staff updated the list of microdata confidentiality references at the American Statistical Association’s Privacy and Confidentiality web site <http://www.amstat.org/comm/cmtepc/images/ConfidReferencesWinkler.pdf>.

Staff sent papers, software, a background reference information to researchers at Statistics Netherlands, University of Leipzig in Germany, Free University of Berlin, the Italian National Statistical Institute, Rutgers University, Statistics Austria, Vienna University of Economics and Business Administration, the medical school at the University of Nice in France, Duke University, Eurostat, University of Erlangen in Germany, the Food and Drug Administration, the National Center for Health Statistics, the University of Plymouth in the United Kingdom, and the University of Michigan.

Staff: Jay Kim (x34907), William Winkler, Robert Creecy, William Yancey

H.2. Survey Design Issues

Aspects of survey design methodology identified for further research includes such areas as decennial long form weighting, collapsing methods in weighting cell development, and smoothing for improved estimation. It is important for the Census Bureau to

maintain an ongoing program of research to aspects of survey design to improve its survey methods and practice.

When rows or columns with different coverage ratios are collapsed, the sample units in a row with a low coverage ratio lose portions of their weights to those in the other row which has better coverage ratio. We specifically expressed, in a mathematical form, the amount of loss and gain (adjustment factor) depending on the coverage ratios and the sizes of control counts of the involved rows. Using the formula, we developed new collapsing rows/columns strategies. When new PSUs are selected for a demographic survey, we try to maximally retain sample PSUs that were in the earlier design. We used Ernst’s algorithm for the 2000 redesigned sample. We investigated several approaches for increasing the rate of PSU overlap between the original and new samples. As multiple solutions exist for a linear programming problem, we tried an approach to pick a solution which provides a higher rate of overlap.

We also showed that the coverage ratios between two collapsed rows/columns are slightly different.

Because of staffing changes, work on this project has been suspended.

Staff: Jay Kim (x34907)

Statistical Computing Methodology

A. General Record Linkage Support and Analytic Uses of Administrative Lists

Under this project, staff will provide advice, develop computer matching systems, and develop and perform analytic methods for adjusting statistical analyses for computer matching error. (This project is also funded under Decennial Project 6121-E)

During FY2003, staff participated in the 2010 Decennial Census Unduplication subgroup that is investigating ideas for preventing or removing census duplicates. Our division will further develop *BigMatch* software for large-scale duplicate detection. Joint Statistical Research Division (SRD) and Decennial Statistical Studies Division proposals for the development of improved string comparators and better date-of-birth comparisons were approved.

Staff has been studying alternative advance string comparators. They have begun developing and testing alternative methods to the standard Jaro-Winkler string comparator in use with the division record linkage software. The ultimate purpose of new string comparators is to determine which string comparator functions best with scanned spelling error and relate the string comparator values to appropriate likelihood ratios. Staff investigated variants of the string comparator metrics and modified likelihoods used in the matching software.

At the request of Decennial staff, we added new features to *BigMatch* and ported the software to the Decennial Statistical Studies Division's VAX. The software is to facilitate Decennial Duplication tests in June and July. In Decennial Census testing with New York files and ten blocking passes, *BigMatch* software processes 100,000 pairs per second. This rate is far faster than commercial software.

Staff produced the reports, "An Adaptive String Comparator for Record Linkage," "Methods for Evaluating and Creating Data Quality," and "Data Cleaning Methods." Staff gave a presentation on record linkage at the National Academies of Science. Staff continued research in blocking, and reviewed a number of papers being presented at record linkage workshops at the International Conference on Machine Learning and the Association of Computing Machinery Conference on Knowledge Discovery and Data Mining. Staff reviewed the QA Guidelines for Record Linkage produced by Census Bureau staff.

Staff sent papers and information about record linkage to researchers at Abt Associates, Paribas in France, Meditech.com, Carnegie Mellon University, University of Southern California, Telecordia, Boston University, Basque Institute of Official Statistics in Spain, Lost Angeles County, Bureau of Labor Statistics, University of Karlsruhe, Germany, National Institute of Statistical Science, MIT, Oxford University, Department of Health in New South Wales, NCE Brazil, University of New Orleans, University of Pittsburgh, Statistical Center of Iran, University of Utrecht in the Netherlands, National Center for Health Statistics, University of Tuebingen, University of Rome, Israel Central Bureau of Statistics, and the Center for European Economic Research, Italian National Statistical Institute, CSIRO in Australia, University of Indiana School of Medicine, University of Minnesota, Hawaii Health Information Corporation, University of Pennsylvania, Statistisches Bundesamt in Germany, IAB.DE in German, University of Konstanz in Germany, dukas.upc in Spain, Australian National University, Eurostat, Princeton University and Rand, Oklahoma State University, NISS, Universitat de Valencia in Spain, NASA, Department of Health, New Jersey; the General Accounting Office; the Iranian Statistical Office; Austrian National Statistical Office; Statistics Canada; Harvard University; Israel Bureau of Statistics, Free University of Berlin; University of Waterloo.

On June 16, staff delivered a preliminary version of *BigMatch* software for two tests of Decennial duplication-detection methods. The main version is still due on January 21, 2004.

Staff: Bill Winkler (x34729), William Yancey

B. General Edit/Imputation Support

Under this project, staff will provide advice, develop computer edit/imputation systems in support of demographic and economic projects, implement prototype production systems, and investigate edit/imputation methods.

The SPEER editing software applies the Fellegi-Holt methodology to continuous economic data under ratio edits and a limited form of balancing. SPEER generates the needed induced edits during error localization for every edit-failing record. In this project, staff wrote new software for generating a subset of the implied edits prior to error localization using Fourier-Motzkin elimination. Using this methodology considerably simplifies SPEER editing because it reduces the computational effort during error localization.

Staff wrote the first version of a C++ program to analyze the edit generation of unobserved implicit edits for edit failing records. We used an example for having a better understanding of the prime covers in error localization. The example has 6 fields and 5 explicit edits. The complete set of edits has 18 edits, 13 of them are implicit edits. Staff compared the prime covers in error localization from the set of explicit edits and the complete set of edits. The total number of all possible data points (the code space) from the 6 fields is 576. The number of data points in any of the 5 explicit edits is 422, i.e., each of the 422 records is an edit-failing record. Staff applied the error localization in the C++ program to each of the 422 records and classified each of the 422 edit-failing records to the following 4 categories: A) records that fail **the explicit edits only**, B) records with which the prime covers from the failed explicit edits and from the failed explicit and implicit edits are **identical**, C) records with which the prime covers from the failed explicit edits and from the failed explicit and implicit edits are **overlapped**, and D) records with which the prime covers from the failed explicit edits and from the failed explicit and implicit edits are **disjoint**.

The initial results are: 1) 120 records (28.4%) are category A; 2) 185 records (43.8%) are category B; 3) 38 records (9%) are category C; 4) 45 records (10.7%) are category D; 5) 16 records (3.8%) are either category C or D because of more than 2 prime cover solutions; 6) 18 records (4.3%) are either category B or D because of more than 2 prime cover solutions.

Combining results 1, 2, and 6, staff would get 76.5% of the all edit failing records that can be obtained without generating new implicit edits. Staff will continue this research. In particular, staff need to consider more examples to show that the majority of the edit failing records will successfully get an imputed (or filled-in) record in error localization. If it is possible to generate as many implicit edits as possible, then the

percentage of the combined items 1, 2, and 6 would be higher.

The 1997 version of SPEER (Draper and Winkler, 1997) generates the needed induced edits during error localization for every edit-failing record. The new SPEER uses an auxiliary SAS program for generating a large subset of the induced edits prior to automatic editing. Using the implied edits generated a priori considerably simplifies SPEER editing since it reduces the computational effort during error localization. Staff wrote new routines for the system and rewrote 1997 versions of the error localization and imputation routines. Staff tested the new SPEER methodology using production data from the 1997 Annual Survey of Manufactures. The results show that more than 98% of the edit failing records are corrected in the first pass through the SPEER system. Staff is writing a first draft of an SRD Research Report describing the software and empirical results.

Staff participated in the 2010 Census Imputation Work Group and gave a presentation about the Nearest Neighbor Imputation Methodology of Statistics Canada. Staff continued work on a C++ program that analyzes the edit generation of unobserved implicit edits for edit failing records.

Staff developed artificial data that can be used for testing editing methods. The data satisfy ratio edits and a limited form of balancing (each field can appear in at most one balance edit). The data set contains 10,994 records, all edit failing records designed to fail different edit patterns (ratio, balance, and implicit). This data set has been used by Professor J.J. Salazar at the University of La Laguna, Islas Canarias, Spain for testing a new editing methodology to be presented at the UNECE meeting in Madrid.

Staff sent information and papers to researchers at: Australian National University, Oregon State University, and the Israel Central Bureau of Statistics.

Staff: Bill Winkler (x34729), Bor-Chung Chen, Maria Garcia, Yves Thibaudeau

C. Graphical Data Analysis

This project entails the promotion of graphics for exploratory data analysis and quality assurance, and presentations of data or findings. A primary goal is to demonstrate that the use of graphical data analysis is an important and easy to implement tool for discovering patterns within data, and assuring that statistical properties have been maintained. Research will be conducted by examining newly developed and already existing graphical software using current Census Bureau data. Papers and reports produced by the Census Bureau will be examined to determine if the display of information can be improved using either existing or new graphical techniques. The use of graphics will be encouraged by providing presentations, training, and

support where needed.

During FY2003, staff completed a broad research agenda. Staff reviewed graphics software packages and conducted an extensive literature review. A request was sent out to selected groups at the Census Bureau to solicit their interest in using graphics in their programs. Meetings with Services Sector Statistics Division have been attended in order to discuss the feasibility of using graphical procedures to enhance the edit program. We have requested data files so that work can begin and a formal proposal can be written. A staff member continues to serve on the Census Bureau Standards/Guidelines for Graphics Committee. The committee is nearing completion of a first draft of the standards to submit to the Methodology and Standards Council for review.

Staff: Pam Ferrari (x34993), Leroy Bailey, Ruben Mera, Todd Williams

D. General VPLX Development and Support

This project will develop new methods and interfaces for VPLX general variance estimation software. Staff will provide support for complex applications such as the Survey of Income and Program Participation (SIPP) and the Current Population Survey (CPS), create training materials, and provide training for applications of VPLX.

During FY2003, staff continued to provide ongoing Hot-Line support for VPLX to the four program directorates at the Census Bureau. Staff assisted with VPLX software to compute variance estimates for the SAIPE project. Staff assisted with VPLX software to compute variances and t-tests for the ACE Survey. Long-term support included research supporting county level variances for the CPS March Supplement, the standard errors of medians, and the Decennial Area's Social Security Number Validation Study. Staff played a minor role with Randal Zu Wallack's effort to code the VPLX median algorithm in SAS.

Staff continued populating the VPLX Content Area of the Oracle 9iAS SRD Portal database with a long-term goal of revamping the VPLX website and integrating it with similar websites that support the various computer applications across the Methodology and Standards Directorate.

It was decided that this group will better support variance estimation activities throughout the Census Bureau when the scope is expanded to support all variance estimation software used at the Census Bureau. Because of this, subsequent reports will be under the header "General Variance Estimation Development and Support."

Staff: Aref Dajani (x31797), Mary Ann Scaggs, Bob Fay (M&S)

E. SRD Portal Development

This project develops a web-based application to provide user controlled knowledge management. Portals will provide access to a wide range of web-enabled applications, will bring together diverse data sources, and will provide secure access to existing information, regardless of where the information resides. The SRD research portal will ultimately become an integrated starting point for accessing all division images and applications.

During FY2003, Release 2 of the Oracle 9iAS Portal was installed on the SRD machine in late March. A new method for submission of ASA and AAPOR papers was proposed using the Notification and Status portlet on the Portal Page. The time between submission and approval will be considerably reduced when all information is consolidated in the central location. Extensive testing of the Notification and Status portlets is being done and training and documentation materials are being written so that these new procedures can be implemented. The new software proved to have several flaws which have been overcome. It is anticipated that the new SRD Portal Home Page will be launched in early December.

Staff: Mary Ann Scaggs (x34966), Aref Dajani, Ned Porter

F. Statistical Computation for Linked Employer Household Data (LEHD)

The Linked Employer-Household Data Project is a cooperative effort among all of the areas of the Census Bureau to combine economic data with demographic data. Sources of data include the American Community Survey, IRS, and Social Security data. Using this data, researchers will now be able to perform analyses that help disentangle the effects of choices that firms make from the choices workers make.

During FY2003, new work was started comparing alternative sparse iterative techniques for both the fixed and the mixed model.

The Conjugate Gradient Program was successfully ported and tested on the new LEHD SUN computers. The Conjugate Gradient Squared (CGS) method was found not to provide faster convergence than the CG method as predicted last quarter. Preliminary work on the programming of treating covariates from the person and firm factors was begun.

Staff: Rob Creecy (x33207)

G. New Techniques for Statistical Disclosure Control in Tabular Data

Current techniques for protecting sensitive information in published tables involve cell suppression, but the problem is notoriously difficult. Its

intractability lies in the fact that there is an explosive number of possible cell suppression combinations that can be used to satisfy protection requirements while maintaining consistency in the table. Cell suppression has several disadvantages, one of which is that the omission of figures in a published table remind the reader that information is deliberately being withheld - a situation in apparent contradiction with the mission of an information dissemination agency.

During FY2003, staff began research on disclosure limitation issues: coding (Fortran) to test ideas for new "controlled tabular adjustment" heuristics; computational testing of large scale suppression using "controlled tabular adjustment" (CTA) techniques. Staff compared computation times with CPLEX 8 in two different platforms using very large data sets (up to 40x40x40).

Staff installed CPLEX 8.0 on SRDU11 and assisted in getting information from ILOG Support about the new licensing system for interfacing with callable library using the Sun Fortran compiler. Staff tested CPLEX 8.0 MILP callable library solver on various protected tables using optimal algorithm to find performance or CPLEX 8.0 and computational limits. The problem of discrete data editing is a certifiably difficult (np-complete) problem. The part of the problem involved with testing for consistency and redundancy can be stated as a prepositional logic formula. New commercial products by ILOG and other vendors provide the possibility of automating the solution to these problems. We explored feasibility of using constraint programming and ILOG's "OPL" product for attacking this problem.

Staff: José Dulá (x33690)

H. ADDITIONAL TOPICS

H.1. IT Accommodation Research

This research is intended to support the further development of new assistive technology and methods developed by the Archimedes Project of Stanford University. Previous research has shown that the initial devices can be applied several times more quickly than traditional aids, and be at least as effective. The purpose of the technology is to isolate assistive devices needed by people with disabilities to use Information Technology (IT). This has enabled people with disabilities to be placed in IT intensive temporary positions throughout the Census Bureau. Success on this project will enable the Census Bureau to take a leadership position, because of our unique environment, in meeting the goals of *Executive Order 13078*, "*Employment of Adults with Disabilities: Executive Order 13078, Employment*, and the requirements of *Executive Order 13164*, "*To Facilitate the Provision of*

Reasonable Accommodation,” and of *Public Law 105-220, Amendments to the Rehabilitation Act, Section 508.*

During FY2003, staff continued to serve as Chairman of NCITS V2, the IT Access Interfaces Technical Committee for IT Standards). NCITS V2 is developing standards that will support the development of Assistive Technology (AT) and Electronics and Information Technology (E&IT) that can be connected without complex tailoring or integration. This will make it easier for suppliers to provide government purchasers with 508 regulations compliant E&IT equipment in the future. This in turn will make it easier for government agencies to comply with Section 508 which will eventually make it easier to ensure that government employees and members of the public who have disabilities have access to government data and information that is comparable to that access provided all other government employees and members of the public.

Staff: Bill LaPlant (x34887)

Survey Methodology

A. Usability Laboratory

The Usability Laboratory conducts research on user-centered design (UCD) and evaluation methods, develops user interfaces for the future, and provides collaboration to Census Bureau projects wishing to employ UCD principles. The lab encourages user-centered design and evaluation of the products produced (or purchased) by the Census Bureau. These include electronic questionnaires, information Websites, and software applications for use by employees and customers.

Outreach Highlights: During FY2003, staff gave presentations about usability to staff from Mexico’s National Institute for Statistics, staff from Ecuador’s National Institute for Statistics and Census, experts from Russia’s GOSCOMSTAT, a distinguished guest from UK Statistics, JPSM student interns, Census Information Centers in Las Vegas, and new members of the Census Bureau’s Director’s staff. Staff consulted with three visitors from New Zealand about usability and two visitors from Statistics Canada. Staff gave a workshop about remote usability methods at a professional association, and presentations on usability to three professional associations. Staff participated in a panel discussion on electronic forms design.

Staff gave a seminar on plans for usability research on mobile computing devices and an SRD Seminar about electronic questionnaire design research.

Usability staff consulted with the IT Remedy Staff, the LEHD staff, and a Computer Assisted Survey Research Office staff about future usability testing. A

usability staff member visited the usability lab at the Social Security Administration in Baltimore, and several staff members assisted a researcher on an NSF-sponsored digital government project.

Contractors gave three presentations about our research at this year’s digital Government Research Conference, two presentations at the University of Maryland Human-Computer Interaction Lab’s annual open house, and two presentations to professional groups in foreign countries.

Usability lab staff continued active participation in the interagency digital government projects, and on the Census Bureau’s Web site Improvement Team. Staff helped respond to two recommendations from the Census Advisory Committee of the American Marketing Association about internet teaching tools and Web site marketing research.

Contracts Highlights: During FY2003, staff continued to rely heavily on our outside contractors: University of Maryland Human Computer Interaction Laboratory, University of Virginia/Erica Inc., Virginia Tech Computer Science Department, UserWorks, Inc., Usability Systems, Inc., and Noldus, Inc. We established a new, long-term contracting vehicle with UserWorks, our primary usability testing contractor. We worked with several Decennial contractors to conduct usability tests including Z-tech and ScanSoft (formerly SpeechWorks.) Staff augmented our contract with Virginia Tech to support a presentation of Census Bureau work at a professional conference. We renewed our contract for eyetracking hardware and software with the University of Virginia.

Facilities Highlights: During FY2003, staff expanded our remote testing capabilities to enable testing at Census Bureau remote sites and completed a project involving remote testing at several regional offices. We closed one of our three Census Information Centers remote sites (Los Angeles), and the Customer Liaison Office began work to establish a replacement site. The new site is affiliated with the Census Information Center at the University of Texas Pan American in Edinburg, Texas. New software was obtained to help evaluate accessibility for applications that run on Microsoft PC platforms and used it to evaluate a Census Bureau data product. We installed new equipment to record both sides of a telephone conversation onto the digital video tape during usability sessions and used that equipment in a usability study of an interactive voice response system. Staff upgraded the Noldus Observer logging software. Additionally, the lab now has the capability to record sessions directly to DVD. The new logging software will permit test administrators to log from DVD video. Written instructions were developed for using the new digital video editing system and conducting accessibility testing for inclusion in the Usability Lab knowledge

base. We configured one of the testing suites for testing mobile computing devices. We continued work on a knowledge repository for usability topics, and an Intranet site.

Staffing Highlights: We added two new researchers to our staff. A staff member was certified as a Level 2 Contracting Officer's Technical Representative (COTR) and will serve as COTR for contracts related to the Usability Lab.

Staff: Kent Marquis (x34719), Joyce Farmer, Safiya Hamid, Leslie Brownrigg, Larry Malakhoff, Betty Murphy, Beth Nichols, Erica Olmsted, Andrew Perez-Lopez, Mike Rosen, Lorraine Randall, Tom Reynolds

B. Usability Collaboration Projects

B.1. Quality Management Repository

The purpose of this joint project is to evaluate the usability of new features for a Census Bureau Intranet site, the Quality Management Repository (QMR). Developed by the Computer Assisted Survey and Research Office (CASRO), this site is intended to enable internal users to share, manage, and disseminate the Census Bureau's best practices, standards, and guidelines in quality management. Usability is key to actual usage and user-satisfaction levels.

During FY2003, staff consulted with CASRO on test planning and development of testing materials. We arranged for CASRO to conduct usability testing in the Usability Lab, and we discussed test findings with CASRO personnel. The report on this testing will be added to the HCI Memo Series. In collaboration with CASRO, usability staff created a card-sorting task for the QMR content. We conducted four card sorting sessions with users, and based on the analysis of the results, determined the primary categories for the main page. In response to a request, we prepared a guidelines template to be reviewed for inclusion in the QMR. Extracted from a presentation to the Federal Economic Statistical Advisory Committee, these guidelines deal with the display of edit messages to respondents in establishment surveys. Adding these guidelines to the QMR will help to build usability into computerized, self-administered questionnaires.

Usability staff conducted four more card sorting studies based on one of the primary categories established in the first card sorting study. Staff analyzed the results and determined the grouping of information on the second level search pages. This information assisted in the prototyping of a redesign of the Web site.

Staff: Erica Olmsted (x34893), Betty Murphy, Kent Marquis

B.2. Automated Listing and Mapping Instrument

The Usability Lab will investigate the usability of the Demographic Area Address Listing (DAAL) process that includes laptop software called Automated Listing and Mapping Instrument (ALMI).

During FY2003, staff and contractors completed data collection focusing on the "Locate Address" updating function. We were unable to arrange observations of the GAIL processes or other features of the ALMI. Staff met several times to reach consensus on central usability issues.

Staff and contractors retrospectively logged the "Locate Address" tasks on the observation tapes, computed individual task times for each field representative, and computed descriptive statistics. User performance times were considerably longer than expected. User satisfaction questionnaire data were analyzed. Users were very satisfied with the terminology used on the interface, but found the system frustrating to use and felt that tasks could not be performed in a straight-forward manner. We constructed an outline of the final report and the principal findings.

The contractor wrote a first draft of the final report. The quantitative results remain the same: user performance times were considerably longer than expected. For satisfaction, users were very satisfied with the terminology used on the interface, but found the system frustrating to use, and felt that tasks could not be performed in a straight-forward manner. In addition, staff identified 12 high priority usability violations and video clips to illustrate each. Staff produced the final report, with video clip highlights and gave the report to the sponsor.

Staff: Erica Olmsted (x34893), Kent Marquis, Betty Murphy, Larry Malakhoff, UserWorks, Inc.

B.3. Administrative and Customer Services Intranet Site Development & Testing

The purpose of this study is to assist the Administrative and Customer Services Division (ACSD) staff redesign its Intranet Website so that it is user-centered. The project will assist the ACSD design team go through all the steps of user-centered design, including information architecture, low-fidelity testing, and usability testing, to come up with a site that is grounded in user-centered design principles. Staff will provide consultation, usability testing, and a final report.

During FY2003, usability staff continued to work on summarizing the results of the 12 usability testing sessions for the final report. Staff finished revising the final report and submitted it for final review. Staff produced the final report, with video clip highlights, and gave the report to the sponsor. This project has been completed.

Staff: Erica Olmsted (x34893), Kent Marquis

B.4. Field Directorate Intranet Site Usability Testing

The Field Division (FLD) at headquarters and the Technologies Management Office (TMO) wish to conduct a usability evaluation, including remote tests at regional offices and telephone centers, of the second phase of their Intranet Web site, based on the Oracle Portal architecture. The Usability Lab will provide usability testing, an interim report, and a final report.

During FY2003, usability staff developed the test protocol for the second iteration usability test. We created typical tasks, defined the user groups, and set usability goals. We then recruited users and conducted six usability test sessions with internal field headquarter users and six with internal field regional staff. We analyzed the results of the study, submitted a quick report to the sponsor.

Usability staff wrote the final usability test report and incorporated video clip highlights of the usability violations into the final product. Staff delivered the final report to the clients. This project has been completed.

Staff: Erica Olmsted (x34893), Kent Marquis

B.5. Housing and Household Economic Statistics (HHES) Information Architecture

The purpose of this project is to incorporate user-centered design principles in a systematic way as we begin the redesign of the HHES poverty, disability, and health insurance Web sites. A goal of the project is to establish a blueprint for incorporating usability and user-centered design features into all the HHES sites that are scheduled to be redesigned.

During FY2003, usability staff met with HHES staff and conducted three iterative rounds of low fidelity prototypes of a mock-up of the HHES poverty, housing and disability sites. Results after the first round showed that the prototype did not give the user an overview of the site. Usability staff worked with HHES staff to create a prototype that gave users a better overview. After each iteration, staff worked with HHES to refine the interface of the main page so that it gives users an overview of what is available deeper in the site.

In addition to the main level page, staff consulted with HHES staff to develop a data dissemination tool. Usability staff constructed a prototype template that uses a cgi script and offers users the opportunity to get an overview of site contents and to select multiple characteristics in order to find tables that contain the desired information.

Staff: Erica Olmsted (x34893), Kent Marquis, Juan-Pablo Hourcade

B.6. Decennial Usability Projects (See Decennial Projects 5010, 5201, 5202, 5604).

B.7. Congressional Web Site

The Director has asked the Usability Lab to work with the Congressional Affairs Office and the DADS staff to create a Web site containing Census Bureau information for congressional staff. The Usability Lab will assist with the site information architecture, usability design, and usability evaluation.

During FY2003, staff arranged for Ben Shneiderman to offer advice to the project through a conference telephone call. Usability staff met with the Web site development team to define the typical users and tasks that they will want to accomplish, and the usability goals for the site. Usability staff created a plan to include information architecture activities in the development of the Web site, including informational interviews and card sorting exercises with congressional staffers.

Usability staff worked on bringing user-centered design into the development of the Congressional Website. We conducted informational interviews with six congressional staffers, and wrote a summary of our findings. We created and conducted card sorting exercises with an additional seven congressional staffers, and presented our results to the Congressional Website team. We conducted low-fidelity prototype testing of the web site with three different users and wrote the initial results. In the results, we included video clip highlights of the usability violations that the users encountered. This project is complete.

Staff: Erica Olmsted (x34893), Kent Marquis, Tom Reynolds

B.8. User-Interface Standards and Guidelines

The purpose of this activity is to develop and maintain user-interface design standards and guidelines to foster consistency, and accessibility across Web-based applications/sites for data-collection, data analysis, data dissemination, and internal Census Bureau applications. An objective is to link these guidelines to other IT standards and style guides developed for user-interface designers and implementers.

In March 2003, a new Standards Development team was formed and trained by the IT Standards and Uniform Products Program. The team undertook the revision of IT Standard 20.0: Design Requirements and Guidelines for Web-based User Interfaces. Sub-groups were formed to address three major issues: integration of related standards, requirements for graphics, and client-side requirements. Staff began serving on the integration sub-group, which focused on combining the related portions of IT Standard 20 and IT Standard 15.0.1: Internet Document Standards. The various

subgroups and the full team continued meeting and agreed on an outline for the revised standard. Staff agreed to focus on informing developers about requirements for Web sites and applications, with supporting rationales to be made available in appendices.

Staff: Betty Murphy (x4858)

B.9. American Community Survey Usability Study

The purpose of this study is to discover usability problems with the American Community Survey (ACS) domain of Census.gov. Our objective is to assess the efficiency, accuracy, and subjective satisfaction of the site, note the usability violations, and make recommendations for how to improve the site.

During FY2003, Usability staff met with ACS staff to understand the typical ACS web site users, and to create representative tasks that users perform. We then conducted keystroke level modeling (KLM) of the tasks. Working with the ACS developers, we set the efficiency, accuracy, and satisfaction goals for the web site. We completed six usability test sessions, data logged the sessions, and conducted analysis of the data. We then wrote the initial quick report, including inserting video clip highlights of the usability problems and sent the initial quick report to the client.

Staff: Erica Olmsted (x34893), Tom Reynolds

B.10. Accessibility Testing

This project focuses on accessibility of desktop applications by blind and low vision users in accordance with the Section 508 regulations.

Desktop Applications: During FY2003, Staff from the Administrative and Customer Service Division (ACSD) contacted the Usability Lab about the accessibility of their retail product - "Census 2000 Modified Race Data Summary File." This product allows a data user to make cross-tabulations of census data by dragging and dropping variables. The initial analysis by the Winscreamer diagnostic tool revealed that some controls were not compliant with the Section 508 regulations. These findings were detailed in a report, "An Accessibility Evaluation of the ACSD Race Browser" to staff from ACSD. The software did not allow a blind user to make data extracts or browse the output tables. Also, documentation, help browser controls, and tables were not accessible. A recommendation was made to pass these comments on to the vendor so they could be corrected.

Staff began an evaluation of Landview (GEO). This product displays a thematic map showing features requested by the user in their query. Staff documented accessibility violations for each screen. A report is being compiled from this documentation.

Staff evaluated an installer program for a member of the division. Staff tested the installer program with Winscreamer accessibility evaluation software and verified findings with the JAWS screen reader program. These findings were reported to the division member, and he then created new versions of the installer program to address the accessibility violations. Three versions were evaluated, each with improved accessibility, but not full compliance. The comments were forwarded to the contractor who created the software for their action

Web Applications Accessibility: This project focuses on accessibility of Internet and Intranet applications by blind and low vision users in accordance with the Section 508 regulations.

Staff collaborated with staff from the Decennial Management Division on an accessibility test plan for the Internet of the 2003 Census Test, performed accessibility testing on the application, and provided a report on accessibility violations. Staff was contacted by staff from the Systems Support Division to assist with checking the accessibility of the American FactFinder web site. A memo was sent detailing the facts.

Staff: Larry Malakhoff (x33688)

B.11. Economic Census Web Usability Project - 2003

The goal of the project is to provide information that can be used by the Economic Census site designers to increase user performance and satisfaction with the site. The usability test has two sections: the prototype test and the live site test. The live site test has three objectives: have changes made after 2001 testing addressed the navigation issues, is the Economic Census site sufficiently easy to find and use for simple queries, and the optional objective of assessing the effectiveness of links to other economic area data sets. The prototype test will consist of testing three prototypes: the existing and graphically rich prototype, a scaled down prototype of the existing site, and a third text-based prototype developed at the Usability Group and based on principles of Information Architecture.

During FY2003, staff completed high-level task analyses and user group definitions by a review of user queries, internal marketing reports, analysis of site structure and information, Aria reports, and communication with site designers. The client approved both task analysis and user group definitions. Sample tasks, representative of those to be used during testing, were created by the Usability Group and approved by the client. An understanding between the Usability Group and the client was reached regarding performance objectives. Staff recruited 21 subjects for the usability study and ran them through the usability

test of 3 versions of the low fidelity prototypes and the live site. Staff data logged the sessions and analyzed the data. Staff composed an initial quick report and sent it to the client. Staff produced the final report, with video clip highlights and gave the report to the sponsor. This project has been completed.

Staff: Eric Olmsted (x34893), Mike Rosen

C. Usability Research Projects

C.1. Bureau of Economic Analysis (BEA) Cognitive and Usability Testing

To assist the Bureau of Economic Analysis (BEA) with improving data quality and response rates to its surveys, this project will focus on reviews of both paper and electronic versions of their current and upcoming forms. It will propose and conduct methodological studies to improve the questionnaires and related information.

During FY2003, staff began assisting the BEA with improving data quality and response rates to its surveys. This project will first focus on improving select surveys. Methodological studies to improve questionnaires and supporting respondent documents will be the focus. The Economic Statistical Methods and Programming Division will receive direct funding from BEA and involve our division staff in methodological research. The methods proposed include expert reviews, cognitive testing, and debriefing.

Staff: Eileen O'Brien (x32695), Beth Nichols

C.2. Usability Research Projects User-Interface Standards and Guidelines

With the April 2002 publication of IT Standard 20.0.0, *Design Requirements and Guidelines for Web-based User Interfaces*, the original team disbanded. The IT Standards and Uniform Products Program has recruited a new team for the purpose of reviewing and revising this standard. One objective is to produce an up-to-date source for developers of Web-based applications/sites for data collection, data analysis, data dissemination, and internal Census Bureau applications. Another objective is to link this standard to other IT standards and style guides developed for user-interface designers and implementers.

During FY2003, staff met with the chair of the new Standards Development Team (SDT) to pass on lessons learned from the initial standards-development effort. The SDT began to meet regularly, and established several sub-groups to work on various aspects of the revision effort. Staff joined the sub-group devoted to integrating IT Standard 20 with portions of IT Standard

15. Staff prepared materials for a card-sorting exercise, which helped the sub-group reach decisions about the organization of topics in the to-be-revised standard.

Staff: Betty Murphy (x34858)

C.3. User-Centered Design Knowledge Base

To raise awareness and assist the practice of usability and user-centered design at the Census Bureau, and maintain control over the quality and consistency of information disseminated by the Usability Lab, this project will develop, organize, implement, and maintain a formal knowledge base. Information in the knowledge base will be disseminated to people who design and acquire hardware, software, Webpages, and Intranet pages for the Census Bureau. The dissemination will be via documents (such as standards and style guides), via Intranet sites, in courses, and with templates for designing Web pages.

During FY2003, staff continued developing the set of content elements for the usability knowledge base. We continued specifying content elements to include in the database and reviewing information for inclusion in the content elements. We revised text for reading from the screen by reducing sentence lengths and adding sub-headings. We identified the need for graphics to illustrate the text.

Staff: Beth Nichols (x31724), Kent Marquis, Betty Murphy, Erica Olmsted, Larry Malakhoff, Tom Reynolds

C.4. Information Visualization Research

This research is to explore possibilities for developing and applying flexible data-visualization tools for users of Census Bureau data, in collaboration with Professor Ben Shneiderman of the Human Computer Interaction Laboratory (HCIL) at the University of Maryland, College Park. Last year, the group developed several proof-of-concept Web versions of a dynamic map interface to data. This year, the group will pursue the most promising possibilities.

During FY2003, the contractor has developed web-accessible dynamic map interfaces to Census Bureau data. These interfaces are designed to work with low-bandwidth connections. A new version is being developed using Java Swing, which should provide for more convenient for both users and programmers. The contractor has also begun to investigate the possibility of using sound to explore data. This user interface approach could prove particularly useful to visually impaired people.

Staff: Juan Pablo Hourcade (x33690), Kent Marquis, Human-Computer Interaction Lab at the University of Maryland, College Park, MD

C.5. DynaMaps for CD-ROMs

The purpose of this study is to create a dynamic map application that allows users to interact with Census Bureau data via dynamic queries, and see results on maps, scatter plots, and in tables. An objective of the usability staff and the contractor, Christopher North of Virginia Tech, is to create an interface that facilitates quick learning by novice users and keeps the needs of all users in the forefront of all design decisions. A successful application will be added to the spring 2003 release of the Counties CD-ROM by the Administrative and Customer Services Division.

During FY2003, the contractor performed several iterations of revisions to the interface at the request of the staff. The staff worked on shortening the names of variables displayed in the software. We also conducted remote usability sessions to evaluate the software and identify problems. The user interface to add and remove variables was substantially improved. Staff expects to receive new data for the software in early FY2004 and plans to release the software to the public soon afterwards.

Staff: Juan Pablo Hourcade (x33690), Erica Olmsted, Kent Marquis, Tom Petkunas, Virginia Tech Department of Computer Science

C.6. Usability Issues in the Design of Electronic Forms for Economic Data Collection

The purpose of this research is to identify usable design solutions that will allow establishment respondents to complete electronic forms quickly and accurately with little perceived cognitive burden. Staff and management from the Economic Directorate were involved in identifying the issues and crafting the test plan. The laboratory research was conceived of as the initial phase of a multi-year effort to investigate electronic form design issues in the establishment context. Later phases may include a test in the field during production or in an experimental panel.

During FY2003, in laboratory testing, test respondents worked with mocked-up survey prototypes incorporating different design options for the issues of interest. Sixty-nine University of Maryland undergraduate students completed the electronic business survey and the attached pre- and post-questionnaires. Data analysis was completed, and results were presented in an SRD Seminar in June. We found that many of the design changes tested did not affect data accuracy or respondent burden. Two design layouts did improve accuracy: 1) using a “stacked” design instead of a grid design improved data accuracy when the grid was wider than the computer screen, and 2) improved user accuracy was found when question text was modified to include key words reminding respondents to report a particular type of data. Eight

additional non-student subjects were run after being screened for experience in business administration and accounting. For the most part, their mean accuracy scores were within two standard deviations of the student scores suggesting that the student data could generalize to the non-student population.

Staff: Beth Nichols (x31724), Betty Murphy, University of Maryland Human-Computer Interaction Lab

C.7. DynaMaps Help Design and Testing

It was observed with previous iterations of the DynaMaps application, that users had a hard time learning basic functions, and the existing help system was not sufficient. Our goal is to create an easy-to-use and effective help system that provides “guided learning” in order to teach the user the basic functions of the application. The Usability Lab has developed several approaches, one based on brief video instructions for each major function, and another based on what look like “Post-it” notes that the user calls up when needed. The latter approach allows the user to work with the current interface while receiving help information. The University of Maryland Human Computer Interaction Lab (HCIL) has developed two different approaches; one gradually reveals parts of the interface to the novice user and demonstrates the key functions, the other uses audio sound files to explain key features.

During FY2003, the usability lab developed, iteratively tested, and implemented a “post-it” note approach for the new counties CD application. The HCIL developed their approaches, demonstrated them to an Interagency Group developing a Statistical Knowledge Network, and presented papers about them at professional association meetings. This project is complete.

Staff: Kent Marquis (x34719)

C.8. Cognitive Aspects of Questionnaire and Web Navigation

This is a long-term basic and applied research project with the University of Maryland, directed by Professor Kent Norman, to understand cognitive principles of questionnaire and Web navigation, with particular emphasis on issues of concern in Census Bureau applications.

During FY2003, students in Professor Norman’s class built a Web site to assist in the design of online surveys. The site provides principles, practice guidelines, examples and links to additional resources. The site is available at http://lap.edu/survey_design/. The students presented their work at an open seminar in December.

Staff: Kent Marquis (x34719), Betty Murphy, University of Maryland Laboratory for Automation Psychology and Decision-Making Processes

C.9. Mobile Computing Devices (MCD) User Modeling Study

This study is designed to obtain empirical usability data about user performance with MCDs as a function of basic features (e.g., input method, form factor, questionnaire design and navigation options). These empirical data will be used in developing general models that predict user performance with different kinds of devices, including those not yet built.

During FY2003, we purchased an ultra-light-weight camera that provides simultaneous views of both the user and the device. With the help of UserWorks, Inc., we evaluated mobile devices as data collection tools using 52 participants recruited from within the Census Bureau. Data analysis focused on developing a model that will be able to predict levels of user performance associated with mobile device interaction. Staff began to investigate the ability of automated key-stroke-level models to predict respondent performance on Web-based survey tasks.

Staff: Tom Reynolds (x34971)

C.10. Eye Tracking

The goals of the eye tracking project include a validation of pre-existing findings in eye tracking literature, developing design recommendations based upon the findings, and developing a benchmark protocol for future eye tracking evaluations.

During FY2003, staff began collaborating on the design of an eye tracking usability evaluation of the Census Bureau's homepage. In preparation for the usability evaluation, we reconfigured the Usability Lab in order to accommodate the eye tracking equipment.

Staff: Erica Olmsted (x34893), Andrew Perez-Lopez

C.11. Assessing Spatial Ability in Relations to Field Activities

This study will investigate alternative methods of assessing the spatial ability of enumerators in relation to their field tasks. The methods derived from this study will be used in future studies to determine whether enhancements to the mobile computing device (MCD) user interface actually help to compensate for the lower spatial abilities of some enumerators. The overall goal is to provide equivalent MCD usability for all enumerators, across a wide range of spatial abilities.

During FY2003, staff collaborated with Professor Sarah Nusser (Iowa State University) in planning the research, which is to be conducted in conjunction with the 2004 Census Test. This was coordinated with staff in the Decennial Management Division, Demographic

Statistical Methods Division and Field Division. The planned study grew out of previous research which staff presented at the Human-Computer-Interface International Conference in Crete, Greece. In a related activity, staff participated in the development and revision of a questionnaire to be administered to enumerators upon their completion of the 2004 Census Test.

Staff: Betty Murphy (x34858)

C.12. Interfaces to Census Bureau Data Designed for Children

This research aims to build educational environments for elementary school children. These educational environments will be designed to give access to Census Bureau data, facilitate the learning of useful skills and information, and generate positive attitudes in children, teachers, and parents toward the Census Bureau. Cooperative inquiry techniques that partner adult researchers with children will be used to generate design ideas. These ideas will in addition, be useful for the development of user interfaces for other populations, such as the elderly, and those without much statistical or geographical knowledge. Preliminary findings point toward the use of games as a form of scaffolding for learning about Census Bureau data.

During FY2003, staff wrote a proposal specifying a research plan for creating educational environments for children with an emphasis on accessing Census Bureau data. The proposal motivates the work, and includes a literature review. It also lays forward a plan to use games as a scaffold to make children interested in Census Bureau data. We also secured cooperation with the University of Maryland's Human-Computer Interaction Laboratory for support on developing design specifications for the software to be built.

Staff: Juan Pablo Hourcade (x33690), University of Maryland's Human-Computer Interaction Laboratory

D. Quick Turnaround Pretesting for Household Surveys

This project involves pretesting new or revised series of questions for insertion into household surveys. The projects are of the short-term, quick turnaround variety, rather than long-term research efforts to redesign a survey. Methods used include cognitive testing and other techniques as appropriate. (Also partly funded under Demographic Project 1465-E.)

During FY2003, staff worked on four short-term projects in addition to the three SIPP projects. Staff revised advance letters and informational brochures for the voluntary and mandatory panels of the American Community Survey (ACS) Voluntary Test (Landreth

2002). Results showed that the most salient piece of information contained in the letters was the data use statements. While these were viewed positively for the most part, they conveyed to respondents that the survey was about local communities. Without a questionnaire as a reference, respondents thought the survey would not contain personal questions, and would ask questions to determine the quality, frequency, etc., of their local community services. Some respondents to the mandatory letter did not understand that the survey was mandatory. Regardless of whether respondents received the voluntary or mandatory letter, they tended to think their own perceptions about the survey would be a larger determinant in their decision to participate than the legal requirements. Recommendations based on these results were incorporated into the final materials for the ACS Voluntary Test.

We conducted cognitive research on the Schools and Staffing Survey (SASS) Public School Questionnaire (Hunter and DeMaio, 2003). We found that revisions to questions related to average daily attendance, number of limited-English-proficient students and programs for them, and measures of school overcrowding worked fairly well. Respondents were familiar with the concept of average daily attendance, although they were more likely to provide it accurately as a percentage rather than a number of students. Limited-English-proficient was a term familiar to all respondents, even though they did not necessarily use it in their schools. With one exception, the measures of overcrowding worked well to identify schools that are crowded. However, due to complexities in the concept of historical building capacity, it did not work well to identify overcrowding. In other areas of the questionnaire: 1) respondents were inconsistent in whether they considered a school nurse to constitute “medical health care services;” 2) the real situations regarding summer school activities for extra assistance or enrichment were not captured properly by these questions; and 3) high school respondents improperly reported programs like immersion in a foreign language and were unsure about whether to report AP classes as part of a talented and gifted program. Recommendations for question revision were proposed in all these areas; most were either accepted or altered jointly between the Census Bureau and its sponsors.

Staff completed documentation of an expert review for the 2002 Police Public Contact Survey (Hughes, 2003). This report summarized problems inherent in the version of the questionnaire proposed by the sponsor (Bureau of Justice Statistics) as well as several iterative versions, and documented changes made to the questionnaire.

Staff conducted cognitive research on the National Crime Victimization Survey School Crime Supplement.

Cognitive interviews have been completed and a research report documenting the results is in progress.

Staff: Terry DeMaio (x34894), Kristen Hughes, Jennifer Hunter, Ashley Landreth, Lorraine Randall, Liz Aaker, Luc Perkins

E. Questionnaire Pretesting Activities

This project involves coordinating the Census Bureau's generic clearance for questionnaire pretesting research. Pretesting activities in all areas of the Census Bureau may use the clearance if they meet the eligibility criteria.

During FY2003, staff monitored the generic clearance, consulted with staff from other areas of the Census Bureau wishing to use the clearance, and kept the Office of Management and Budget (OMB) informed of all pretesting activities. Thirty-eight letters were submitted to OMB for work done under the generic clearance, with a total of 4,038.5 burden hours. Staff also chaired an interdivisional group including participation by the Economic Statistical Methods and Programming Division, Decennial Systems and Contracts Management Office, and Demographic Surveys Division, which was charged by the Methodology and Standards Council with developing corporate standards for pretesting censuses and surveys. The “Census Bureau Standard: Pretesting Questionnaires and Related Materials for Surveys and Censuses” was issued on July 25, 2003.

Staff: Terry DeMaio (x34894), Kristen Hughes

F. Questionnaire Design Experimental Research Survey - 1999 (QDERS)

Staff developed, coordinated, and implemented an omnibus questionnaire design experimental research survey (QDERS). This survey was a moderate-sized (target of 2000 completed interviews averaging 15 minutes each) RDD survey conducted through the Hagerstown Telephone Center in spring, 1999. The QDERS allowed the staff an opportunity to conduct questionnaire design field experiments in a timely and flexible manner.

This project has been completed.

Staff: Jennifer Rothgeb (x34986), Joanne Pascale, Jeff Moore

G. Questionnaire Design Experimental Research Survey - 2000 (QDERS)

QDERS 2000 is an omnibus survey designed to facilitate independent research related to questionnaire design issues, interviewer training, and other survey methodology issues. The QDERS 2000 was conducted from the Hagerstown Telephone Center in September

2000, using an RDD sample. Five researchers conducted questionnaire design experiments and two researchers conducted interviewer training experiments.

During FY2003, staff presented research results from the QDERS 2000 experiments at several domestic and international conferences. Data from QDERS 2000 also served as the bases for book chapters. The interviewer training experiment results led to additional field testing and are being included in three demographic production surveys beginning in 2004.

Staff: Jennifer Rothgeb (x34986), Jeff Moore, Joanne Pascale, Eileen O'Brien, Ashley Landreth

H. Questionnaire Design Experimental Research Survey -2002 (QDERS)

QDERS 2002 is an omnibus survey designed to facilitate independent research related to questionnaire design issues and other survey methodology issues. The QDERS 2002 will be conducted from the Hagerstown Telephone Center in June/July 2002, using an RDD sample. Three researchers conducted questionnaire design and survey methods experiments

During FY2003, it came to our attention that the QDERS 2002 data files had been erroneously deleted. Staff immediately put plans in place to re-conduct the QDERS 2002 data collection during summer, 2003.

Staff: Jennifer Rothgeb (x34986), Joanne Pascale, Ashley Landreth, Terry DeMaio, and Gordon Willis (National Cancer Institute)

I. Questionnaire Design Experimental Research Survey -2003 (QDERS)

QDERS 2003 is an omnibus survey designed to facilitate independent research related to questionnaire design issues and other survey methodology issues. The QDERS 2003 was a repeat of the content of the QDERS 2002.

During FY2003, the QDERS 2003 was conducted from the Tucson Telephone Center over an eight week period in July and August. Three researchers conducting questionnaire design and survey methods experiments participated. A researcher external to the Census Bureau also participated through collaborative research with staff. Data collection went smoothly and datafiles became available from the Technology Management Office (TMO) in late August. Staff immediately transferred the datafiles from TMO to our division. Staff currently have preparations underway to process the QDERS 2003 through the Demographic Surveys Division's TRANSCASES program in order to generate a SAS dataset for the involved researchers.

Staff: Jennifer Rothgeb (x34986), Joanne Pascale, Ashley Landreth, Terry DeMaio, and Gordon Willis (National Cancer Institute)

J. Questionnaire Design Experimental Research Survey - 2004 (QDERS)

QDERS 2004 is an omnibus survey designed to facilitate independent research related to questionnaire design issues and other survey methodology issues. Research proposals were reviewed and accepted for inclusion in the QDERS the QDERS2004 which is tentatively scheduled for April/May 2004. In order to increase sample size, additional funding support is being provided by the Demographic Surveys Division and the Social Security Administration.

During FY2003, preparations are underway to obtain cost estimates from involved divisions, and to develop questionnaire specifications for CAPI instrument authoring. With past administrations of QDERS, it has become clear that a higher number of sample cases are needed in order to obtain our targeted goal of completed interviews. For QDERS 2004, we will work with the Demographic Statistical Methods Division to determine if more sophisticated sampling procedures now exist which allow the RDD sample vendors to produce more efficient samples (i.e., fewer nonresidential numbers, cell phones, fax machines, non-working numbers, etc.).

Staff: Jennifer Rothgeb (x34986), Joanne Pascale, Jennifer Hunter, Nancy Bates (DSD), Jason Fields (POP)

K. Privacy and Confidentiality Research

The purpose of this project is to investigate the effects of public attitudes and perceptions about privacy and confidentiality on Census Bureau activities, and to research methods to address respondent concerns and discourage nonresponse behavior.

During FY2003, staff chaired and participated in an Interagency Household Nonresponse Group. Meetings of the subgroup provided a forum to discuss related research with which various agencies are involved, shared information, and discussed the possibility of collaborative research on issues related to privacy and confidentiality.

Staff: Tom Mayer (x34930), Eleanor Gerber, Melinda Crowley

L. Ethnography: Methods and Culture

Apply ethnographic research methods to ground key Census Bureau concepts, processes, and operations in evidence from direct observation. Ethnographic methods are generated from anthropological theory and studies of particular societies, and have been widely adopted by other social sciences.

During FY2003, staff participated in the external peer review of the program of ethnographic research at the Census Bureau, the internal working group on

residence rules, 2010 census content planning, and the interagency groups on Household Survey Non-response (Privacy and Confidentiality sub-group) and on migrant and seasonal farm workers. Staff also submitted draft and final reports for the Census 2000 ethnographic studies, disseminated results at professional meetings, and expanded the analysis of this research.

Staff: Leslie Brownrigg (x34995), Manuel de la Puente, Eleanor Gerber, Laurie Schwede, Melinda Crowley, Kristen Hughes, Yuling Pan

M. Continuities in Research on the Foreign-Born Population and Immigration Issues

This research effort calls for the application of varied qualitative research methods to solving problems that will improve the Census Bureau's ability to collect and provide timely, relevant, and quality data about the social and economic circumstances of the foreign-born population in the United States. It will conduct collaborative research with anthropologists, sociolinguists, and other behavioral scientists on foreign-born persons from different national origins to assess the barriers they experience to participation in the decennial census and demographic surveys.

During FY2003, staff started a research project on how recent immigrants from China interpret and respond to questions in censuses and demographic surveys, and how they react to the participation in government surveys. There are two parts to this project. The first is to conduct cognitive interviews in Chinese to assess respondents' understanding of key census questions, and to identify linguistic and cultural issues in translations, and in conducting cognitive interviews in a language other than English. The second is to conduct in-depth ethnographic interviews to identify sociocultural and attitudinal issues that are barriers to immigrants' participation in census and demographic surveys.

By August 2003, staff completed cognitive interviews and ethnographic interviews. We also completed linguistic analysis of all interviews. A paper outlining preliminary discussions on findings from the Chinese cognitive interviews was written and presented at the Annual Meeting of the American Association for Applied Linguistics. Staff made suggestions for methodological considerations in translations, pre-testing translated instruments, and in conducting cognitive interviews in languages other than English. Staff drafted a final report summarizing the main findings and recommendations from the research. The final report, "Impact of Language and Culture: Immigrants' Experiences and Participation in Census Surveys" is under review.

Staff: Yuling Pan (x34950), Eleanor Gerber, Suzanne

Scollon (Georgetown University), and Barbara Craig (Georgetown University)

N. ADDITIONAL TOPICS

N.1. Evaluating Pretesting Techniques for Finding and Fixing Questionnaire Problems

The objective of this research is to determine how well laboratory question testing methods predict the types of problems that will actually be experienced in the field, and to what extent the laboratory testing contributes to improved questions. This project includes research to determine not only the relative effectiveness of different methods for detecting questionnaire problems, but will also evaluate the methods in terms of their ability to provide information on question wording repairs.

During FY2003, we found from our research, that problems observed during laboratory pretesting do predict problems observed when the same items are administered in a field setting, therefore pretesting is effective at detecting problems. However, we found mixed results for the effectiveness of laboratory pretesting providing information necessary to "repair" problematic questions. Question problems identified as "respondent problems" are more easily "repaired" than question problems identified as "interviewer problems." Research results were presented at the Questionnaire Development and Evaluation Techniques Conference, and are contained in a chapter of a forthcoming Wiley publication, "Questionnaire Evaluation and Testing Methods."

Staff: Jennifer Rothgeb (x34986), Gordon Willis (National Cancer Institute), Barbara Forsyth (Westat)

N.2. Analyzing the Data from Cognitive Interviews

This independent research project focuses on the process by which cognitive interviews are conducted and analyzed by survey research organizations. The objective is to conduct an experiment to evaluate different types of interviewing and analytic procedures in terms of the substantive conclusions that are drawn and the recommendations for questionnaire revisions that they produce. This will facilitate a comparison of the analysis procedures used by the Census Bureau versus those used by other organizations.

During FY2003, staff presented the paper, "Cognitive Interviews: Do Different Methodologies Produce Different Results?" at the Questionnaire Development and Evaluation Techniques (QDET) Conference. Staff revised the paper in accordance with reviewers' comments for publication in the QDET monograph.

Staff conducted follow-up research in 2002 QDERS for this project. This involved a split panel test of the recommended questionnaires resulting from the application of three alternative cognitive interviewing methods as well as the control, under actual field conditions in a telephone survey. We completed behavior coding of the interviews taped during the data collection, and began analyzing the data. We also created self-study interviewer training materials and refusal conversion aids for 2003 QDERS, when the data collection had to be repeated due to the loss of the original data.

Staff: Terry DeMaio (x34894), Ashley Landreth, Tina Arbogast, Luc Perkins, Gloria Prout, Joyce Farmer

N.3. Interviewer Refusal Aversion Training

Based on the concepts of tailoring and maintaining interaction (Groves and Couper, 1998), Groves and McGonagle (2000) developed a theory-guided training protocol to sharpen interviewers' skills in averting a refusal.

Staff in our division adapted, experimentally researched, and field tested the design over the past three years - first in an omnibus RDD telephone research survey in 2000, then in the National Health Interview Survey (NHIS) in 2001. Now called the Census Response Achievement Field Training (CRAFT), final preparations were made in 2002-2003 to scale the program up to a production setting. In January 2004, it will be implemented nationally in the NHIS and the Survey of Income and Program Participation. Additional program areas are expected to adopt CRAFT in the future. Transitional support with the Field and Demographic Surveys Divisions is ongoing.

Research collaboration with other organizations continued in FY2003. Researchers shared research design, content, or budget advice with the Joint Program in Survey Methodology, the Research Triangle Institute, Nielson Media Research, Arbitron, the Office of National Statistics, United Kingdom, Statistics Canada, Westat Inc., and the Survey Research Center of the Institute for Social Research at the University of Michigan. Most have adopted or adapted the training design. A panel is being developed for a major conference in 2004. Also, a journal submission and final division report are planned.

Staff: Eileen O'Brien (x32695), Tom Mayer

N.4. Research Method Comparison for Computer Crime Questions

This project involves comparing three methods of pretesting used on the National Crime Victimization Survey (NCVS) computer crime questions. The three methods consist of cognitive interviews, respondent debriefing, and behavior coding.

During FY2003, all cognitive interviews were completed and have been summarized. Data from the respondent debriefing and behavior coding has also been analyzed. At this time, the results from the three methods are being compared, and analysis of this data will be used to determine if each of the three methods revealed similar or different results. Results of this research are not yet available, but will be presented at the 2003 Federal Committee on Statistical Methodology Conference.

Staff: Kristen Hughes (x38458)

N.5. Immigrants' Perception of the "Survey"

This project involves using ethnographic research methods to gain some understanding in how immigrants from four different countries (El Salvador, Ethiopia, Russia, and China) interpret and perceive the "survey" process. Specifically, we will be using both the Census short and long forms as examples of federal surveys.

During FY2003, using ethnographic research methods, a total of thirty-two in-depth interviews will be conducted with immigrants from four different countries: El Salvador, Ethiopia, Russia, and China. These interviews are being conducted to compare the different cultural understandings, socialization processes, experiences and perceptions of "the survey" that these immigrants possess. Due to problems recruiting respondents, only 20 interviews have been completed at this time and the results are not yet available.

Staff: Kristen Hughes (x38458)

Research & Development Contracts

The Research and Development Contracts are indefinite delivery, indefinite quantity task order contracts for the purpose of obtaining contractor services in highly technical areas to support research and development activities across all Census Bureau programs. The contracts provide a pool of contractors to assist the Census Bureau in conducting research on all survey and census methods and processes to improve our products and services through FY2007. The prime contractors include educational institutions, university supported firms and privately owned firms that concentrate in sample survey research, methodology and applications to create a pool of specialists/experts to tackle some of the Census Bureau's most difficult research. Many of the prime contractors are teamed with one or more organizations and/or have arrangements with outside experts/consultants to broaden their ability to meet all of the potential needs of the Census Bureau. These 5-year contracts allow Census Bureau divisions and offices to obtain outside advisory and assistance services to support their research and development efforts quickly and easily.

A. R&D 2002 Contracts

Multiple contracts were awarded in July 1997 in five technical areas: 1) technology services, 2) assessment, planning, and analysis, 3) statistical analysis, 4) methodological research, and 5) minority-focused and special populations research. These contracts expired in July 2002, and will end with completion of research on the active task orders.

During FY2003, no new task orders were awarded, 29 modifications were made to the task orders, and 25 task orders were completed. To date, 101 tasks with a value of over \$39 million have been awarded. Currently, there is one active task order. Many custom reports were prepared for various purposes, and various program managers to summarize the task orders by type of research, cost, task order, type of research, etc. Monthly progress reports were prepared and year end accruals were prepared.

Staff: Ann Dimler (x34996)

B. R&D 2007 Contracts

Multiple contracts were awarded in six technical areas: assessment, planning, and analysis; statistical analysis; methodological research; survey engineering; sub-population research; and data analysis and dissemination.

During FY2003, 18 new task orders were awarded, 13 modifications were made to the task orders, and one task order was completed. To date, 18 task orders with a value of over \$13 million have been awarded. Currently, there are 17 active tasks. Many custom reports were prepared for various purposes and various program managers, to summarize the task orders by type of research, cost, task order, type of researcher, etc. Monthly progress reports were prepared and year end accruals were prepared.

Staff: Ann Dimler (x4996), Tina Arbogast

Research Assistance

This staff provides research assistance, technical assistance, and secretarial support for the various research groups.

Staff: Tina Arbogast, Maria Cantwell, Safiya Hamid, Judi Norvell, Gloria Prout, Lorraine Randall, Nita Rasmann

3. PUBLICATIONS

3.1 JOURNAL ARTICLES, PUBLICATIONS

- Aston, J.A.D., Turkheimer, F.E., Cunningham, V.J., and Gunn, R.N. (In Press). "Wavelet Variance Components in Image Space for Spatio-Temporal Neuroimaging Data," *Proceedings of Wavelet X*.
- Bell, W.R. and Martin D.E.K. (In Press) "Computation of Asymmetric Signal Extraction Filters and Mean Square Error for ARIMA Component Models." *Journal of Time Series Analysis*.
- Bryant, Kang, Zick, and Chan, A. (2003). "His and Hers: Evaluating Husbands' Reports of Wives' Housework," *Family and Consumer Science Research Journal*, 32 (1), 8-26.
- Dougherty, M. R.P. and Hunter, J.E. (In Press). "Probability Judgment and Subadditivity: The Role of Working Memory Capacity and Constraining Retrieval," *Memory and Cognition*.
- Dougherty, M.R.P. and Hunter, J.E. (2003). "Hypotheses Generation: Probability Judgment, and Individual Differences in Working Memory Capacity," *Acta Psychologica* 111, 263-282 .
- Druin, A., Bederson, B., Weeks, A., Farber, A., Grosjean, J., Guha, M.L., Hourcade, J.P., Lee, J., Liao, S., Reuter, K., Rose, A., Takayama, Y., and Zhang, L. (2003). "The International Children's Digital Library: Description and Analysis of First Use," *First Monday*, 8 (5).
- Druin, A., Revelle, G., Bederson, B.B., Hourcade, J.P., Farber, A., Lee, J., and Campbell, D. (2003). "A Collaborative Digital Library for Children: A Descriptive Study of Children's Collaborative Behaviors and Dialogue," *Journal of Computer-Assisted Learning*, 19(2), 239-248.
- Findley, D.F. and Wills, K. (In Press). "Comments on 'Shrinkage Estimators for Damping X-12-ARIMA Seasonals,'" *Journal of Forecasting*.
- Gunn, R.N., Gunn, S.R., Turkheimer, F.E., Aston, J.A.D., and Cunningham, V.J. (2002). "Position Emission Tomography Compartmental Models: A Basis Pursuit Strategy for Kinetic Modeling," *Journal of Cerebral Blood Flow and Metabolism*, 22(12): 1425-1439.
- Hawala, S., Zayatz, L., and Rowland, S. (2003). (In Press). "American FactFinder: Disclosure Limitation for the Advanced Query System," *Journal of Official Statistics*.
- Hourcade, J.P., Bederson, B.B., Druin, A., Rose, A., Farber, A., and Takayama, Y. (2003). "The International Children's Digital Library: Viewing Digital Books Online." *Interacting with Computers*, 15, 151-167.
- Martin, D.E.K. (2003). "An Algorithm to Compute the Probability of a Run in Binary Fourth-Order Markovian Trials," *Computer and Operations Research*, 30, 577-588.
- Martin, D.E.K. (In Press). "Markovian Start-Up Demonstration Tests with Rejection of Units Upon Observing d Failures," *European Journal of Operational Research*.
- Maples, J., Murphy, S., and Axinn, W. (2002). "Two Level Proportional Hazards Models," *Biometrics* 58 (4), 180-188.
- McElroy, T. and Politis, D. (2002). "Robust Inference for the Mean in the Presence of Serial Correlation and Heavy-Tailed Distributions." *Economic Theory*, 18, 1019-1039.
- Pan, Y. (2002). "Discourse and Business Communication," *International Encyclopedia of Linguistics*, 2nd Edition. Editor, William Frawley, New York: Oxford University Press.
- Rothgeb, J., "Questionnaire Development, Evaluation, and Testing Methods (QDET) Conference: Good Discussions, Debates, and Dining in Charleston." Ongoing Research Report in *Bulletin De Methodologie Sociologique*, N. 77.
- Thibaudeau, Y. (2002). "Model Explicit Item Imputation for Demographic Categories," *Survey Methodology*, Vol. 28, No. 2, 135-143.
- Turkheimer, F.E., Aston, J.A.D., Banati, R.B., Riddell, C., and Cunningham, V.J. (2003). "A Linear Wavelet Filter for Parametric Imaging with Dynamic PET," *IEEE Transactions on Medical Imaging*, 22:289-301.
- Winkler, W. (In Press). "Methods for Evaluating and Creating Data Quality," *Information Systems*.
- Wright, T. (In Press). "U.S. Bureau of the Census," *Encyclopedia of Statistical Sciences*, 2nd Edition, (Editors: Balakrishnan, Read, and Vidakovic), New York: John Wiley & Sons, Inc.

3.2 BOOKS/BOOK CHAPTERS

- Weeks, A.C., Druin, A., Bederson, B.B., Hourcade, J.P., Rose, A., Farber, A., Reuter, K., Takayama, Y., Lee, J., White, E.J., and Kahle, B. (2003). "Creating an International Digital Library for Children," In M.A. Mardis (Ed.), *Developing Digital Libraries for K-12 Education*, 13-28, Syracuse, NY: ERIC IT Clearinghouse.

3.3 PROCEEDINGS PAPERS

American Association for Public Opinion Research, St. Pete Beach, FL, May 16-19, 2002.

2002 Proceedings of the American Statistical Association

- Hughes, K. and DeMaio, T. (2002). "Does This Question Work? Evaluating Cognitive Interview Results Using Respondent Debriefing Questions," 1535-1541.
- Leslie, T., Raglin, D., and Schwede, L. (2002). "Understanding the Effects of Interviewer Behavior on the Collection of Race Data," 2067-2068.

International Conference on Universal Access Human-Computer Interaction, Crete, Greece, June 17-July 1, 2003.

- Murphy, E.D. and Nusser, S.M. (2003). "Evaluating User Interfaces for Accommodation of Individual Differences in Spatial Abilities and Way-Finding Strategies," in C. Stephanidis (Ed.), *Universal Access in HCI: Inclusive Design in the Information Society, Vol. 4*, Proceedings of HCI International 2003, 1005-1009. Mahway, NJ: Erlbaum.

2002 Joint Statistical Meetings (American Statistical Association), New York, NY, August 11-15, 2002.

2002 Proceedings of the American Statistical Association.

- Aston, J., Cunningham, V., and Gunn, R. (2002). "Statistical Based PET Partial Volume Correction," 84-93.
- Cantwell, P. (2002). "Accounting for Imputation When Estimating Variances in the Economic Surveys at the Census Bureau," 379-384.
- Chen, B.-C., Thibaudeau, and Winkler, W. (2002). "A Comparison Study of ACS If-Then-Else, NIM, and DISCRETE Edit and Imputation Systems Using ACS Data," 461-466.
- Chen, B.-C., Thibaudeau, Y., and Winkler, W. (2002). "Item Imputation with the Discrete Edit System.
- Crowley, M. (2002). "The Political Economy of Privacy Among Special Populations: The Power Factor," 659-664.
- Findley, D., Martin, D.E.K., and Wills, K. (2002). "Generalizations of the Box-Jenkins Airline Model," 2266-2271.
- Garcia, M. (2002). "Implied Edit Generation and Error Localization for Ratio and Balancing Edits," 1122-1127.
- Huang, E. and Bell, W. (2002). "Comparing Alternative Models for Using Decennial Census Data in SAIPE State Poverty Estimates," 1494-1499.
- Schwede, L., Leslie, T., and Griffin, D. (2002). "Interviewers' Reported Behaviors in Collecting Race and Hispanic Origin Data," 3134-3139.
- Groves, R., Mayer, T., O'Brien, E., and O'Neill, G. (2002). "Interviewer Training to Increase Survey Participation," 2502-2507.
- Hawala, S. (2002). "Simulation Study of the Effectiveness of Masking Microdata with Mixtures of Multivariate Normal Distributions," 1406-1410.
- Hood, C., McDonald-Johnson, K., Monsell, B., and Nguyen, T. (2002). "Improving the Automatic RegARIMA Model Selection Procedures of X-12-ARIMA Version 0.3," 2304-2309.
- Kegan, B. and Williams, T. (2002). "Flexible Matching Imputation in the Manufactured Homes Survey," 1799-1804.
- Marquis, K. and Olmsted, E. (2002). "A Usability Evaluation of Key Parts of Census.Gov," 2518-2523.
- Massell, P. (2002). "Optimization Models and Programs for Cell Suppression in Statistical Tables," 2288-2293.
- Mulry, M. and Petroni, R. (2002). "Error Profile for PES-C as Implemented in the 2000 A.C.E.," 2436-2441.
- Slud, E. (2002). "Model-Assisted Weighting for Surveys with Multiple Response Modes," 3268-3748.
- Winkler, W. (2002). "Methods for Record Linkage and Bayesian Networks," 3743-3748.
- Yancey, W. (2002). "Improving EM Algorithm Estimates for Record Linkage Parameters," 3835-3840.

3.4 STATISTICAL RESEARCH DIVISION RESEARCH REPORTS

RR (Statistics #2002-05), William Winkler, "Methods for Record Linkage and Bayesian Networks," November 4, 2002.

RR (Statistics #2002-06), Maria Garcia and Roger Goodwin, "Developing SAS Software for Generating a Complete Set of Ratio Edits," November 6, 2002.

RR (Statistics #2002-07), William Winkler, "Using Simulated Annealing for K-Anonymity," November 19, 2002.

RR (Statistics #2002-08), William Winkler, "Single Ranking Micro-Aggregation and Re-Identification," November 22, 2002.

RR (Statistics #2002-09), Yves Thibaudeau and William Winkler, "Bayesian Networks Representations, Generalized Imputation, and Synthetic Microdata Satisfying Analytic Constraints," November 22, 2002.

RR (Statistics #2002-10), Jay Kim, Danielle Corteville, and Patrick Flanagan, "Maximizing Retention of Primary Sampling Units in a Two Primary Sampling Unit Per Stratum Design," December 4, 2002.

RR (Statistics #2003-01), Jay Kim and William Winkler, "Multiplicative Noise for Masking Continuous Data," April 17, 2003.

RR (Statistics #2003-02), David Findley and Donald E.K. Martin, "Frequency Domain Analyses of SEATS and X-11/12-ARIMA Seasonal Adjustment Filters for Short and Moderate-Length Time Series," April 25, 2003.

RR (Statistics #2003-03), Michael Ikeda and Julie Tsay, "Transparent File Construction for the State of New Jersey in Census 2000," July 30, 2003.

RR (Statistics #2003-04), Maria Garcia, "Error Localization and Implied Edit Generation for Ratio and Balancing Edits," September 2, 2003.

RR (Statistics #2003-05), Patrick J. Cantwell, Howard Hogan, and Kathleen M. Styles, "The Use of Statistical Methods in the U.S. Census: Utah v. Evans," September 4, 2003.

RR (Computing #2003-01), Bor-Chung Chen and William E. Winkler, "Preorder and Set Covering in the DISCRETE Edit System," September 10, 2003.

3.5 STATISTICAL RESEARCH DIVISION STUDIES

SS (Statistics #2002-01), Maria Garcia, Roger Goodwin, and Katherine Thompson, "The Genbounds Software for Generating a Complete Set of Ratio Edits: %Implied User's Guide," November 18, 2002.

SS (Survey Methodology #2003-01), Theresa DeMaio, Ashley Landreth, and Kristen Hughes, "Report of Cognitive Research on the School Crime Supplement for the 2001 National Crime Victimization Survey," May 9, 2003.

SS (Survey Methodology #2003-02), Kristen Hughes and Theresa DeMaio, "Final Report of Cognitive Research on the New Computer Crime Questions for the 2001 National Crime Victimization Survey," May 9, 2003.

SS (Survey Methodology #2003-03), Kristen Hughes and David Mingay, "Final Report of Cognitive Research on the Veteran's Questions for the American Community Survey," May 9, 2003.

SS (Survey Methodology #2003-04), Theresa DeMaio and Tracy Wellens, "Cognitive Evaluation of Proposed Disability Questions for the 1998 Dress Rehearsal," June 24, 2003.

SS (Survey Methodology #2002-05), Jessica Jakubowski and Theresa DeMaio, "Results of Cognitive Pretesting for the Survey of Income and Program Participation (SIPP) Work Related Expense Module," June 30, 2003.

SS (Survey Methodology #2003-06), Theresa DeMaio and Kristen Hughes, "Report on Cognitive Testing of Question to Address the Respondent Identification Policy," July 2, 2003.

SS (Survey Methodology #2003-07), Ashley Landreth, "SIPP Advance Letter Focus Group Results - Issues and Recommendations," July 3, 2003.

SS (Survey Methodology #2003-08), Theresa DeMaio and Kristen Hughes, "Report of Cognitive Research on the Residence Rules and Seasonality Questions on the American Community Survey," July 7, 2003.

SS (Survey Methodology #2003-09), Jennifer Hunter, “Results and Recommendations from the Research on Assisted Living Facilities and the Tenure Question,” July 7, 2003.

SS (Survey Methodology #2003-10), Ashley Landreth, “Results and Recommendations from Cognitive Interviews with Selected Materials Accompanying the American Community Survey,” July 9, 2003.

SS (Survey Methodology #2003-11), Ashley Landreth, “Report of Cognitive Interview Results from the Alternative Administrative Records Use Paragraph of the SIPP Advance Letter,” July 11, 2003.

SS (Survey Methodology #2003-12), Jennifer Hunter and Kristen Hughes, “Results and Recommendations from the Cognitive Pretesting of the SIPP Cash Balance Pension Question,” July 26, 2003.

SS (Survey Methodology #2003-13), Jennifer Hunter and Kristen Hughes, “Results and Recommendations from the Cognitive Pretesting of the 2003 SIPP Welfare Reform Module,” July 26, 2003.

SS (Survey Methodology #2003-14), Ashley Landreth, “Results and Recommendations from Cognitive Interviews with the Employment History Topical Module to the Survey of Income and Program Participation,” July 26, 2003.

SS (Survey Methodology #2003-15), Kristen Hughes and Jennifer Hunter, “Results and Recommendations from the Cognitive pretesting of the 2003 Reciprocity History Module,” July 26, 2003.

SS (Survey Methodology #2003-16), Jennifer Hunter and Theresa DeMaio, “Results and Recommendations from the Cognitive Pretesting of the 2003 Public School Questionnaire from the Schools and Staffing Survey,” July 26, 2003.

SS (Survey Methodology #2003-17), Lorena Carrasco, “The American Community Survey (ACS) En Espanol - Using Cognitive Interviews to Test the Functional Equivalency of Questionnaire Translations,” August 5, 2003.

3.6 OTHER REPORTS

Hunter, J., “Results and Recommendations from the Research on Assisted Living Facilities and the Tenure Question,” November 19, 2002

Landreth, A., “Final Report on Cognitive Testing of Voluntary and Mandatory Mailing Materials for the American Community Survey (ACS) Field Experiment,” December 19, 2002.

Landreth, A., “Results and Recommendations from Cognitive Interviews with the Employment History Topical Module to the Survey of Income and Program Participation,” March 21, 2003.

Hughes, K. and Hunter, J., “Results and Recommendations from the Cognitive Pretesting of the 2003 Reciprocity History Module,” March 28, 2003.

Hunter, J., and DeMaio, T., “Results and Recommendations from the Cognitive Pretesting of the 2003 Public School Questionnaire from the Schools and Staffing Survey,” April 21, 2003.

Hughes, K., “Expert Review of Proposed Changes for the 2003 Police Public Contact Survey,” April 29, 2003.

Hunter, J., Schwede, L., and Aaker, E., “Exploratory Research on Person in Custodial Care Relationship Category,” July 30, 2003.

4. TALKS AND PRESENTATIONS

Biostatistics and Brain Imaging Seminar, Columbia University, New York, NY, October 3, 2002.

- John Aston, “Statistical Based Partial Volume Correction I: The Theory,” and “Statistical Based Partial Volume Correction II: The Method.”

Special Tabulations Conference, Bureau of Labor Statistics, Washington DC, October 28, 2002.

- Laura Zayatz, “Disclosure Limitation for Census 2000 Special Tabulations.”

MD-DC-VA Chapter Meeting of the Mathematical Association of America, College Park, MD, November 2, 2002.

- Paul Massell, “Two Algorithms for Solving the Cell Suppression Problem.”

International Conference on Questionnaire Development and Evaluation Techniques Conference, Charleston, SC, November 14-17, 2002.

- Manuel de la Puente, discussant, “Methods and Issues in Developing Multilanguage, Multicultural Questionnaires.”
- Terry DeMaio and Ashley Landreth, “Cognitive Interviews: Do Different Methods Produce Different Results?”
- Sylvia Kay Fisher and Eleanor Gerber, “Issues in Translating Surveys: Methods and Approaches.”
- Barbara Forsyth, Jennifer Rothgeb, and Gordon Willis, “Does Question Pretesting Make a Difference: An Empirical Test Using a Field Survey Experiment.”
- Eleanor Gerber, presented a short course on cognitive interviewing.
- Jeff Moore, Joanne Pascale, Pat Doyle, Anna Chan, and Julia Klein Griffiths, “Using Field Experiments to Improve Instrument Design.”
- Elizabeth Murphy and Courtney Stapleton, “Evaluation of Users’ Experience of the Census 2000 Internet Form.”
- Joanne Pascale and Tom Mayer, “Alternative Methods for Exploring Confidentiality Issues Related to Dependent Interviewing.”

SAS Institute, Cary, NC, November 20, 2002.

- Brian Monsell, “An Overview of the X-12-ARIMA Seasonal Adjustment Program.”

National Academies of Science, Washington, DC, November 20, 2002.

- William E. Winkler, “Record Linkage.”

Statistics Department Seminar Series, University of Maryland, College Park, MD, November 21, 2002.

- John Aston, “Spatio-Temporal Wavelet Methods for Neuroimaging Datasets.”

Washington Statistical Society Seminar, Washington, DC, November 21, 2002.

- Laura Zayatz, discussant for “Confidentiality for a Mandatory Reporting System: Challenges and Solutions.”

Society for Judgement and Decision Making Annual Meeting, Kansas City, MO, November 21-25, 2002.

Psychonomic Society 43rd Annual Meeting, Kansas City, MO, November 21-25, 2002.

- Jenny Hunter, poster, “Effects of Time Pressure on Probability Judgments and Subadditivity.”

National Institutes of Health, Bethesda, MD, November 22, 2002.

- John Aston, “Partial Volume Correction: A Tensor Implementation.”

Combining Data from Different Sources: Applications of Record Linkage Methodology and Estimation Using Administrative Data, Sienna, Italy, January 9, 2003.

- Bill Winkler, “Record Linkage.”

Workshop Data Quality in Cooperative Information Systems 2003, Siena, Italy, January 10, 2003.

- Bill Winkler, “Methods for Evaluating and Creating Data Quality.”

Italian National Statistical Institute, Rome, Italy, January 13-17, 2003.

- Bill Winkler, Record Linkage short course.

- Bill Winkler, a series of lectures including; “Masking and Re-Identification Methods for Public-Use Microdata,” “Machine Learning Methods for Text Classification.” “Computational and Theoretical Aspects of the Expectation-Maximization and Boosting Algorithms,” “Data Quality in Cooperative Information Systems” (Panel), and “Integrated Systems.”

University of Maryland, Department of Mathematics, College Park, MD, February 20 and February 27, 2003.

- Yves Thibaudeau, “Approximations of the Posterior Variance with the Method of Laplace.”
- Yves Thibaudeau, “Approximating the Variance of the MLE of the Transition Probabilities of a Quasi-Markov Process.”

MINI STATFest: A Conference for Undergraduates, Meharry Medical College, Nashville, TN, March 1, 2003.

- Tommy Wright, “The U.S. Bureau of the Census: People, Economy, Geography, Research.”

National Academies’ Panel on Research on Future Census Methods, Seventh Plenary Meeting & Working Group Meetings, Washington, DC, March 12, 2003.

- Don Malec, “Using the ACS to Improve Net Domestic County Migration.”

Annual Meeting of American Association for Applied Linguistics, Arlington, VA, March 22-25, 2003.

- Yuling Pan, “Sociolinguistic Research and Survey Development.”

International Workshop on Comparative Survey Design and Implementation, Brussels, Belgium, March 31-April 2, 2003.

- Manuel de la Puente, “Progress Report the Language Research Initiative and Translation Guidelines at the U.S. Census Bureau.”

Biostatistics and Brain Imaging Seminar, Columbia University, New York, NY, October 3, 2002.

- John Aston, “Statistical Based Partial Volume Correction I: The Theory,” and “Statistical Based Partial Volume Correction II: The Method.”

Joint ECE/Eurostat Work Session on Statistical Data Confidentiality, Luxembourg, April 7-9, 2003.

- Laura Zayatz, “Disclosure Limitation for Census 2000 Tabular Data.”
- Yves Thibaudeau, “Bayesian Networks Representations of Contingency Tables for Reducing Disclosure and Preserving the Accuracy of Sufficient Statistics.”

Hearing Center for Enterprise Science, School of Business, University of Mississippi, Oxford, MS, April 18, 2003.

- José Dulá ,P.B. Massell, and J.T. Fagan , “Optimization Techniques in Tabular Disclosure Control.”

Seminar for Survey Methodologists, General Accounting Office, Washington, DC, April 22, 2003

- Terry DeMaio and Ashley Landreth, “Cognitive Interviews: Do Different Methodologies Produce Different Results?”

Statistics Talks, Department of Decision Sciences, University of Texas at Arlington, Arlington, TX, April 22, 2003.

- Mary Mulry, “Census 2000.”

Bureau of Transportation Statistics, Washington, DC, April 29, 2003.

- Laura Zayatz, “Overview of Statistical Disclosure for Microdata.”

DIMACS Workshop on Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications, Rutgers University, Piscataway, NJ, May 14, 2003.

- José Dulá , “Fast Algorithms for Frames and Point Depth.”

Annual Conference of the American Association for Public Opinion Research, Nashville, TN, May 15-18, 2003.

- Anna Chan and Jeffrey Moore, “Educational Attainment and Vocational/Technical Training: Questionnaire Design and Data Quality.”
- Manuel de la Puente, Discussant, “Response Rates for Mail Surveys.”
- Terry DeMaio, round table discussion, “Cognitive Interview Theory and Practices: Implications of Results from Recent Experiments.”

- Eleanor Gerber, “Respondents’ Understanding of Confidentiality Language.”
- Jennifer Hunter, poster session, “The Impact of Senior Housing Arrangements on Demographic Surveys.”
- Ashley Landreth, poster session, “Survey Letters: Message Sent, Message Received?”

Census Information Center’s Yearly Meeting, Las Vegas, NV, May 16, 2003.

- Erica Olmsted, Panel Discussion, “Census Usability Labs Initiative: Remote Usability Testing.”

Field Directors’ Conference, Nashville, TN, May 18-21, 2003.

- Larry Malakhoff, “2003 National Census Test Census Short Form IVR Development.”

National Academy of Sciences Conference on Statistics and Counter Terrorism, Washington, DC, May 29-30, 2003.

- Laura Zayatz, “Disclosure Limitation for Demographic Data.”
- Bill Winkler, “Record Linkage.”

NSF Funding Opportunity in Survey and Statistical Research, Bureau of Labor Statistics, Washington DC, June 9, 2003.

- Bill Winkler, discussant for “Multiple Imputation for Statistical Disclosure Limitation and Selective Multiple Imputation of Keys for Statistical Disclosure Control in Microdata.”

Statistical Society of Canada’s Annual Meeting, Halifax, Nova Scotia, June 10, 2003.

- Don Malec, Invited Discussant for “Recent Developments in Small Area Estimation.”

Military Personnel Research Science Workshop, University of Mississippi, Oxford, MS, June 17, 2003.

- José Dulá, N.K. Womer, and H.T. Li, “Applications of Constraint Programming: The Data Detection Problem and the DDX Manpower Planning Problem.”

Human-Computer Interaction International, 2003, Crete, Greece, June 23-27, 2003.

- Elizabeth Murphy, “Evaluating User Interfaces for Accommodation of Individual Differences in Spatial Abilities and Way-Finding Strategies.”

School of Business Administration, Virginia Commonwealth University, Richmond, VA, June 24, 2003.

- José Dulá, P.B. Massell, and J.T. Fagan, “Optimization Techniques in tabular Disclosure Control.”

Department of Mathematics and Computer Science, University of Puerto Rico, San Juan, PR, July 23, 2003.

- Yves Thibaudeau, “Imputation Methods Used by the Demographic, Decennial, and Economic Areas.”
- Tommy Wright, “Probability Sampling Concepts & Overview at the Census Bureau.”

Joint Statistical Meetings, American Statistical Association, San Francisco, CA, August 3-7, 2003.

- John Aston and Siem Jan Koopman, “A Non-Gaussian Airline Model for Seasonal Adjustment.”
- Pat Cantwell, Ryan Fescina, and Melvin McCullough, “Estimating Low-Value Exports from the U.S.”
- Adam Carle, “Measurement Invariance in Parental Reports of Child Symptomatology: Comparing Parents With and Without Psychopathology.”
- Anna Chan, “New Roster Procedures and Probes to Improve Coverage in the Survey of Income and Program Participation.”
- Jay Casselberry, Joyce Weil, Philip Steel, Gerald Gates, Kimberly McGuigan, Alvn Zarate, and Robin Smith, “Statistical Methods for Data Disclosure Limitation: An Introduction to the Committee on Privacy and Confidentiality’s Website Initiative.”
- Aref Dajani and Thomas Mathew, “Comparison of Some Tests in the One-Way ANOVA with Unequal Error Variances.”
- David Findley, John Aston, Catherine Hood, Kellie Wills, and Roxanne Feldpausch, “Diagnostics for ARIMA Model-Based Seasonal Adjustment.”
- Eric Leifer, Gregory DiRienzo, Stephen Lagakos, Eric Slud, and David Haberman, “Small-Sample Behavior of Robust Score and Wald Tests Arising from a Misspecified Cox Proportional Hazards Model.”
- Don Malec, “Small Area Estimation for Production: ‘Smoothing’ the Obstacles for Routine Use and ‘Shrinking’ the Gap with Design-Based Estimates.”
- Don Malec, “Modeling Small Areas in the Presence of Informative Cluster Sampling.”

- Donald Martin, “A Recursive Method for Computing Run Probabilities in Binary Higher-Order Markovian Sequences.”
- Paul Massell, “Comparing Cell Perturbation to Cell Suppression for Statistical Control of Tables.”
- Brian Monsell and John Aston, “Towards X-13?”
- Jeffrey Moore and Julia Griffiths, “Asset Ownership, Program Participation, and Asset and Program Income: Improving Reporting in the Survey of Income and Program Participation.”
- Mary Mulry, “Evaluation of A.C.E. Revision II Estimates of Census Coverage Error.”
- Erica Olmsted and Sarabeth Rodriguez Salibi, “Card Sorting: A Usability Method to Suggest Usable Navigation Strategies for Federal Internet and Internet Web Sites..”
- Joanne Pascale, “Improving the Health Insurance Section of the Survey of Income and Program Participation..”
- Laurie Schwede, “Rostering, Residence Rules, and Coverage: Where We’ve Been and Where We’re Going.”
- Eric Slud, “Comparison of Models for Household Response in the 1990 and 2000 Census.”
- Philip Steel, “The Feasibility of Aggregating Small Blocks for Census 2010.”
- Philip Steel, Joyce Weil, Deborah Tress, Gerald Gates, Kimberly McGuigan, Alvan Zarate, and Robin Smith, “Personal Health Information and HIPAA: An Introduction to the Committee on Privacy and Confidentiality’s Website Initiative.”
- Kellie Wills, John Aston, and David Findley, “Generalizations of the Box-Jenkins Airline Model with Frequency-Specific Seasonal Factors.”
- William Winkler, “A Contingency-Table Model for Imputing Data Satisfying Analytic Constraints.”
- William Winkler, “Current Record Linkage Practices.”
- William Winkler and Bor-Chung Chen, “Preorder and Set Covering in the DISCRETE Edit System.”
- William Yancey, “Developing and Evaluating a Learning String Comparator.”
- Laura Zayatz, “Protecting Confidentiality in Microdata Files” (Short Course).

Wavelet X, San Diego, CA, August 4, 2003.

- John Aston, “Wavelet Variance Components in Image Space for Spatio-Temporal Neuroimaging Data.”

International Workshop on Microdata, Stockholm, Sweden, August 21-22, 2003.

- Sam Hawala, “Microdata Disclosure Protection Research and Experiences at the U.S. Census Bureau.”

Department of Decision Sciences, University of Texas at Arlington, Arlington, TX, September 18, 2003.

- Mary Mulry, “Census Data for Business Use.”

NSF/NBER Time Series Analysis Conference, Chicago, IL, September 19-20, 2003.

- John Aston, Discussant, “Estimation for State-Space Models: an Approximate Likelihood Approach.”
- John Aston, “A Non-Gaussian Airline Model for Seasonal Adjustment” (Poster Session).
- Tucker McElroy, “Mean Estimation for Heavy-Tailed, Long-Range Dependent Data” (Poster Session).

5. STATISTICAL RESEARCH DIVISION SEMINAR SERIES

Seminar Series Team: Barbara Palumbo, Joanne Pascale, Yves Thibaudeau, Julie Tsay

Nathan Conklin, Virginia Polytechnic Institute and State University, Blacksburg, VA “A Web-Based, Run-Time Extensible Architecture for Interactive Visualization and Exploration of Diverse Data,” December 4, 2002.

Phil Kott, USDA/NASS/RDD, “Clarifying Some Issues in the Analysis of Survey Data,” January 14, 2003.

Herman A. Alvarado, The National Academies, “Getting Older in Puerto Rico: Are Policymakers Ready? Characteristics of the Population 65 Years and Older,” January 22, 2003.

Paul Biemer, RTI, G. Gordon Brown, RTI, Christopher Wiesen, Odum Institute, “Triple System Estimation with Erroneous Enumerations and Heterogeneity,” February 5, 2003.

James V. White, Consultant, Cambridge, Massachusetts, “Estimating Geoid Undulations from Bathymetry via Automated State Space Modeling of Multiple Time Series,” February 26, 2003.

Richard Gagnon, University of Maryland, “Time Series Analysis Using State Space Models,” March 4, 2003.

Dongping Fang, SPSS, Inc., “Automatic Model Selection and Automatic Outlier Detection for ARIMA and Transfer Function Models,” March 13, 2002.

Tucker McElroy and Dimitris Politis, University of California, San Diego, “Tall Index Estimation for Stochastically Volatile Time Series,” April 1, 2003.

José Dulá, James Fagan, and Paul Massell, Statistical Research Division, “Optimization Techniques in Tabular Disclosure Control,” April 10, 2003.

Stephen M. Belz, Eastman Kodak Company, Rochester, New York, “User Action Framework - An Emerging, Reliable Foundation for Usability Engineering,” April 15, 2003.

Juan Pablo Hourcade, University of Maryland, “The International Children’s Digital Library: Giving Children Developmentally Appropriate Interfaces to Access Online Content,” April 16, 2003.

Barbara Mirel, University of Michigan, “Field-Based Observation and Useful Support for Complex Data Questions,” April 30, 2003.

Thomas Reynolds, SRD, Census Bureau and George Mason University, “Modeling User Interaction with Mobile Computing Devices,” May 5, 2003.

Irwin Anolik, USDA, “Making Sense of Census Data via the World Wide Web: A Case Study Using the 1997 Census of Agriculture,” May 14, 2003.

Tony Hak, Erasmus University, Rotterdam, The Netherlands, “Ethnographic Methods for Understanding the Response Process in Demographic and Establishment Surveys: The Concept of Ecological Validity,” May 28, 2003.

Antonio Bruce and J. Gregory Robinson, Population Division, “The Planning Database: Its Potential Use in Current Surveys and Census 2010 Planning,” June 11, 2003.

Kent Norman, Anna Rivadeneira, and Cyntica Eaton, University of Maryland, “Design Issues in Electronic Business Surveys,” June 25, 2003.

Brian Harris-Kojetin and Karen Lee, OMB, “Insights into the OMB Clearance Process for Survey Information Collections,” June 26, 2003.

David L. Word, Charles D. Coleman, and Robert Nunziata, Census Bureau, "Putting a Demographic Face on Names from Census 2000 - An Interim Report," July 8, 2003.

Robert Fay, Census Bureau, "More on Census Duplication: Probabilistic Models for Duplication," July 15, 2003.

Craig McLaren, Australia Bureau of Statistics, "SEASABS and Time Series Analysis in the Australia Bureau of Statistics," September 2, 2003.

Paul Biemer, Jay Levinsohn, and William Wheaton, Research Triangle Institute, "New Technologies and Methodologies at Research Triangle Institute International," September 10, 2003.

6. PERSONNEL ITEMS

6.1 HONORS/AWARDS/SPECIAL RECOGNITION

Silver Medal Award, U.S. Department of Commerce

- **Laura Zayatz** - for developing and successfully implementing the disclosure limitation methods used to mask data from Census 2000 and protect the confidentiality of respondents. In addition, for testing the disclosure filter for the American FactFinder, allowing the Census Bureau to make its data available with minimal risk of disclosure; for sharing her expertise as Chair of the Disclosure Review Board; as organizer of a conference on privacy and disclosure limitation; and as co-editor of the book, *Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies* (awarded December 5, 2003).

Bronze Medal Award, Bureau of the Census

- **Kent Marquis** - for leadership and contributions that use behavioral science methods for improving data products. He founded the Census Bureau's Usability Laboratory, which focuses on human-computer interactions; established a research program to understand and reduce response error in the Survey of Income and Program Participation; and undertook the basic research and development for the Census Bureau's telephone survey capabilities.

Equal Employment Opportunity Award

- **Bill LaPlant, Lawrence Malakhoff, Kent Marquis, Betty Murphy, and Other Members of the Section 508 Implementation Team** - for development of a successful, on-going approach to address the implementation of Section 508 of the Rehabilitation Act, which requires access to Information Technology (IT) resources by the disabled. They dealt with multiple issues such as access to the Internet, Intranet, network systems, specialized accessibility software, applications software, computer equipment, establishing usability and accessibility laboratories, and access by the disabled to the IT resources for the new headquarters building. In recognition of the team's accomplishment, GSA selected this team's Section 508 acquisition process for its survey of "Best Practices" for Section 508 implementation found in the federal government.

Customer Service Award

- **Juanita Rasmann** - for expert and timely technical editing which improves the following products before their release by Census Bureau researchers in the Statistical Research Division: division reports, research reports, studies, proceedings papers, book chapters, books, and journal articles (awarded October 30, 2003).

Certificate of Recognition

- **Larry Malakhoff and Betty Murphy** - for their contribution to the success of the 2003 National Census Test.

6.2 SIGNIFICANT SERVICE TO PROFESSION

John Aston

- Reviewed a paper for *The American Statistician*.
- Reviewed a book proposal for Wiley (Statistics).

Leslie Brownrigg

- Department of Commerce Delegate, Federal Interagency Group on Migrant and Seasonal Farm Workers.
- Member, Inter Agency/Inter University Seminar on Andean Culture and Society.

Pat Cantwell

- Associate Editor, *Journal of Official Statistics*.
- Methodology Section Chair, Washington Statistical Society.
- Refereed papers for *Journal of the American Statistical Association* and *Communications in Statistics*.

Manuel de la Puente

- President, District of Columbia Sociological Society.
- Member, American Sociological Association Working Group on the Definition of Race.
- Member, American Sociological Association Nominations Committee.
- Refereed a paper for *Latino Studies Journal*.
- Invited Discussant, 2003 Meeting of American Association for Public Opinion Research.

Terry DeMaio

- Reviewed several papers for *Public Opinion Quarterly*.
- Secretary, Washington-Baltimore chapter of American Association for Public Opinion Research.

Eleanor Gerber

- Instructor, Cognitive Interviewing Course, Behavioral Science Research Center, Bureau of Labor Statistics.

Sam Hawala

- Member, National Center for Education Statistics Disclosure Review Board.
- Refereed a paper for *Survey Methodology*.
- Member, Confidentiality and Data Access Committee.

Juan Pablo Hourcade

- Paper reviews for Computer Human Interaction (CHI) 2003, Interaction Design and Children 2003.
- Papers chair, Interaction Design and Children 2004.

Jay Kim

- Refereed a paper for the *Journal of Official Statistics*.

Bill LaPlant

- Member, Advisory Board, International Center for Disability Resources.

Don Malec

- Refereed manuscripts for *Statistics in Medicine*, *Journal of the American Statistical Association*, *Journal of Official Statistics*.

Donald Martin

- Member, ASA Committee on Minorities in Statistics.
- Refereed paper for *European Journal of Operational Research*.

Paul Massell

- Member, Bureau of Transportation Statistics Disclosure Review Board.
- Refereed papers for *The American Statistician* and *Journal of Official Statistics*.
- Member, Confidentiality and Data Access Committee.

Brian Monsell

- Webmaster and *AMSTAT Online* Assistant Editor, Business and Economic Statistics Section, ASA.
- Organized session, 2003 Federal Forecasters Conference.

Jeff Moore

- Reviewed papers for the *Journal of Official Statistics* and *Public Opinion Quarterly*.

Mary Mulry

- Associate Editor, *The American Statistician*.
- Refereed a paper for *The Journal of Official Statistics*.

Eileen O'Brien

- Member, American Statistical Association's Survey Review Committee.
- Member, Interagency Response Error Group.

- Member, Interagency Household Survey Nonresponse Group.
- Teaching Assistant, Joint Program in Survey Methodology Course, Fall, 2002.

Erica Olmsted

- Co-facilitator, Workshop, “Remote Testing : Practices and Procedures,” Annual Usability Professional Association.

Yuling Pan

- Member, Interagency Response Error Group.
- Member, American Association for Applied Linguistics.
- Session Organizer, “Applying Sociolinguistics to Improving Multilingual Survey Instruments, Annual Meeting of American Association for Applied Linguistics.

Jennifer Rothgeb

- Session Organizer, “Cognitive Interviewing Theory and Practices: Implications of Results from Recent Experiments,” American Association for Public Opinion Research.
- Session Organizer, “A Sampling of the Best of QDET: Questionnaire Evaluation Methods Contributed Topic Session,” American Statistical Association meetings.
- Member, Planning Committee for QUEST 2003, Mannheim, Germany.
- Conference Chair and Organizer, International Conference on Questionnaire Development, Evaluation, and Testing Methods (QDET).
- Session Chair, “Comparison of Question Evaluation Methods,” QDET conference.
- Co-editor, QDET Monograph, *Questionnaire Testing and Evaluation Methods*.

Phil Steel

- Member, American Statistical Association’s Committee on Privacy and Confidentiality.
- Refereed a paper for *The American Statistician*.
- Member, Confidentiality and Data Access Committee.
- Chair, Washington Statistical Society Seminar on Synthetic Tabular Data.

Kellie Wills

- Member, Retention and Recruitment Committee, American Statistical Association.

Bill Winkler

- Refereed papers for *Survey Methodology*, *The Journal of the American Statistical Association*, *Information Systems*, *Journal of the Royal Statistical Society*, and *The Journal of Official Statistics*.
- Member, Organizing Committee, Work Session on Statistical Data Editing (Madrid).
- Member, Program Committee, Workshop on Data Cleaning, Record Linkage and Object Identification, Association of Computing Machinery Knowledge Discovery and Data Mining Conference.
- Organized Record Linkage Session, “Statistics for Counterterrorism Workshop,” National Academies of Sciences.
- Member, Eurostat-sponsored Data-Clean Project.
- Member, Program Committee for *Data Quality in Cooperative Information Systems*.

Tommy Wright

- Associate Editor, *The American Statistician*.
- Refereed papers for *The American Statistician*.
- Member, Editorial Board, *American Journal of Mathematical and Management Sciences*.

Laura Zayatz

- Member, Confidentiality and Data Access Committee
- Member, Interagency Team on Implementation Guidance for the Confidentiality Information Protection and Statistical Efficiency Act (CIPSEA) of 2002.
- Member, Social Security Administration Disclosure Review Board.
- Rapporteur, Technical Session, National Science Foundation Confidentiality Workshop.
- Refereed papers for *Journal of Official Statistics*.

6.3 PERSONNEL NOTES

Beverly Causey, a principal researcher, died October 17, 2002 of internal medical problems. With hometown in Alexandria, VA, he received a bachelor's degree in mathematics from Princeton University, and a Ph.D. in statistics from the University of Chicago. After several years in private industry, Bev joined the Census Bureau in 1970 and remained in the Statistical Research Division for his entire career. Bev worked independently and collaborated with others on research in mathematical statistics. The variety of his research is demonstrated by the topics on which he published: the transportation problem, optimal allocation, census adjustment, small-area estimation, estimating the number of classes in population, and many others. His research generally took a theoretical approach to solve practical problems of interest to the Census Bureau's survey programs. He published papers in the *Journal of the American Statistical Association*, the *Annals of Mathematical Statistics*, *American Demographics*, and the *Journal of Applied Statistics*, among others. Most recently, Bev had developed a model using SIPP data for estimating the length of time individuals participated in various government programs.

RETIREMENTS from the Census Bureau: Sam Highsmith

Tom Reynolds, a graduate student in the Human Factors and Cognition Program at George Mason University, joined the usability group as an academic year intern. He left the Census Bureau to accept a position in private industry.

Erica Olmsted earned her masters degree in Technical Communication from the University of Central Florida.

Lynn Weidman returned to the Census Bureau to lead our Sampling Research Group.

Jerry Maples joined our Small Area Estimation Research Group.

Adam Carl, a Census Bureau Postdoctoral Researcher, joined the division as a member of the Questionnaire Design and Measurement Research - I Group.

Mohammed Chaudhry joined our division as a member of the Computer Support Staff.

Richard Gagnon, a graduate student at the University of Maryland in applied mathematics, joined the Time Series Research Group.

Summer Visitors:

Joint Program in Survey Methodology Junior Fellows:

- Elizabeth Aaker (junior in psychology at Mount Holyoke College) joined our Questionnaire Pretesting for Household Surveys Group.
- Muhammad F. Bhuiyan (senior in economics, mathematics, and mathematical methods for social sciences at Northwestern University) joined our Disclosure Limitation Research Group.
- Michael Furchtgott (junior in economics and mathematics/statistics at Columbia University) joined our Time Series Research Group.
- Katie Gagne (junior in economics at Colby College) joined our Questionnaire Design & Measurement Research (1) Group.
- Lucas Perkins (senior in political science at Reed College) joined our Questionnaire Pretesting for Household Surveys Group.

Andrew Perez-Lopez (graduate student in computer science at Virginia Tech University) joined our Human Factors and Usability Research Group.

Michael Rosen (graduate student in English-technical writing at University of Central Florida) joined our Human Factor and Usability Research Group.

Jay Kim accepted a position at the National Center for Health Statistics.

José Dulá's temporary appointment ended, and he returned to the University of Mississippi.

Tom Mayer left the Census Bureau for personal pursuits.

Juan Pablo Hourcade joined our division as a member of the Human Factors and Usability Research Group.

Tucker McElroy joined the Time Series Research Group.

APPENDIX A

**Statistical Research Division's FY 2003 Program Sponsored Projects/Subprojects
With Substantial Activity and Progress and Sponsor Feedback
(Basis for PERFORMANCE MEASURES)**

Project #	Project/Subproject Sponsor(s)	SRD Contact	Sponsor Contact
	DECENNIAL		
5010	Special Purpose Tests		
5201	Content and Forms Development		
5202	Questionnaire and Public Use Forms Design and Development		
5604	Special Populations		
	1. <i>Alternative Questionnaire Experiment (AQE2000)</i>	Eleanor Gerber . . .	Randall Neugebauer
	2. <i>2010 Census Residence Rule Working Group</i>	Laurie Schwede	Ed Byerly
	3. <i>2010 Coverage Planning Research and Development Group</i>	Laurie Schwede	Frank Vitrano
	4. <i>Decennial Privacy Research</i>	Tom Mayer	Gerald Gates
	5. <i>Decennial Questionnaire Design and Pretesting</i>	Terry DeMaio	Jane Ingold
	6. <i>Response Options Strategies Working Group</i>	Betty Murphy	Suzanne Fratino
	7. <i>Census 2000 Modified Race Data Summary File Accessibility Improvement</i>	Larry Malakhoff	William Savino
6121	Statistical Requirements		
	8. <i>Decennial Coverage Research</i>	Don Malec	Rick Griffin
	9. <i>Research on Accuracy of the Census and the A.C.E. Estimates</i>	Mary Mulry	Rita Petroni
	10. <i>2010 Census Coverage Measurement Research</i>	Mary Mulry	Donna Kostanich
	11. <i>A.C.E. Missing Data Research and Development</i>	Mike Ikeda	Michael Beaghen
	12. <i>Administrative Records Linkage Support</i>	Ned Porter	Robert Rothhaas
	13. <i>Decennial Disclosure Limitation Research</i>	Laura Zayatz	Raj Singh
6367	14. <i>Census 2000 Evaluation Projects</i>	Manuel de la Puente . .	Debbie Bolton
	DEMOGRAPHIC		
1461	15. <i>SIPP METHODS PANEL</i>	Jeff Moore	Pat Doyle
1465	Survey of Income and Program Participation (SIPP) Research		
	16. <i>Continuous Instrument Improvement Group (CIIG)</i>	Jeff Moore	Bob Kominski
	17. <i>Longitudinal Weighting</i>	Leroy Bailey	Tracy Mattingly
	Quick Turnaround Pretesting of Household Surveys		
	18. <i>ACS Voluntary Experiment</i>	Terry DeMaio	Debbie Griffin
	19. <i>SIPP Welfare Reform Module</i>	Terry DeMaio	Chuck Nelson
	20. <i>Research on Assisted Living</i>	Terry DeMaio	Len Norry
	21. <i>SASS Public School Questionnaire</i>	Terry DeMaio	Steve Tourkin
	22. <i>SIPP Reciprocity History Module</i>	Terry DeMaio	Pat Doyle
	23. <i>SIPP Employment History Module</i>	Terry DeMaio	Pat Doyle
	24. <i>Group Quarters Change to Relationship Category</i>	Terry DeMaio	Annetta Smith
	25. <i>SASS Public School Teacher Questionnaire</i>	Beth Nichols	Steve Tourkin
TBA	26. <i>Evaluation of the National Long Term Care Survey</i>	Jennifer Rothgeb .	Denise Quesenberry
5385	Continuous Measurement/American Community Survey (Now Decennial)		
	27. <i>ACS Language Research: Focus Groups with FRs and SFRs and Cognitive Testing of the CAPI/CATI Spanish Language Instrument</i> . . .	Manuel de la Puente . . .	Debbie Griffin
	28. <i>ACS Small Area Estimation Research</i>	Don Malec	Freddie Navarro
	29. <i>ACS-Edit/Imputation Research</i>	Bill Winkler	Barbara Diskin
	30. <i>Nonresponse Weighting Adjustments for ACS Application</i>	Eric Slud	Sue Love
	31. <i>ACS Weighting Simplification Research</i>	Lynn Weidman	Freddie Navarro
	32. <i>Program of Integrated Estimates</i>	Lynn Weidman	Signe Wetrogan
TBA	33. <i>Research on Item Imputation for the 2001 RFS</i>	Yves Thibaudeau	Howard Savage
7165	34. <i>Research for Small Area Income and Poverty Estimates (SAIPE)</i>	Elizabeth Huang	Dave Waddington
TBA	35. <i>Analysis and Forecasting of Demographic Time Series</i>	Kellie Wills	Fred Hollman

Project #	Project/Subproject Sponsor(s)	SRD Contact	Sponsor Contact
	ECONOMIC		
3420051	36. <i>Disclosure Limitation Methods</i>	Laura Zayatz	Carol Blatt
3420052	Time Series Research		
	37. <i>X-12-ARIMA Development and Evaluation</i>	Brian Monsell	Catherine Hood
	38. <i>Research on Seasonal Time Series - Modeling and Adjustment Issues</i> .	Don Martin/Kellie Wills .	Catherine Hood
3320054	39. <i>Development of Software for Economic Programs</i>	Maria Garcia ...	Katherine Thompson
	METHODOLOGY AND STANDARDS		
7676	40. <i>Statistical Consulting/Postal Rate Commission</i>	Leroy Bailey	Robert Cohen
7566091	41. <i>National Park Service Grant Program</i>	Carol Corby	Joel Lynch
8863	42. <i>National Institute of Standards & Technology - Bayesian Statistical Methodology</i>	Don Malec	Nell Sedransk
Other	Usability		
	43. <i>Quality Management Repository</i>	Erica Olmsted	Deb Stempowski
	44. <i>Automated Listing and Mapping Instrument</i>	Kent Marquis ...	Mary Ann M. Chapin
	45. <i>Administrative and Customer Services Intranet Site Development & Testing</i>	Erica Olmsted	Susan Boyer
	46. <i>Field Directorate Intranet Site Usability Testing</i>	Erica Olmsted	Steve Angel/ Dennis Van Longen
	47. <i>Housing and Household Economic Statistics Information Architecture</i>	Erica Olmsted	Lori Gido
	48. <i>Congressional Web Site</i>	Erica Olmsted	Marian Brady
	49. <i>American Community Survey Usability Study</i>	Erica Olmsted	Margaret Gill
	50. <i>Economic Census Web Usability Project - 2003</i>	Michael Rosen	Paul Zeisset

APPENDIX B

**FY 2003 PROJECT PERFORMANCE MEASUREMENT QUESTIONNAIRE
STATISTICAL RESEARCH DIVISION
Methodology and Standards Directorate**

Dear

In a continuing effort to obtain and document feedback from program area sponsors of our projects or subprojects, the Statistical Research Division will again attempt to provide *seven measures of performance*, this time for the fiscal year 2003. For FY 2003, the *measures of performance* for our division are:

- Measure 1. Overall, Work Met Expectations*
Percent of FY 2003 Program Sponsored Projects/Subprojects where sponsors reported that work met their expectations.
- Measure 2. Established Major Deadlines Met*
Percent of FY 2003 Program Sponsored Projects/Subprojects where sponsors reported that all established major deadlines were met.
- Measure 3a. At Least One Improved Method, Techniques Developed, Solution, or New Insight*
Percent of FY 2003 Program Sponsored Projects/Subprojects reporting at least one improved method, techniques developed, solution, or new insight.
- Measure 3b. Plans for Implementation*
Of the FY 2003 Program Sponsored Projects/Subprojects reporting at least one improved method, techniques developed, solution, or new insight, the percent with plans for implementation.
- Measure 4. Predict Cost Efficiencies*
Number of FY 2003 Program Sponsored Projects/Subprojects reporting at least one "predicted cost efficiency."
- Measure 5. Journal Articles, Publications*
Number of journal articles (peer review) and publications documenting research that appeared or were accepted in FY 2003.
- Measure 6. Proceedings Publications*
Number of proceedings publications documenting research that appeared in FY 2003.

These measures will be based on input from our sponsors as well as from members of our division. We will use these measures and associated detail to help improve our efforts. This action is consistent with the spirit of the *Government Performance Results Act (GPRA) of 1993* ". . . to provide for the establishment of strategic planning and performance measurement in the Federal Government."

To construct these seven measures for our division, we will combine the information for each of our program area sponsored projects or subprojects obtained during October 6-20, 2003 using this questionnaire. As indicated on this questionnaire, much of the information will be provided by researchers in the Statistical Research Division. Your assistance is requested for the remaining information on

Project Number and Name _____
Sponsoring Division(s) _____

After all information has been provided, the SRD Contact _____ will ensure that the signatures are obtained in the order indicated on the last page of this questionnaire.

We very much appreciate your assistance in this undertaking.

Tommy Wright, Chief
Statistical Research Division
Date

Brief Project Description (SRD Contact will provide from Division's Quarterly Report):

Brief Description of Results/Products from FY 2003 (SRD Contact will provide):

TIMELINESS: Established Major Deadlines/Schedules Met

1(a) Were all established major deadlines associated with this project or subproject met? **(Sponsor Contact)**

- Yes No No Established Major Deadlines

1(b) If the response to 1(a) is No, please suggest how future schedules can be better maintained for this project or subproject. **(Sponsor Contact)**

Comments: _____

**QUALITY & PRODUCTIVITY/RELEVANCY:
Improved Methods /Techniques Developed/Solutions/New Insights**

2. Listed below are at most 3 of the top improved methods, techniques developed, solutions, or new insights offered or applied on this project or subproject in FY 2003 where an SRD staff member was a significant contributor. Review the list, **(provided by SRD Contact)** and make any additions or deletions as necessary. For each, please indicate whether or not there are plans for implementation. If there are no plans for implementation, please comment. Add any comments, and certify with your initials.

No improved methods/techniques/solutions/new insights developed or applied.

Plans for Implementation?

- | | | |
|----------|------------------------------|-----------------------------|
| a. _____ | Yes <input type="checkbox"/> | No <input type="checkbox"/> |
| b. _____ | Yes <input type="checkbox"/> | No <input type="checkbox"/> |
| c. _____ | Yes <input type="checkbox"/> | No <input type="checkbox"/> |

Comments (**Sponsor Contact**): _____

Related reports, software/hardware, professional publications (e.g., in the proceedings of professional/scientific organizations, through inter-agency publications, etc.) or peer-reviewed publications by SRD staff that appeared during FY 2003 are listed. An abstract or summary for each listed document is attached.

No reports, software/hardware, professional publications, or peer-reviewed publications appeared during FY 2003.

- a. _____
- b. _____
- c. _____
- d. _____
- e. _____
- f. _____
- g. _____

COST: Predict Cost Efficiencies

3. Listed (**provided by SRD Contact**) below are all research results or products produced for this project or subproject in FY 2003 that predict cost efficiencies. Review the list, and make any additions or deletions as necessary. Add any comments.

No cost efficiencies predicted.

- a. _____

- b. _____

Comments (**Sponsor Contact**): _____

OVERALL: Expectations Met/Improving Future Communications

4. Overall, work on this project or subproject by SRD staff during FY 2003 met expectations. (**Sponsor Contact**)

Strongly Agree Agree Disagree Strongly Disagree

If you checked "disagree" or "strongly disagree," please comment (**Sponsor Contact**): _____

5. Please provide suggestions for future improved communications or any area needing attention on this project or subproject.

Suggestions (**Sponsor Contact**) _____

(SRD Contact will coordinate first two signatures as noted and pass to SRD Chief.)

First _____ Second _____
Sponsor Contact Signature Date SRD Contact Signature Date

(SRD Chief will coordinate last two signatures as noted.)

Third _____ Fourth _____
Sponsor Division Chief Signature Date SRD Division Chief Signature Date

Statistical Research Division

Assistant Division Chief for Computing and Technology

Robert Creecy
Barbara Palumbo

Computer Support Staff

Chad Russell
Mohammed Chaudhry
Tom Petkunas

Statistical Computing Research

Bill Winkler
Bor-Chung Chen
Maria Garcia
Judi Norvell
William Yancey

Computing Applications

Aref Dajani
Ned Porter
Mary Ann Scaggs
VACANT

Computing Research

Yves Thibaudeau
Carol Corby
Bill LaPlant
Nita Rasman
Todd Williams
VACANT
VACANT

Assistant Division Chief for Mathematical Statistics

Pat Cantwell
Alice Bell

Sampling Research

Lynn Weidman
Ann Dimler
Mike Ikeda
Mary Mulry
Gloria Prout
Julie Tsay
VACANT

Small Area Estimation Research

Don Malec
Elizabeth Huang
Jerry Maples

Statistical Estimation & Analysis Research

Leroy Bailey
Tina Arbogast
Pam Ferrari
Ruben Mera
Eric Slud (U of MD)
VACANT

Disclosure Limitation Research

Laura Zayatz
Sam Hawala
Paul Massell
Phil Steel
VACANT

Time Series Research

Brian Monsell (Acting)
John Aston (NISS Post Doc)
Richard Gagnon (S)
Donald Martin (Howard U)
Tucker McElroy
Kwon Moon (V)
Thomas Trimbur (Post Doc)
Kellie Wills

Assistant Division Chief for Survey Methodology

Center for Survey Methods Research

Manuel de la Puente
Maria Cantwell

Questionnaire Design & Measurement Research -1

Jeff Moore
Adam Carle (Post Doc)
Anna Chan
Julia Klein-Griffiths
Beth Nichols
Joanne Pascale
Jennifer Rothgeb

Questionnaire Design & Measurement Research -2

Eleanor Gerber
Melinda Crowley
Safiya Hamid
Eileen O'Brien
Yuling Pan
Laurie Schwede
VACANT

Questionnaire Pretesting for Household Surveys

Terry DeMaio
Kristen Hughes
Jennifer Hunter
Ashley Landreth
Lorraine Randall

Human Factors and Usability Research

Kent Marquis
Leslie Brownrigg
Joyce Farmer
Juan Pablo Hourcade
Larry Malakhoff
Aaron Miller (S)
Betty Murphy
Erica Olmsted

Office of the Chief

Tommy Wright
Hazel Beaton