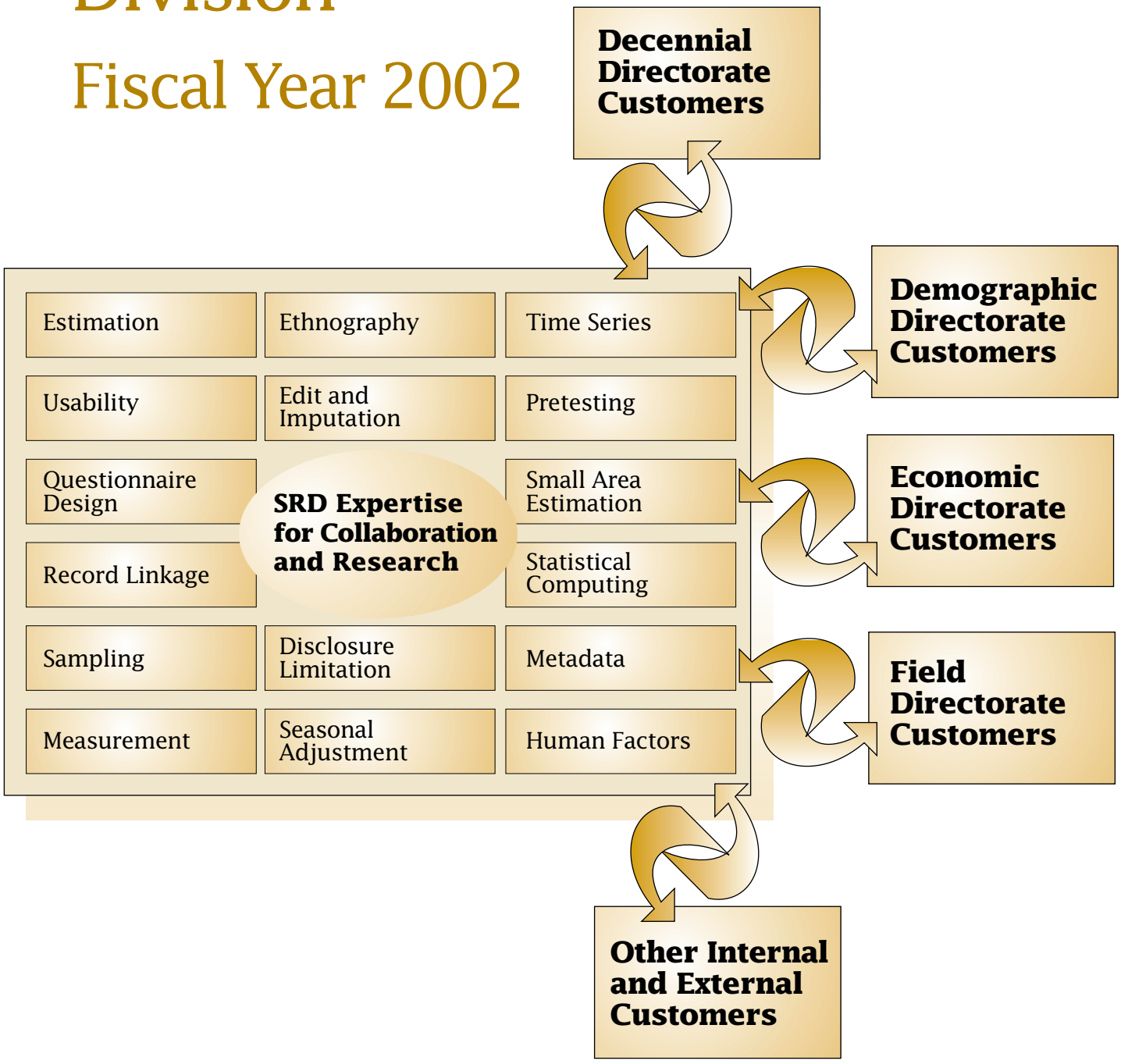
Methodology and Standards Directorate

Annual Report *of the* Statistical Research Division

Fiscal Year 2002



U S C E N S U S B U R E A U

Helping You Make Informed Decisions • 1902-2002

U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU

We help the Census Bureau improve its processes and products. For fiscal year 2002, 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 2002 follow, and more details are provided within subsequent pages of this report:

- tested and evaluated the disclosure limitation techniques for the Advanced Query System; tested the new 3D cell suppression and auditing program to determine feasibility for use in the 2002 Economic Census.
- completed Release 0.2.10 of X-12-ARIMA; investigated an approach to improving survey estimates by using time series models with a sampling error component and its application to a large set of "Construction" series with high sampling variability.
- developed a model and estimation procedure that uses American Community Survey immigration to extend the Population Division's county immigration estimation procedures; nearly completed comparisons between an exact and approximate small area estimation method that accounts for unknown within-small-area sampling variances.
- constructed a research (transparent) file for New Jersey in Census 2000.
- analyzed 1992-1993 panel data from the Survey of Income and Program Participation (SIPP) relating to the effects of attritional bias on selected survey items.
- developed a new set covering algorithm for implicit-edit generation for discrete data and a method of applying them in the integer programs used for error localization (finding the fields to change in an edit-failing record); demonstrated noticeable improvements with the application of the Flexible Matching Imputation procedure to the Manufactured Homes Survey.
- opened our accessibility evaluation facility; collaborated on the design or usability evaluation of at least 10 Census Bureau electronic products.
- conducted or coordinated over twenty-five pretesting activities across the decennial, demographic, and economic areas.
- provided advisory services to the Systems Support Division (SSD) Corporate Metadata Repository project with particular focus on applications required to perform the 2002 Economic Census.
- reported results from a literature review that explored a number of issues regarding privacy concerns and attitudes, confidentiality beliefs, and their relationship to the functions of the U.S. Census Bureau.
- developed new methods for using small (0.01%) amounts of training data to improve record linkage parameters and error rate estimates.
- provided question wording, content, and design recommendations to the National Cancer Institute in preparation for the final draft of the Current Population Survey Tobacco Use Supplement.
- designed and initiated a round of cognitive interview testing of the new SIPP Wave 2 interview focusing on dependent interviewing procedures and the "RIP" (respondent identification policy) question; results suggest that the new procedures work well and do not elicit any respondent concerns.

How Did We Do..

For a fourth year, we received feedback from our sponsors. Near the end of fiscal year 2002, our efforts on fifty-six of our program (Decennial, Demographic, Economic, CASRO) 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-six questionnaires provided the following results:

*Measure 1. Overall, Work Met Expectation*s

Percent of FY2002 Program Sponsored Projects/Subprojects where sponsors reported that overall work met their expectations (agree or strongly agree) (51 out of 56) 91%

Measure 2. Established Major Deadlines Met

Percent of FY2002 Program Sponsored Projects/Subprojects where sponsors reported that all established major deadlines were met (36 out of 45 responses) 80%

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

Percent of FY2002 Program Sponsored Projects/Subprojects reporting at least one improved method, techniques developed, solution, or new insight (49 out of 56 responses) 88%

Measure 3b. Plans for Implementation

Of these FY2002 Program Sponsored Projects/Subprojects reporting at least one improved method, techniques developed, solution, or new insight, the percent with plans for implementation (42 out of 49 responses) 86%

Measure 4. Predict Cost Efficiencies

Number of FY2002 Program Sponsored Projects/Subprojects reporting at least one "predicted cost efficiency" 18

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 (6) or were accepted (5) in FY2002 11

Measure 6. Proceedings, Publications

Number of proceedings publications documenting research that appeared in FY2002 . . . 34

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

Statistical Research Division

TABLE OF CONTENTS

1. COLLABORATION

1.1 - 1.5 DECENNIAL TOPICS (Decennial Projects)

A. 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 FY2002, staff continued to develop and apply methodology and software for Transparent File construction for the state of New Jersey for Census 2000. A preliminary result is that the overall difference between the count from the Accuracy and Coverage Evaluation and the count from development of the Transparent File was one person for the entire state. Results from subpopulations are still being reviewed. The memorandum "Transparent File Construction for the State of New Jersey in Census 2000" is in preparation and outlines the methodology and results.

Staff: Julie Tsay (x4915), Michael Ikeda

B. Decennial Edit/Imputation Research

(For FY2002 progress, see Demographic Project 4200-H.)

Staff: Bill Winkler (x4729), Bor-Chung Chen, Yves Thibaudeau, Todd Williams

C. 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 FY2002, staff completed sensitivity analysis of the synthetic comparisons using artificial population analysis with staff in DSSD. Results were disseminated in a report and used as part of the decision making process on whether to recommend adjusting Census 2000 or not. (See Griffin & Malec (2001) under *Publications/Other Reports*). The artificial populations were used to estimate the additional effects the synthetic assumption could have on the numerical loss function used to decide for, or against, adjustment. While in most scenarios the addition of synthetic error has no effect on the loss function decision, there were a few cases that would change the decision, illustrating that the error from the synthetic assumption is not trivial.

Short term assistance to DSSD has been completed. Staff continues research on small area unit-level models at the Local Census Office level for census adjustment.

A simple hierarchical model incorporating census capture, A.C.E. capture and A.C.E. erroneous enumeration has been developed to assess the degree of variability of coverage. Programming for estimation is underway.

Staff: Donald Malec (x4892), Michael Ikeda

D. Alternative Questionnaire Experiment (AQE2000)

The objectives of AQE2000 are to continue efforts to develop a user-friendly mail package that can be accurately 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 FY2002, analysis was completed for the three components of the Alternate Questionnaire Experiment. Final versions of the reports were submitted for the branching instructions component of the experiment and for the race and Hispanic origin components of the experiment. The residence rules component of the experiment, which presented a new reader-friendly format for the residence rules and improved coverage of low mail return strata, was circulated in draft and revised according to comments received.

Staff: Eleanor Gerber (x4890), Manuel de la Puente, Aref Dajani, Mary Ann Scaggs, Yves Thibaudeau

E. 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 FY2002,the RM&IE report, prepared by the contractor Westat, was submitted to the Planning, Research, and Evaluation Division for final approval. We found that the Computer-Assisted Telephone Interview brought about a small but statistically significant improvement in the overall response rate. The calling card was very effective in promoting the use of the alternative response mode. However, the impact of the calling card may not justify its cost.

Staff: Larry Malakhoff (x3688)

F. 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 FY2002, staff provided information to the Decennial Statistical Studies Division (DSSD) on assorted issues related to A.C.E. missing data. Staff also provided consulting support to the Planning, Research, and Evaluation Division on assorted issues related to the evaluation of A.C.E. missing data procedures for Census 2000. Staff produced the final version of the memorandum, "Accuracy and Coverage Evaluation Survey: Some Notes Related to Accuracy and Coverage Evaluation Missing Data Procedures." The paper, "Handling Missing Data in the 2000 Accuracy and Coverage Evaluation Survey" has been submitted to *Survey Methodology* and presents the extent of missing data in the survey, describes the methods applied, comparing them when appropriate to alternatives used in the past, and provides analytical summaries of the methods. Because the resulting levels of missing data were low, it appears that alternative procedures would not have affected the results substantially.

Staff produced a draft specification for the noninterview adjustment in the Revised A.C.E. Staff modified the production noninterview adjustment program for use in the Revised A.C.E. and ran the modified program on preliminary Revised A .C.E. data files. Staff proposed and described procedures for imputing a value of missing age for Revised A.C.E. that is consistent with the imputed age category from the production A.C.E. Staff also provided general support to DSSD on handling missing data in the Revised A.C.E.

Staff: Mike Ikeda (x4864), Pat Cantwell

G. Privacy Implications of the Decennial Census

The goal of this project is to conduct a qualitative study of belief structures that influence survey respondents' perceptions of and reactions to survey information requests, focusing on privacy concerns. A team of ethnographers will use a combination of observation, interview debriefing, and semi-structured cognitive interviews to explore how respondents assess the consequences of survey participation and survey response, their sense of information ownership, their reactions to confidentiality statements, and their reasons for choosing to participate in survey data collection activities.

During FY2002, results from this research indicated that there is a wide common area of agreement among respondents in the way they make decisions about requests for information. Our major findings, such as the concern with sponsorship, relevance, and risks vs. benefits, can be found in every group we interviewed. All are concerned with the possibility of fraud, and give high priority to the protection of financial resources. Suspicions of the security of the Internet occurs in all groups, regardless of the degree of experience with the mode or technical expertise with computers. These results were included in a draft report circulated for comment. Additional analysis of this data provided insight into respondents' ideas of the role of respondents in the survey process, and the situation of the survey process within a particular cultural matrix. Results of these additional analyses were presented at conferences in the United States and Germany.

Staff: Eleanor Gerber (x4890), Melinda Crowley

H. Complex Households and Relationships in the Decennial Census and in Ethnographic Studies

This ethnographic research project has three primary objectives: 1) to explore the range and functioning of complex households within different ethnic groups in the United States; 2) to examine how the response categories of the decennial relationship question capture the emerging diversity of household types; and 3) to compare the household composition and relationship information collected by the ethnographic interviews to those collected in Census 2000. This study is designed to assess how well census methods, questions, relationship categories, and household composition typologies describe the emerging diversity of household types in the United States. Seven contract ethnographers conducted small-scale ethnographic interviewing subprojects of African American, Inuipiaq Eskimo, Navajo, and white, as well as Hispanic and Korean immigrant complex households.

During FY2002, staff received a commitment that AltaMira Press would publish our proposed ethnographic book on complex households and cultural conceptions of "households" and "families" among the six race/ethnic groups. Among other things, this research documented that complex households can affect the reliability of some of the Census Bureau's distributions of family and non-family households (e.g., unmarried partner households with children of one of the partners only may be classified as family households or non-family households depending on

whether or not the parent is Person 1) and that relationship categories are not mutually exclusive (e.g., grandchildren who are adopted may be listed as "adopted" rather than "grandchild"). Staff also participated in helping the 2010 Within Household Coverage Working Group identify root causes of coverage errors.

Staff: Laurie Schwede (x2611), Anna Chan

I. Generation X Speaks Out on Censuses, Surveys, and Civic Engagement: An Ethnographic Approach

The purpose of this nationwide ethnographic research is to examine civic engagement behaviors and attitudes towards censuses and surveys among gen-Xers (individuals born during the years 1968-1979) from varied socio-economic backgrounds and ethnicities, including individuals from hard-to-enumerate categories (such as young, minority males). Patterns of civic engagement have consequences for government data collection efforts in terms of survey nonresponse, trust and privacy concerns, policy-oriented issues and effective educational outreach campaigns. Baby Boomers and Millennial Generation individuals will also be interviewed in order that comparative life-cycle experiences and cultural explanations emerge in regard to census and survey nonresponse, government engagement, and civic responsibility and obligation.

During FY2002, a comprehensive report, "Generation X Speaks Out on Civic Engagement and the Census," was written, submitted to the Planning, Research, and Evaluation Division, and is currently under Census 2000 Testing and Experimentation Program review. Findings from this research will be presented at the 2002 American Evaluation Association Conference in Washington, DC as an invited panel paper in November 2002. This research suggests that the following factors may contribute to decennial noncompliance and undercoverage errors: non-citizenship status or unstable immigration status, respondents not knowing about or understanding the decennial census, and increased levels of distrust among respondents toward the government.

Staff: Melinda Crowley (x2726), Eleanor Gerber, Yves Thibaudeau

J. Enumeration Barriers Specific to Colonias in the Context of Census 2000

The objective of this ethnographic research is to examine potential barriers to census enumeration in selected colonias along the US/Mexico border in the context of Census 2000. This objective will be accomplished through: a) participant observation; b) ethnographic interviews; c) focus groups with colonia residents, census enumerators and crew leaders; and d) in-depth interviews with colonia residents. The field work will be conducted by professional ethnographers from outside the Census Bureau working under purchase orders and staff from our division and the Planning, Research, and Evaluation Division (PRED). Specific issues to be examined in this project include: irregular housing; knowledge of English and literacy; concerns regarding confidentiality of census data; and complex and fluid household structure. Products from this research will include ethnographic reports authored by ethnographers from outside the Census Bureau, and a final report, authored by staff from our division and PRED, summarizing all relevant findings. This report will also make recommendations for improving the census enumeration of colonias.

During FY2002, staff worked on the final PRED report, which recommended the use of "cultural facilitators" in all colonias, provided suggestions for improving enumerator and crew leader training, and observed that almost all colonia residents did in fact, want to participate in Census 2000. This report is expected to be final during the second quarter of FY2003.

Staff: Manuel de la Puente (x4997)

K. Comparative Ethnographic Research on Mobile Populations

This project conducts and reviews research and technology to recommend or develop applications, methods, outreach, and communications appropriate for population groups defined in categories other than "race" or ethnicity which are identified as requiring special methods for listing, enumeration, and/or enumeration support. Examples include people in migratory or seasonal occupations, communities of language other than English such as recent immigrants, institutions, and group quarters (GQ), or rural remote or other areas difficult to enumerate by conventional methods and those who require accommodations or tailored approaches.

In FY2002, a summary report was composed of four ethnographic studies of mobile populations that had been conducted in the months surrounding Census 2000. Each researcher had used a unique method to study his/her particular group that varied, depending on the nature of the socio-cultural setting. Across

these four distinct groups (i.e., gang members, Irish Travelers, Arizona Snowbirds, and American Indians), the summary presents the following common barriers to enumeration: residential mobility; distrust and/or fear of outsiders or the government; irregular and complex household arrangements; and disinterest. These barriers have been mentioned in previous studies of hard-to-enumerate groups. We demonstrated, using novel ethnographic methods, that these problems are real and consistent across mobile groups. Through this research, we discovered that involvement of community organizations is critical in gaining trust from these groups; non-traditional site enumeration was helpful in Census 2000, but should be expanded in Census 2010; and transients and their families are unsure how to apply the residence rules to a person who has more than one usual residence. The final report summarizes recommendations made by ethnographers for Census 2010 enumeration strategies.

Staff: Jennifer Hunter (x4927), Matt Salo

L. Ethnographic Social Network Tracing in the Context of Census 2000

This evaluation project applies ethnographic field and analysis methods to trace the domiciles of people interacting in social networks over a six month period overlapping Census Day 2000. (*Decennial Evaluation J-2*)

During FY2002, staff edited, formatted, and distributed final reports from the six ethnographers and completed contract administration for the J2 (Ethnographic Social Network Tracing) and J3 (Mobile Peoples) projects. By sustaining contact with a small number of participants over a six month period, ethnographers were able to document in great depth, the greater context of community social networks and migration circuits. The findings suggest areas for improving our lists of units of enumeration such as: short-term rural rental housing of labor camps, workers' group quarters, informal boarding houses, and transient quarters (hotels and motels) where migrant workers typically stay. Staff responded to requests for information, materials, and advice from the National Marine Fisheries Services (NMFS), which is considering developing a new survey of U.S. fishermen inspired in part by the Memorandum of Understanding ethnographic research between the Census Bureau and NMFS in this project.

Staff also wrote, formatted, and completed the final J2 report, and circulated it for review.

Staff: Leslie Brownrigg (x4995)

M. Evaluation of the Decennial Frame of Group Quarters and Sources

This project evaluates the coverage, content, comparability, and sources of information used to construct the Decennial List of Special Places/Group Quarters (SP/GQ) through linkage and matching with the contemporary Business Register (BR) and examination of contributing sources. The SP/GQ list is built separately and with different sources and methods from the Master Address File list of housing units. (*Decennial Evaluation E4*)

During FY2002, staff submitted three drafts of the final report for E4 and presented "Lessons Learned" before a joint meeting of members of three Group Quarters working groups and Decennial Management Division (DMD) staff.

Reviewers requested a separate report limited to questions, methods, and results of the match between the final Census 2000 SP/GQ records and the Business Register. Relevant text was culled from the longer report to format into the newly requested product. Among findings, the automated match between the decennial area's list of Group Quarters and economic area's Business Register documents the equivalence between certain Group Quarters type codes and certain economic Standard Industrial Codes. The results suggest a potential future role for use of the Business Register in the construction of the decennial's Group Quarters Frame. Matching work continued as well as service on the Frame Development Working Group.

Staff: Leslie Brownrigg (x4995), Ned Porter

N. Administrative Records Linkage Support

The purpose of this project is to assist in the application of record linkage software to a variety of projects in the Decennial, Demographic, and Economic Directorates.

During FY2002, staff joined the Special Places and Group Quarters Unduplication Review Board and started an unduplication project which will require new techniques and software for the matcher. Staff developed one such technique with the address standardizer and made some improvements to the matcher reporting systems, incorporated an Internet page into the SRD web portal and added documentation of Record Linkage Software to the SRD web portal. Staff researched and began to learn XML (extensible markup language - the Web's database language of Choice). Staff taught part of a Record Linkage Course to Census Bureau employees. Web-based applications for Record Linkage appear at http://recordlinkage.srd.census.gov/.

Staff: Ned Porter (x4729)

P. Decennial Privacy Research

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

In FY2002, staff continued to work with the Privacy Policy and Research Committee (PPRC) to conduct research to assess public opinion on privacy-related issues and to assist with the design of privacy related policies. We participated in meetings of the PPRC and participated in subgroups regarding the construction of Census Bureau Privacy Principles developing a Privacy Confidentiality Program Initiative. We chaired a PPRC subgroup charged with closing the gaps between the Privacy Principle of Informed Consent and the current Census Bureau policies and procedures, which has led to an analysis and possible redesign of Census Bureau informed consent materials. We completed a review of the literature regarding privacy and confidentiality issues as they relate to the Census Bureau and released it as an official SRD Research Report which was subsequently used as a key background document for the creation of a business case for the research component of the Privacy and Confidentiality Program Initiative. We participated in an interchange between Statistics Canada and the Census Bureau and made a presentation on cognitive testing of privacy and confidentiality statements at the Census Bureau.

Staff: Tom Mayer (x4930), Jeff Moore

Q. Coverage, Rostering Methods, and Household Composition: A Comparative Study of Current Population Survey and the Census 2000

The Statistical Research Division proposed evaluating the coverage, rostering methods and household composition differences between the Census 2000 and the Current Population Survey (CPS). The study which was originally developed in FY2001, was given the evaluation number I.6., and was titled, "Coverage, Rostering Methods, and Household Composition: A Comparative Study of the Current Population Survey and the Census 2000."

To conduct this evaluation, staff was to use the matched CPS to Census 2000 dataset. Staff worked during FY2001 and part of FY2002 on the preparations for the computer match, clerical match, and proposed field follow-up. Budget cuts eliminated approximately 50 or so Census 2000 evaluations, one of which was evaluation I.6. Prior to these cuts, staff drafted a survey for the field follow-up portion and conducted cognitive interviews on the draft. Staff then finalized a report, "CPS-and NHIS-Census Match Field Follow-up Study: Questionnaire Development, Cognitive Testing Results and Recommendations for Coverage Study Research." This report documents the results of the cognitive interviews and the entire match process. Using information gleaned from the cognitive interviews, the report speculated as to the lack-of-success expected by reinterviewing household over two years later than the original data collection. The report also suggests coverage questionnaire research for 2010 planning.

Staff: Beth Nichols (x4865), Laurie Schwede

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

This project examines the quality of Census 2000 and two sets of dual system estimates of census coverage error using data collected in the A.C.E. Survey. The two sets of estimates of census coverage error are the A.C.E. estimates published in March 2001 and the Revised A.C.E. estimates with a completion date in December 2002. The assessment of the quality includes analyses of the biases and random errors in the dual system estimates arising from data collection, data processing, missing data, identification of census duplicate enumerations, and model bias. The investigation includes a synthesis of the nonsampling and sampling errors and a comparison of the relative accuracy of the census and the revised A.C.E. estimates. The analyses add refinements as new evaluation data become available, and are tailored for the specific application.

During FY2002, staff developed the design for an evaluation program for the Revised A.C.E. estimates of coverage error in Census 2000 and plans for its implementation. The implementation began in FY 2002 and will be completed in FY2003. The planning included preparing overviews of 12 individual projects and a draft summary of all projects for a chapter in the documentation of Revised A.C.E. Plans. New methodology in the evaluation includes the use of administrative records in an evaluation of the accuracy of duplicate census enumerations identified by computer. The 12 projects combine to assess the relative quality of the census and the estimates from the Revised A.C.E. Additional contributions were the designs of several of the individual evaluation projects for the Revised A.C.E. estimates. Organizing the

evaluation program included writing a Statement of Work for a contractor funded by the Planning, Research, and Evaluation Division (PRED) to participate in the evaluation of the Revised A.C.E. Preparation for the designing of the evaluation program included reviewing the evaluations of the 2000 A.C.E. estimates and creating an error profile. Also provided was technical advice to PRED and outside researchers on an analysis of erroneous enumerations in the census using latent class models.

Staff: Mary Mulry (x8462)

S. 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 FY2002, the contractor submitted a final version of the report to the Census Bureau. Two hundred and fifty telephone interviews were behavior coded using two different versions of the race question. The purpose of the coding was to look for evidence of possible problems in interviewer behavior or respondent performance. Problems proved to be infrequent: over 90% of the interviewer behavior was coded as appropriately, following instructions. However, interviewers did have some difficulty with skip patterns within the race questions, probably caused by the necessity of paper and pencil, rather than CATI administration. This research supports the accuracy of the data collected by the Census Quality Survey.

Staff: Jennifer Rothgeb (x4968), Eleanor Gerber, Betsy Martin (M&S)

T. Legal Case Brought by Utah on Count Imputation

The purpose of this project is to assist with the preparation of declarations, briefs, and arguments for the second lawsuit brought by the state of Utah against the Department of Commerce which focused on Count Imputation in Census 2000.

Staff conducted research and prepared and reviewed materials to be used in the lawsuit whereby Utah claimed that the Census Bureau's use of count imputation in Census 2000 was illegal. Staff produced documentation of past practices and uses of imputation, and developed arguments to help explain the differing roles that imputation and sampling have in a data collection activity such as a census or sample survey. On June 20, 2002, the U.S. Supreme Court ruled that the Census Bureau's use of hot deck imputation methodology, which added over one million persons to the Census 2000 count, was permitted in producing the final census count.

Staff: Pat Cantwell (x4982)

U. 2010 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 FY2002, we worked with colleagues from other divisions to develop alternative presentations of the Census residence rules for the 2004 Census Test sites to improve coverage. Division staff reviewed the literature on alternative questionnaire designs used in successive Census Bureau tests of residence rules and coverage conducted from 1994 to 2000 and prepared a report on the findings. Based on that review, we recommended that the existing plans for cognitive testing of person-count formats be expanded to include resumed testing of roster-based formats. Colleagues in the Decennial Management Division agreed to this recommendation and added funding for additional cognitive testing. We participated in subgroups to design five questionnaires and reviewed four successive iterations of the five questionnaire versions. We participated in another subgroup to plan the cognitive testing to be done by an outside contractor. We took the lead in writing the proposal for cognitive testing by an outside contractor and worked with staff to select the contractor and finalize the contract. We participated in the kickoff meeting with the contractor and reviewed the contractor's draft protocol and screening instruments. We suggested developing a new method of rolling coverage follow-up as part of the 2004 Census Test.

Staff: Laurie Schwede (x2611), Eleanor Gerber, Ann Dimler

1.6 CURRENT POPULATION SURVEY (CPS) (Demographic Project 0906)

A. CPS March Supplement Research

Design, conduct, and analyze exploratory

cognitive and other research to investigate measurement problems in the CPS March supplement, and recommend questionnaire design solutions. Recently, this research has focused on questions measuring participation in welfare reform benefits.

No work on this project was requested of our division during this fiscal year.

Staff: Jennifer Rothgeb (x4968), Lorraine Randall

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

Develop the questionnaire content of the Tobacco Use Supplement (TUS) to the 2003 Current Population Survey (CPS) 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 FY2002, staff provided question wording, content, and design recommendations to the National Cancer Institute in preparation for the final draft of the Current Population Survey (CPS) Tobacco Use Supplement (TUS). Staff produced two data set files from behavior coding analysis and provided significant findings to the sponsor, including guidance on placement of some questions. For example, by placing the menthol/non-menthol question before the cigarette-brand strength question, we found that this helped respondents avoid confusing these characteristics when reporting on their cigarette strength. Also, revised and simplified wording of some questions resulted in improved understanding by all respondents. In preparation for the 2003 CPS TUS, all cognitive pretesting, pre-field research, and behavior coding analysis was completed. The final behavior coding report is in progress. Administration of the supplement will be conducted during February 2003.

Staff: Melinda Crowley (x2726), Kristen Hughes, Aref Dajani, Susie Resper-McFadden, Michelle Jiles

1.7 SIPP 2000 METHODS PANEL (Demographic Project 1461)

The SIPP Methods Panel (MP) 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 staff carry out the research activities necessary to implement the recommendations of the Continuous Instrument Improvement Group (CIIG).

During FY2002, staff designed and initiated a round of cognitive interview testing of the new Wave 2 interview focusing on dependent interviewing procedures and the "RIP" (respondent identification policy) question; results suggest that the new procedures work well and do not elicit any respondent concerns. Staff attended two debriefing sessions held with MP2001 Wave 2 Field Representatives (FRs). In general, FRs' evaluations of the Wave 2 instrument were highly favorable, and with few exceptions, they much preferred the new instrument design to the old. We summarized the main results from all research activities through the MP2001 field test, and prepared background documents for subject-area staffs in the Housing and Household Economic Statistics and Population Divisions concerning final recommendations for Wave 1 instrument design for the final MP2002 test. With their concurrence, we presented those final, section-by section recommendations to the SIPP Executive Staff, which approved them for the MP2002 test. Staff completed detailed design work for MP2002 instrument sections that are largely new in the test instrument (i.e., the labor force section, which is being re-formatted back to a more SIPP-like design; and the health insurance section, whose general design plans have long been in place but which has not yet been implemented in the first two field tests), and for new components of existing sections (i.e., a new series on non-English language use). We analyzed data from the MP2001 test, focusing particularly on seam bias and other Wave 2+ issues; initial results show evidence of seam bias improvement with the new procedures, as well as continued improvement in income item nonresponse. We implemented Wave 1 of the MP2002 test, completing 1182 test (MP instrument) and 1181 control (standard SIPP) interviews; for the first time in the field test series, the test instrument response rate (87.8%) was actually slightly higher than the control (86.8%). We conducted a debriefing session for

MP2002 Wave 1 FRs, who continued to report highly favorable attitudes toward almost all of the new procedures.

Staff: Jeff Moore (x4917), Anna Chan, Julia Klein-Griffiths, Joanne Pascale

1.8 SURVEY OF INCOME AND PROGRAM PARTICIPATION 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 FY2002, we conducted research on advance letters for SIPP. We conducted research for the administrative record use paragraph in the advance letters for SIPP Methods Panel administrative records experiment and prepared a report of research results. Staff also completed and distributed a research report documenting cognitive research on proposed changes to the SIPP advance letter. Recommendations were incorporated into the production SIPP advance letter and the MPSIPP experimental letters.

We began finalizing the questionnaire for, and making preparations to conduct cognitive research on the SIPP Welfare Reform module.

Staff: Terry DeMaio (x4894), Kristen Hughes, Jenny Hunter, Ashley Landreth, Lorraine Randall

B. Continuous Instrument Improvement Group (CIIG)

The 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 FY2002, with the assistance of subject-area experts from the Housing and Household Economic Statistics and Population Divisions, and from the Demographic Surveys Division's SIPP production and processing branches, CIIG reviewed and recommended improvements to the following SIPP topical modules: Recipiency History, Work Disability History, Education and Training History, and Marital History. We also initiated a review of the Fertility History module. A review of the recommended revisions to the Employment History topical module by division staff led to the inescapable conclusion that cognitive pretesting with paper-and-pencil prototypes of the redesigned modules will not be feasible, but will instead require the development of automated versions of the modules.

Staff: Jeff Moore (x4975), Julia Klein-Griffiths, Joanne Pascale, Anna Chan

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 FY2002, staff completed an analysis of the effectiveness of the current longitudinal nonresponse weighting procedure based on the 1992-93 panel results. We also developed general strategies for assessing the effects of nonignorable attrition on the principal survey items and for producing empirical estimates of the likelihood of program participation for selected survey items.

Staff adapted a procedure designed to compare different overall longitudinal weighting schemes and to determine the effects and interactions of specific weight modifications on SIPP estimates. In addition, the staff developed a theoretical framework for the derivation and further testing of model-based alternatives to the current weighting cells adjustment procedure for longitudinal nonresponse.

Staff: Leroy Bailey (x4917), Pam Ferrari, Todd Williams

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 FY2002, a paper summarizing the major results of this research was completed and reviewed. It provided additional insight into alternatives for variance estimation, spell length estimation and analysis, and longitudinal analysis applicable to the SIPP design. A detailed proposal based on the principal recommendations resulting from this

research is being developed for presentation to the SIPP Research and Evaluation Committee.

Staff: Ruben Mera (x4934), Leroy Bailey

E. Spell Length Analysis

This project seeks to address the accurate estimation of such statistics as multiple occurrences of spells of program participation for analytical units, right censoring of spells due to limited observation periods, and dependencies among analytical units. The work will extend the analysis to include the use of weights and variance estimation for the model parameters.

During FY2002, a revised paper, "Maximum-Likelihood Estimation for Spell Durations Under Censoring" was reviewed internally and submitted for consideration for journal publication. The procedure described in the paper is designed to permit spell length estimation for doubly censored spells, and it compensated for nonstationarity, which posed problems in previous reports.

Staff: Bev Causey

F. Quick Turnaround Pretesting for Household Surveys

(See Projects 0351 and 1871, Survey Methodology-D)

1.9 SURVEY OF PROGRAM DYNAMICS (Demographic Project 1467)

The purpose of this research is to test questions proposed for the Survey of Program Dynamics (SPD), a survey that provides panel data to evaluate welfare and health care reform, especially as they influence income, program participation, employment, and child well-being.

During FY2002, staff provided consultation through attendance at monthly steering committee meetings.

Staff: Jennifer Rothgeb (x4968)

1.10 AMERICAN HOUSING SURVEY (Demographic Project 7455)

Since 1985, the Census Bureau has been involved in efforts to explain and reduce the disparity between various income estimates derived from the American Housing Survey (AHS) and those produced by the Current Population Survey (CPS) and the Survey of Income and Program Participation (SIPP). The AHS has historically reported a higher percentage of households below the poverty level than the CPS; moreover, the median household income reported by the AHS has generally been lower than that reported by the CPS or SIPP. This research is intended to provide a systematic analysis of the problem, which is expected to result in the identification of possible causes of the discrepancies and the nature and extent of their effects, and, if appropriate, the derivation of plausible compensation procedures.

During FY2002, staff proceeded with a more extensive analysis of the discrepancies between AHS and CPS income estimates, and identified what is ostensibly the primary cause of the disparity. A proportion of the AHS respondents report having zero non-wage income while they have positive non-wage income. Those who report positive values seem to be accurate. An appropriate adjustment methodology was derived and is being tested for consistency. A more definitive analysis of the problem and final recommendations will be contingent upon the availability of more detailed data files and subject matter specialists to assist us in understanding the intricacies of the data collection and processing procedures that resulted in the respective income estimates. A final report has been written and the project is completed.

Staff: Ruben Mera (x4934), Pam Ferrari, Leroy Bailey

1.11 2000 SAMPLE REDESIGN RESEARCH (Demographic Project 4000)

The Demographic Statistical Methods Division (DSMD) has been redesigning samples based on new Census 2000 data. The Statistical Research Division has been assisting DSMD in its redesign activities. The activities include: maximizing PSU overlap between 1990 and 2000 redesigns (Work Group 3.4) and stratification of Primary Sampling Units (PSUs).

During FY2002, staff performed a thorough review of the PSU Selection Methodology Plan. We observed errors and helped DSMD correct them.

Staff verified test run data for the Survey of Income and Program Participation (SIPP), the Current Population Survey (CPS), and the National Crime Victimization Survey (NCVS) for selecting PSUs within the strata. In many strata for which SUNSET Software did not run, we used CPLEX to find reasons for the infeasibility and helped DSMD resolve the problems. We ran CPLEX using the dual procedure and performed sensitivity analyses for a stratum. We compared SUNSET Software solutions to CPLEX solutions. With our help, DSMD picked the 2000 sample PSUs for CPS, SIPP, and NCVS.

Staff: Jay Kim (x4907)

1.12 CONTINUOUS MEASUREMENT/ AMERICAN COMMUNITY SURVEY (ACS) (Demographic Project 4200)

A. Questionnaire Design and Measurement: Race and Ethnicity Research

ACS staff have hypothesized that interviewer behavior may offer one explanation as to why more Hispanics reported their race as "White" in the nation-wide administration of the American Community Survey (Census 2000 Supplementary Survey) than in Census 2000. Staff will use a combination of qualitative and quantitative methods to learn more about how ACS Field Representatives (FRs) report they handle a variety of potential responses to the Hispanic origin and race questions.

During FY2002, staff developed iterative versions of the "Race and Hispanic Origin Questionnaire for American Community Survey CAPI Field Representatives and Supervisory Field Representatives" and finalized it after a review by some field representatives. We implemented it and integrated qualitative data from debriefings and questionnaire write-in comments with quantitative data in our analysis. Our findings suggested that the handling of the race questions by FRs may be correlated with regional office affiliation, that FRs differ on what they think the race question is asking for, and that FRs differ on how they reply to respondents confused by the race question. These findings point to the need for revised instructions and training for the FRs.

Staff: Laurie Schwede (x2611), Eleanor Gerber

B. Questionnaire Design and Measurement: Border Communities (Colonia) Research

This study uses ethnographic methods and other qualitative techniques to better understand unique barriers that may exist for the American Community Survey in communities along the US/Mexico border. The study uses focus groups and in-depth interviews to determine and understand the perceptions of community residents regarding various American Community Survey materials and methods. This study will also conduct a qualitative assessment of potential housing unit coverage problems.

During FY2002, the contractor delivered a draft report and staff provided comments. The second draft was submitted from the contractor and the staff is reviewing int. The final report is expected by the second quarter of FY2003.

Staff: Manuel de la Puente (x4997)

C. Questionnaire Design and Measurement: Translating Demographic Surveys

This effort has two major goals. The first goal is to convene a panel of expert translators in order to develop best practices and guidelines for the translation of Census Bureau demographic surveys from English into Spanish. The second goal of this effort is to use these best practices and guidelines to develop a conceptually and linguistically equivalent Spanish language translation of English language ACS forms. A long term goal is to use this effort as a prototype to do translations of Census Bureau demographic surveys from English into other selected languages.

During FY2002, staff prepared materials for an expert panel meeting. The expert panel was convened, and a summary of the proceedings was produced but not finalized. A final report on this panel meeting is expected during the second quarter of FY2003. Staff developed and delivered presentations on the development of guidelines and best practices of the Census Bureau at the Center for Survey and Methodology (ZUMA) in Mannheim, Germany.

Staff: Manuel de la Puente (x4997), Yuling Pan, Eleanor Gerber

D. 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 FRs and SFRs who interview persons whose primary language is Spanish.

During FY2002, staff and the contractor worked collaboratively and developed and pretested a

cognitive interview protocol, and conducted over 40 Spanish language cognitive interviews using the Spanish language ACS CAPI instrument in the Washington DC, Chicago, and Los Angeles metropolitan areas, and El Paso, Texas. A draft report was produced. A final report is expected during the second quarter of FY2003.

Staff also worked collaboratively with a contractor and developed a focus group protocol, and conducted 20 focus groups with ACS field representatives and supervisory field representatives who interview Spanish speaking households using the Spanish and English language ACS instruments. These focus groups took place in Tucson, AZ, Houston, TX, San Antonio, TX, Miami, FL, and the Los Angeles metropolitan area. Staff and contractor produced preliminary reports of findings. A draft report was produced and circulated for comments. Extensive comments were received. These comments will be incorporated, and a final report will be issued early in the third quarter of FY2003.

Recommendations based on findings include: provide standard procedures to FRs and SFRs for the conduct of ACS interviews with Spanish speaking households; provide specific training to FRs and SFRs for conducting ACS interviews with Spanish speaking households; and conduct additional cognitive testing using the ACS Spanish language CATI/CAPI instrument in order to make it more useful to the FRs and SFRs and more understandable to Spanish speaking respondents.

Staff: Manuel de la Puente (x4997), Lorena Carrasco (DMD), Alfredo Calvillo (NCHS)

E. ACS Language Research: Assessment of Data Quality Obtained from Linguistically Isolated Households

Very little research has been conducted on the quality of data obtained by Census Bureau surveys from linguistically isolated households. The objective of this effort is to assess the quality of survey data obtained from these households. Data quality measures or indicators will be developed and used to compare and assess the quality of survey data obtained from linguistically isolated households with households that are not linguistically isolated. Mode comparisons as well as comparisons between English language and Spanish language survey instruments will be conducted. This research will be conducted by a team of researchers selected to participate in the Census Bureau's Mid-Career Enrichment Program. This is a new program open to selected permanent Census Bureau employees. Employees selected for the program will be detailed to the sponsoring division for 6 to 12 months.

During FY2002, staff prepared a proposal for the research which was submitted to the Human Resources Division and used for the Mid-Career Enrichment Program. Staff reviewed responses to the proposal from individuals who applied, and one individual was selected. Staff reviewed the proposal and began the research.

Staff: Manuel de la Puente (x4997), Pamela McGovern (DIR), Debbie Griffin (DSMD)

F. American Community Survey New York City Special Study (Effects of September 11)

The New York City Planning Department requested the Secretary of Commerce to conduct a special study in New York City to learn about the job-related and transportation-related effects of the September 11 attack. The Demographic Surveys Division (DSD) requested our division to test the proposed questions.

During FY2002, staff conducted an expert review of the proposed questionnaire and suggested revisions. Using a revised questionnaire, staff developed a cognitive interviewing protocol and conducted cognitive interviewers with persons in New York City. Testing results, recommended revisions, and final decisions were provided in a report to DSD.

Staff: Jennifer Rothgeb (x4968)

G. 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.

As part of this goal, and during FY2002, staff developed a statistical model of the Population Division's (POP) domestic county migration cohort component method which is based on matching IRS households for two consecutive years. This model has been extended to incorporate ACS data for estimating migration of individuals who were not in households who filed taxes in two consecutive years. Work had begun with staff from POP and the Demographic Statistical Methods Division to use the model to derive estimates and evaluate the impact of these new estimates on estimates of county migration. A

complete draft of the model and estimation was completed. However, work has slowed down due to higher priority responsibilities by both divisions, and difficulties with obtaining permission to access IRS data through POP division.

In addition, staff began developing a project to model county level 1 differences between current and usual residence. The aim of this project is to develop better ACS benchmarks.

Staff: Donald Malec (x4892)

H. 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 FY2002, staff continued to work on the comparison study of ACS If-Then-Else, NIM, and DISCRETE edit and imputation systems. Staff completed a draft of the paper of comparing ACS if-then-else, NIM, and DISCRETE. Staff also completed the paper, "An Efficient Formulation of Age Comparisons in the DISCRETE Edit System." Staff had ASA papers, "Comparison Study of ACS If-Then-Else, NIM, and DISCRETE Edit and Imputation Systems Using ACS Data" and "Item Imputation with the Discrete Edit System." If several systems are compared, then the comparison can show discrepancies between the edited/imputed outputs. In some situations, we found that the discrepancies suggest that the ACS If-Then-Else can be improved. While we found no major issues, there were some minor ones.

Staff: Bill Winkler (x4729), Bor-Chung Chen, Yves Thibaudeau, Todd Williams

I. Nonresponse Weighting Adjustments for ACS Application

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

Staff prepared, for the Demographic Statistical Methods Division and ACS application, a set of formulas for mean-squared error of survey estimates.

Staff applied mean-squared error formulas, previously developed for survey estimates corrected with model-based weighting adjustments to 1990 Census data. The resulting MSEs were used to assess the desirability of such model-based corrections. These were contrasted with results when misspecified models were used in weighting adjustment.

Results were presented at the ASA meetings and written into a proceedings paper. This year's progress on this project suggests model-based weighting adjustment is feasible and desirable but has risks. The next step, for future work, is to test the methodology on ACS data.

Staff: Eric Slud (x4991)

1.13 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 or small demographic domains (e.g., proportion of children age 5 to 17 in poverty 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 G under Projects 0351 and 1871, General Research, Statistical Methodology).

During FY2002, staff evaluated alternative ways of using the census data in state poverty ratio models paying particular attention to results for income year (IYs) 1989 and 1999. [(Note that the 1990 (2000) census poverty estimates were for IY1989 (1999).] The results showed that models using census data only (omitting the administrative records predictor variables) were preferred for census years (IY 1989 and 1999). Staff also evaluated the prediction error variance estimates of the model-based state poverty ratio estimates from the current model which uses the "census residuals" and the administrative records variables, and a reduced model which uses census data only for census income years (IYs 1989 and 1999). The results showed that the estimated prediction error variances of the state poverty ratio estimates from the reduced model were lower than those from the current model for all 51 states (including DC as a "state" for

estimation purposes) for each of 4 age groups (0-4, 5-17, 18-64, 65+). These results were presented at the 2002 Joint Statistical Meetings and are documented in a paper titled, "Comparing Alternative Models for Using Decennial Census Data in SAIPE State Poverty Estimates." In considering what to do for IY 1999 production state estimates, it was realized that, having done the statistical testing involved in making these comparisons, to quote prediction error variances from the reduced model as if the reduced model were known to be true, would understate statistical uncertainty in the estimates. As a compromise, the full model (using both census and administrative records predictors) was used, but with an informative prior distribution for the regression coefficients of the administrative records predictors. This produced prediction error variances somewhat lower than those for the full model with a noninformative prior distribution, but that were not nearly as low as those for the reduced model.

Staff produced the estimated sampling variances and covariances of the direct Current Population Survey (CPS) state poverty ratio and median income estimates for IYs 1995-2000 (extending to include IY 1999 and IY 2000 with the results for IYs 1995-1998 that were produced last year). For IYs 1999 and 2000, CPS estimates based on Census 2000 controls were used. These variance and covariance estimates provided the input data used to develop the CPS sampling error model whose results fed into the model for the CPS state estimates. Staff also developed a program in SAS to test potential models for producing income and poverty estimates at the state and county levels for SAIPE. The model uses a bivariate normal structure that is more flexible than allowed for by other programs SAIPE has developed, and so will be useful for testing alternative models. This program could be used to develop models at the School District level.

Staff: Elizabeth Huang (x4915), David Smith, William Bell (M&S)

1.14 LATENT CLASS MODELS (LCM)
(Demographic Project TBA)

The goal of this project is to increase staff understanding of latent class methods, and their applicability to Census Bureau projects, especially the modeling of response error. Simulation studies may be performed with the goal of determining how sensitive the results of latent class modeling are (e.g., estimation of error rates) to violations in the assumptions about data commonly made in latent class models. If these simulation studies point toward the usefulness of latent class models to certain Census Bureau projects, training in latent class models may be given. Associated training in specialized software may also be given. (This work is also supported by Projects 0351 and 1871.)

During FY2002, staff assisted a researcher in the Quality Assurance and Evaluation Branch (QAEB) who was testing latent class modeling software developed at a university and used widely in the field. The goal was to determine if this software was computing correctly all the quantities needed to analyze the types of latent class models used in QAEB. Related to this work, staff helped improve understanding of the algorithms underlying LCM software so that QAEB could develop its own specialized software.

Staff: Paul Massell (x4954)

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

The purpose of this project is to produce specifications for statistically tested item imputations for the 2001 Residential Finance Survey (RFS).

During FY2002, and in coordination with the Housing and Household Economic Statistics Division and the Demographic Surveys Division, staff designed and produced an imputation matrix for the item imputation process for the RFS. Staff tested the statistical validity of the imputation formulae and the nearest neighbor hot-deck schemes referenced in the matrix with proven statistical techniques, such as general linear models, logistic regression, and discrete multivariate analysis. The formulae and hot-deck schemes are implemented in the imputation/edit process for the RFS.

Staff: Yves Thibaudeau (x4729), Richard Levy (HHES)

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

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

This is new work with the Population Division (POP). Staff met with POP staff to discuss directions for the project. Issues discussed included implications of race detail and incorporation of parity specific (birth order) fertility rate calculations.

Staff: Kellie Wills (x1425), William Bell (M&S)

1.17 DISCLOSURE LIMITATION METHODS (Economic Project 3320051)

The purpose of this research is to develop disclosure limitation methods to be used for Census Bureau publicly available economic 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 cell suppression for selected Economic surveys. Methods will be developed, tested, evaluated, and documented. We will also aid in the implementation of the methods.

During FY2002, staff worked with staff from Economic Statistical Methods and Programming and Economic Planning and Coordination Divisions to assess the state of disclosure software to be used for the 2002 Economic Censuses. Based on testing, it appears that the linear programming (LP) based suppression programs would not be fast enough to handle many of the large 3D tables. Only 3D tables need be considered for treatment by LP based programs. The older, network- based programs are suitable for 1D and 2D tables.

Staff developed multistage cell suppression software that may be applied to large three-dimensional tables such as those that are included in the County Business Patterns set of tables. This software involves running programs on two platforms (DEC-VMS and SUN-UNIX) and the transfer of data using FTP (file transfer protocol). A new program (called CAIF) was developed on the DEC-VMS to create an input file for the 3D audit program. The CAIF program will produce a file for the audit program which is more compact than the audit input file that is currently produced by the suppression program. For the County Business Patterns, the audit file is very large, so a reduction in its size is important for both disk space and run time considerations. A variant of the 3D LP-audit programs on SUN-UNIX was developed to identify the subset of the primary suppressions in the master data file that did not receive adequate protection during the initial suppression stage (from the network flow-based cell suppression program on DEC-VMS). Finally, a variant of the 3D LP-based suppression program was developed to find adequate protection for the identified set of primaries. Integration of the components will begin soon; it will be followed by testing of the whole process.

Staff sent documentation to the Energy Information Administration that describes the Census Bureau's successful testing and implementation of the interagency (Federal Committee on Statistical Methodology) Disclosure Auditing System (DAS), specifically the version for a UNIX platform that runs on a recent (8.1 or later) version of SAS. This documentation was included in a readme.txt file that was distributed with the final version of DAS.

Staff documented the application of the p% rule to survey data and the use of a newly-developed, short, user-friendly program for cell suppression that can be used by the economic directorate on single (i.e., unlinked) small two-dimensional tables. The Manufacturing and Construction Division used this program for cell suppression on tables of data on housing starts and completions in New York City.

Staff: Laura Zayatz (x4955), Paul Massell, Phil Steel, Sam Hawala

1.18 TIME SERIES RESEARCH (Economic Project 3320052)

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 Statistics Directorate.

During FY2002, X-12-ARIMA or general seasonal adjustment support was provided to Centers for Disease Control, Bureau of Labor Statistics, Federal Reserve Board, Eurostat, Statistics Netherlands, INEGI (Mexico), INDEC, OECD, Statistics Austria, Statistics Canada, Statistics New Zealand, Statistics Norway, Secretariat of the Pacific Community, Australian Bureau of Statistics, Canadian Department of Finance, Niigata Prefectural Office, Washington State Employment Office, Virginia Department of Employment and Training, City of Virginia Beach Management Services, the Office of Management and Budget of the District of Columbia, Korea Railroad Research Institute, Bank of England, Bank of Japan, Bank of Mexico, Bank of South Africa, Central Bank of Chile, International Monetary Fund, Federal Reserve Bank in Dallas, Hitotsubashi University, The University of Montpelier (France), Universidad Nacional de Tucuman, Southern Methodist University, University of Illinois at Carbondale, University of Indiana, Tongji University, Korea University, City University of New York, University of Dresden, University of Austral (Argentina), George Washington University, University of Wollongong, Universidad Nacional de Colombia, Institute of Economics (Taiwan), the Swiss Institute for Business Cycle Research, BCA Research, Pitcairn Trust, AHSA Associates, CreditRe Corporation, Frito-Lay, Toyota, Vamberk Chains Ltd., SPSS, DIR-WEFA London, dc Soft, Standard and Poors, Bank of America,

DGBASEY, and SAS Institute.

Staff: Brian Monsell (x2985), 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 or external training in the use of X-12-ARIMA and the associated programs, such as X-12-Graph, when appropriate.

During FY2002, versions 0.2.9 and 0.2.10 of X-12-ARIMA were made available on the Census Bureau website for the PC, Unix, and Windows XP. Improvements implemented in these versions of the X-12-ARIMA include a new procedure for generating the default outlier critical value, an option to reweight trading-day coefficients generated by the irregular regression procedure to ensure negative daily weights are not used to generate trading-day factors (requested by Statistics Canada), a chi square statistic for groups of related regressors and specific types of user defined regressors, new runtime options to reduce program output, an option to rescale the original series by a power of ten (requested by Statistics Canada), extensive improvements for procedural and code consistency, and many bugs fixes. Staff created a utility program to convert X-11-ARIMA/2000 input files to X-12-ARIMA input specification files to assist analysts at Statistics Canada in converting their seasonal adjustment production runs to X-12-ARIMA.

The automatic model selection procedure in Version 0.3 of X-12-ARIMA was revised to utilize an AICC test to determine if trading-day or Easter regressors should be included in a default time series model, rather than t-tests of the regressors as in TRAMO. An additional AICC test was added to the procedure to check if trading-day or Easter regressors should still be included in the model if it is different than the default model. Staff also improved error handling and checking in the automatic model selection procedure from Version 0.3 after testing from other staff members and staff from TSMS.

Courses on seasonal adjustment for statisticians and economists were taught by staff at the Bank of Thailand, the Sri Lankan Department of Census and Surveys, and the Bank of England. The final course was co-taught with staff from TSMS.

Staff: Brian Monsell (x2985), Kellie Wills, Thuy Trang Ta Nguyen (MCD)

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 included 1) collaboration with the Time Series Methods Staff of the Economic Statistical Methods and Programming Division in the further evaluation of the TRAMO/SEATS model-based seasonal adjustment program and 2) 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 Program Directorate.

During FY2002, staff studied new variants of the most widely used time series model for seasonal time series which permit more independence between the properties of the trend component and the properties of the seasonal component. Software provided by visiting ASA/Census Research Fellow Siem Jan Koopman, and further developed by staff, made it possible to fit the new models to a large number of Census Bureau time series. At least one of the new models had better model fitting diagnostics for fourteen percent of the series. Staff developed software to produce model-based seasonal adjustments from the new models and diagnostics to assess various properties of the competing adjustments.

Staff developed time series models for 70 value of construction put in place (VIP) time series which included a component for the sampling errors in the direct VIP survey estimates and a component for the true underlying VIP series (ignoring nonsampling errors). The series were all for private nonresidential types of construction categories. Preliminary results were obtained with the fitted models for investigating the use of time series signal extraction to produce estimates that have lower variances than do the direct survey estimates. While the results indicated some potential for improving the estimates, difficulties were encountered due to the short length of the time series and the lack of estimates of sampling variances until recent years.

Staff completed a study to determine how many

years of data are needed to detect Easter holiday effects with high reliability and to obtain useful estimates when the effects conform to the models of X-12-ARIMA. The study used synthetic series obtained from the Time Series Methods Staff of the Economic Statistical Methods and Programming Division. The study shows that very strong Easter effects (about 8%) are detectable and estimable with four years of data; almost as much success can be achieved for moderate effects with six years of data, depending somewhat on the dates of Easter in the six-year span.

Versions of X-12-ARIMA/SEATS were developed that incorporated new features and code changes from 2002 releases of SEATS, procedural corrections developed by staff, and new features and code updates from Versions 0.2.9, 0.2.10, and 0.3 of X-12-ARIMA.

Staff: Donald Martin (x3689), Kellie Wills, Thuy Trang Ta Nguyen (MCD), John Aston, 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 adjustors and experts to use the program as effectively as their backgrounds permit. Additional text explaining how to interpret the roots from an ARMA process, how to avoid conflict between the seasonal adjustment mode and the transformation used for regARIMA modeling, and options for the sliding spans and history analysis related to the irregular regression component, was developed and added to the X-12-ARIMA Reference Manual with assistance from the staff from Time Series Methods Staff.

During FY2002, staff assisted the staff of TSMS in completing a revision of "Getting Started with X-12-ARIMA Input Files." Staff completed a final version of the paper "Modeling of Time Series Arrays by Multistep Prediction or Likelihood Methods" that has been accepted by the *Journal of Econometrics.* The article provides a theory for the fitting of possibly misspecified time series (ARIMA models) to data that are subject to regular revisions.

Staff: Brian Monsell (x2985), Donald Martin, 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 Ration Module and the SPEER edit systems. This is a joint project with the Economic Statistical Methods and Programming Division (ESMPD).

During FY2002, we developed and tested SAS software to generate the complete set of edits for a given set of explicit ratio edits. Staff did a thorough literature search into available algorithms and participated in organizing a seminar where an alternative approach to implicit ratio edit generation was discussed. Staff wrote software corresponding to two separate implicit ratio edit generation methodologies, one written in SAS/IML and the other written in SAS/OR. Both versions use a network flow approach. Each program checks the edit set for redundancies, determines if the edit set is consistent, and finally generates the implicit ratio edits.

Staff verified the edit generation programs using six separate data sets developed for the Plain Vanilla Service Sector Statistics Division (SSSD) test. Staff also tested the edit generation software using the Census Bureau's Annual Survey of Manufactures production edits. Staff presented a briefing and did a software demonstration for several interested individuals and end users from SSSD and the Economic Planning and Coordination Division. Staff delivered the SAS programs and a detailed user's guide along with the software demo. Staff incorporated suggested changes and extra software features per customer's requests. Staff wrote project reports describing the research and products of this project.

Staff: Maria Garcia (x4729), Bill Winkler, K.J. Thompson (ESMPD), Roger Goodwin (ESMPD)

1.20 FLEXIBLE MATCHING IMPUTATION RESEARCH (Economic Project TBA)

The purpose of this project is to apply the Flexible Matching Imputation (FMI) method to applicable economic surveys and compare the item imputations with those that are created by the existing method. The FMI method fits regression models to the observed data to determine a set of ranked matching variables for records containing missing item values. These variables are then used to match the missing item record with a donor record in a hot-deck imputation procedure.

During FY2002, staff imputed missing price, length, width, and number of bedrooms found on the Manufactured Homes Survey using the Flexible Matching Imputation (FMI) method. To compare the accuracy of the FMI procedure to the current method of imputation, we simulated missing data, performed the

imputations, and compared the results to the original observed data. We found that the FMI procedure performed better than the current method for imputing length, width, and price. We did not see as strong of an improvement for number of bedrooms.

Staff: Todd Williams (x4863)

1.21 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 FY2002, staff provided a technical review of the survey methodology and analysis of a parcel post weight study, and provided recommendations to the Commission regarding the utility of the survey results for products costing purposes. We also conducted an extensive review of a study of the impact of specific postal agreements relating to the regulation of international mail.

Staff: Leroy Bailey (x4917)

1.22 IT & AT FOR THE DISABLED/GENERAL SERVICES ADMINISTRATION
(Methodology and Standards Project 8083)

(See Projects 0351 and 1871, General Research, Statistical Computing and Technology - Additional Topics I.2.)

1.23 TESTING IMPROVEMENTS/U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
(Methodology and Standards Project 7225)

The purpose of this research is multi-faceted. First, differences in government health plan reporting will be explored by experimentally manipulating two survey design features: the sequencing of the Medicare and Medicaid items, and the definitions of these programs embedded within the questions. Four questionnaire treatments (based on a 2x2 experimental design) will be administered in the 2002 Questionnaire Design Experimental Research Survey (QDERS), to be fielded in the summer of 2002. A second goal is to collect additional detail on all types of health plans through respondent debriefings; these data will be used to analyze the validity of the initial reports. A third goal of the research is to gather sufficient detail from householders regarding the source of their health insurance coverage in order to conduct a later, separate study to validate that coverage with the insurance provider (e.g., employers and unions for employment-based plans, the Health Care Financing Administration for Medicaid and Medicare, etc.). Finally, order effects will be explored by experimentally manipulating the response categories in a question on general health status, and by manipulating the order of two items on physical exercise.

During FY2002, staff developed the 2x2 research design experiment for the Medicare/Medicaid research questions, the debriefing questions for the validation of all health plan reports, and identified key variables that would be necessary for a later follow-up establishment survey to validate householder reports. Staff also developed instrument specifications for each of the four questionnaire treatments, as well as the order-effects experiments, conducted testing and instrument debugging and completed the Interagency Agreement with the U.S. Department of Health and Human Services for financial support which helped increase the overall QDERS study sample size.

Staff: Joanne Pascale (x4920)

1.24 BAYESIAN STATISTICAL METHODOLOGY/NIST
(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 (NIST) towards setting up Bayesian Computation software and developing Bayesian methods of meta-analysis.

Staff started consultation on Bayesian methodology with NIST researchers in April 2002 which continues at the level of .20 FTE. Staff developed partition model methodology and accompanying software for doing meta-analysis when a uniform similarity among experiments is questionable. Staff is also working with a programmer to develop a MATLAB library of Bayesian analysis programs.

Staff: Don Malec (x4892)

1.25 COMPUTER ASSISTED SURVEY INFORMATION COLLECTION (CASIC)
(Methodology and Standards Project 4100)

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

(See Decennial Project E.)

B. Metadata Systems Research

The purpose of this project is to conduct research into the collection, use, and dissemination of metadata. This research includes development of metadata

standards, repositories, tools, and educating agency personnel about the latest developments and assisting with the implementation of these new methodologies. New to this research is the usage of a web based portal product to provide user controlled knowledge management. The scope of this research is being determined.

During FY2002, the Service Sector Statistics Division (SSSD) completed laying out the content for over 200 paper questionnaire forms for the 2002 Economic Censuses using an Economic Metadata Repository (EMR) graphical user interface tool built by the Corporate Metadata Repository (CMR) project team. The extensive SSSD requirements for this user interface tool evolved over time as the users became more familiar with the approach. During the latter part of the year, the requirements to allow the Manufacturing and Construction Division to input, review, and correct collected and published NAICS codes as well as ensure comparability with Bureau of Labor Statistics codes, were added and incorporated into the user interface. The economic EMR user interface makes use of the CMR data element registry to generate the dataset variable names and allowable values to be used in data collection. This data element registry is also in use at the Federal Aviation Administration and has generated interest at other federal agencies. A significant amount of time has been devoted, particularly in the fourth quarter, in briefings to the Executive Staff and key Directorate staff. These briefings have been both to provide accountability to the CMR stakeholders and to discuss funding issues. The big cost savings that the Economic Directorate expects its new automated questionnaire design system to provide is the ability to image and key all forms without the keyer having to enter a keycode for the data item. There will also be a huge cost saving provided by the ready availability of all forms content for the 2007 Economic Census.

Staff: Sam Highsmith (x1928), Aref Dajani, Mary Ann Scaggs, Ned Porter

1.26 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 (x1030), 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 (x4978), Tom Petkunas

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 data 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 3320051.)

During FY2002, staff implemented Gina Roque's method for masking data using simulated data, and added noise made up from a mixture of normal distributions. Matching software was used to test the effectiveness of the masking procedure. Staff developed microaggregation software and are testing and modifying the software using longitudinal Social Security earnings data.

Staff developed a webpage with URL "www.census.gov/srd/sdc/index.html" that will be used for a variety of research purposes related to statistical disclosure control work at the Census Bureau.

The Disclosure Review Board (DRB) completed 92 requests for review. The DRB issued an addendum to its checklist requiring further masking of certain data items on public use microdata files. Staff assisted the demographic area in interpreting and meeting the new requirements. Staff did much consulting with different people to develop and refine disclosure limitation procedures for various data: Population Division: migration data products, school district special tabulations, Public Use Microdata Samples; Demographic Surveys Division: National Crime Victimization Survey, National Epidemiologic Survey on Alcohol and Related Conditions; Bureau of Labor Statistics: Current Population Survey; Department of Housing and Urban Development: special tabulations; Department of Veterans Affairs: special tabulations; Department of Transportation: Census Transportation Planning Package; Planning, Research, and Evaluation Division: Census Quality Survey.

The data swapping was successfully performed on the Census 2000 sample data and the American Community Survey data. The data swapping for the Public Use Microdata Samples (PUMS) was also successfully completed. For the PUMS, weights were swapped because this produced a better national file, and Puerto Rico was processed separately from the states. Staff drafted the Census 2000 data swapping evaluation.

Staff coached and directed three summer interns on the microdata reidentification (data fusion) project. The interns found some potential disclosure problems with the Survey of Income and Program Participation. These will be corrected in future releases. They also wrote code that will make it easier and faster to use public and semi-public data. The interns also worked on reidentifying respondents from the National Survey of College Graduates, the Survey of Program Dynamics, and the Current Population Survey. This work continues.

At a meeting of the Data Stewardship Executive Policy (DSEP) Committee, staff presented an issue paper on using unswapped Census 2000 data for sampling for the National Survey of College Graduates. The committee decided that the Demographic area should use swapped data for software development, and the Decennial Systems and Contracts Management Office would then run the software on the unswapped data. At the same meeting, staff was asked to contact three federal statistical agencies (Social Security Agency, Agency for Health Care Research and Quality, and National Center for Health Statistics) to see if some detail in data they or their affiliates currently publicly release could be removed. Such detail is causing potential disclosure problems for Census Bureau public use files. Staff documented their findings. We are acting as consultants for a committee writing another Data Stewardship Executive Policy Committee issue paper on using waivers to collect data via email.

Staff tested the American FactFinder Advanced Query System. Staff analyzed the Advanced Query System filters for 100% Census 2000 data and produced some tallies from the data that show how often online queries passed or failed and why. They found that the median and sparsity filters are sufficient (the mean filter is unnecessary). Staff met with people from the State Data Centers, the Census Information Centers, and library data repositories, as well as outside beta testers, to answer questions on disclosure limitation procedures for the American FactFinder Advance Query System and other Census Bureau data products.

Staff: Laura Zayatz (x4955), Philip Steel, Paul Massell, Sam Hawala

B. Research File Development-Decennial Application

(See Decennial Project A.)

C. Small Area Estimation-Demographic Applications

Staff continued development of small area models, within small areas, for both the American Community Survey and census coverage estimation. These models incorporate a hierarchical structure.

During FY2002, staff presented a paper on a unit-level model for tract-level poverty rates using the ACS at the FCSM conference. (See Malec (2001) under *Section 4. Talks and Presentations*.) A new model and accompanying MCMC methods based on the logistic transformation, had been implemented for making ACS poverty estimates. Estimates have successfully been made from the model, and comparisons with a corresponding aggregate-level model have been completed and written in a report.

Staff: Don Malec (x4892)

D. Compensating 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 that 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 of Demographic Project 1465-C.

Staff: Leroy Bailey (x4917)

E. Some Special Topics in Mathematical Statistics

(See highlights of Demographic Project 1465 E.)

F. Seasonal Adjustment

(See Economic Project 3320052.)

G. 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.)

The data analysis, which examined the fit of models both with respect to CPS-estimated child poverty rates, and (in the 1990 data) with respect to the Decennial Census estimated rates, was conducted in Splus. This work is in a draft joint paper with Tapabrata Maiti.

Staff prepared for journal submission, an analysis of the fix to 1990 and 1994 SAIPE county-level data of the current (Fay-Herriot log-linear model based) small area estimation methods versus several variants based on unit-level mixed-effect logistic models. This data analysis also examined the differences between the CPS and decennial census child poverty rates. This work was written and submitted.

As part of the internal review for this paper, staff prepared and disseminated an internal document, "Linear Model Coefficients and Nondisclosure," addressing the accuracy to which coefficients in the log-child poverty rate SAIPE model (e.g., corresponding to 1990 and 1994 data) might be disclosed without in any way compromising CPS and IRS nondisclosure requirements.

Staff: Eric Slud (x4991)

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-coding and rounding. These approaches significantly distort variance and correlation between variables. Alternative approaches or remedies for this weakness will be investigated.

During FY2002, staff developed new computational procedures for the mixtures of additive noise approach of Roque (2000). Staff wrote new C++ rank swapping software with far faster algorithms than those of Domingo-Ferrer (2000). Staff applied rank swapping, ordinary additive noise of Kim (1986), and mixtures of additive noise to the Kim-Winkler data (1995) using the information-loss and disclosure-risk framework introduced by Domingo-Ferrer and Mateo-Sanz (2001).

Staff developed the expected value and variance formula for rounded data. Staff also investigated the assumption of the uniform distribution of 0, 1-9 in one's digit using 1991 Income Tax Return data.

Staff built multiple linear regression and logistic

regression models using a data set created by staff's additive noise and transformation approach. The estimates of regression coefficients for the regression obtained from the masked data were very close to those obtained from the unmasked data. The estimates from the two data sets for the logistic regression were found to be different.

"Disclosure Risk Assessment in Perturbative Microdata Protection" appeared in the monograph, *Inference Control in Statistical Databases.*

Staff: Jay Kim (x4907), William Winkler, Robert Creecy, William Yancey

H.2. Survey Design Issues

Aspects of survey design methodology identified for further research include such areas as decennial long-form weighting, collapsing methods in weighting cell development, and smoothing for improved estimation.

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. Staff 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. Staff developed bias and variance formula using the adjustment factor. We showed that current collapsing procedures can introduce huge non-sampling errors to the estimates.

Staff: Jay Kim (x4907)

Statistical Computing and Technology

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 partly funded under Decennial.)

During FY2002, staff modified re-identification software to include a new 1-metric that can yield higher reidentification rates with certain types of skewed distributions. Staff wrote the monograph chapter, "Disclosure Risk Assessment in Pertubative Microdata Protection" that will appear in *Inference Control in Statistical Databases*, Springer Lecture Notes on Computer Science (ed. Professor Josep Domingo-Ferrer). Related methods are given in SRD *Research Report (Statistics #2002-03*).

Staff taught the Record Linkage Course on February 25-28, 2002. Eighteen students participated. Each student was given extensive course notes, software, and documentation.

BigMatch software is designed for matching files of 100 million or more records against large administrative files having upwards of 4 billion records. The software is far faster than commercial software, yet still employs advanced string comparators and other sophisticated matching methods introduced in the Statistical Research Division. Staff created several new versions of BigMatch software. Decennial Census and division staff are testing the software on national census files for identifying duplicates. Documentation of BigMatch software is available in SRD *Research Report (Computing #2002-01).*

Staff investigated advanced methods of parameter estimation for record linkage. In "Improving EM Algorithm Estimates for Record Linkage Parameters," staff showed how to better apply the EM Algorithm for obtaining optimal parameter estimates in large matching situations where the proportion of matches within a set of pairs falls far below the minimum of 3 percent needed for mixture distributions. The methods should yield much better estimates of parameters for Decennial duplicate identification and large administrative record projects. In "Methods for Record Linkage and Bayesian Networks," staff showed how to apply a combination of maximum entropy methods from the machine learning literature and existing maximum likelihood methods for optimizing record linkage parameters when small or moderate amounts of training data may be available. The theoretical methods generalize methods for maximum entropy estimation due to Della Pietra and Lafferty (1997 *IEEE Transactions on Pattern Recognition and Machine Intelligence*). For dealing with additional levels of partial agreement for effectively comparing strings having moderate and large amounts of typographical error, staff developed new combinatorial algorithms for fitting I-Projections that generalize methods of Haberman (1973 *Journal of the Royal Statistical Society, C*). The software is significant because it yields new exploratory tools for determining when record linkage comparison algorithms can be improved. Because the algorithms are closely connected to the theoretical model of Fellegi and Sunter (*JASA 1969*), the algorithms may be more suitable than neural net methods that have also been applied to record linkage.

Staff: Bill Winkler (x4729), Ned Porter, William Yancey, Robert Creecy

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.

In the Fellegi-Holt model as exemplified by the SPEER editing software for continuous economic data, implicit edits are needed for error localization (i.e., finding the minimum number of fields to impute so that a record satisfies all edits). In the current version of SPEER, the easiest failed implicit edits (called *induced* edits) are generated during error localization for every edit-failing record. In the current stage of this project, staff applied the Fourier-Motzkin elimination method to generate a large subset of the implied edits prior to error localization in the SPEER edit system. The new SPEER version uses the implicit edits generated a priori which considerably simplifies error localization. Staff modified SPEER software routines for error checking, error localization, and imputation.

Staff made further progress in developing the imputation systems to complement the DISCRETE edit software. The focus was the implementation of discriminant analysis to impute relationship and marital status for all the members of a household when the ages are available. When it is reported, the age information is very reliable, as it is double-coded: age itself is reported, and birth date is also reported. Thus imputation rules that change age when it is so reported are suspicious. Discriminant analysis evaluates each configuration of relationship-marital status for the members and imputes the most likely pattern for a household, based on the relative frequencies of the relationship-marital status pattern for households, based on the relative frequencies of the relationship-marital status pattern for households with similar age configuration.

Staff completed two research papers, "Implied Edit Generation and Error Localization for Ratio and Balancing Edits," and "Bayesian Network Representation, Generalized Imputation, and Synthetic Microdata Satisfying Analytic Constraints."

Staff: Bill Winkler (x4729), Bor-Chung Chen, Maria Garcia, Yves Thibaudeau

C. Measuring Variance with Imputed Values

In this project, we investigate approaches to measure the variance of survey estimates when some of the data are missing and imputed. We focus on procedures that can be easily accommodated within the Standard Economic Processing System (StEPS), and work well with different types and mixtures of imputation.

During FY2002, we extended an alternative method – the use of an inflation factor – that requires very little additional computation and the creation of no additional data files, and compared it to Kim's method via simulation. We verified that, as expected, Kim's method eliminates most of the downward bias of the usual variance estimator. However, the variances of the usual estimator and the inflation estimator are often smaller than Kim's estimator. Thus, under mean squared error in the presence of mixed imputation, no one estimator consistently dominates the others, leaving more investigation to be done.

Staff: Pat Cantwell (x4982)

D. 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.

The second half of the fiscal year involved a transition from Exploratory Data Analysis (EDA) and Graphics to Graphical Methods Research for the fiscal year of 2003. Under EDA and Graphics, we taught a graphical data analysis course to a group of students from Russia, and sponsored a seminar presented by Dr. Jim Filliben of NIST. The seminar highlighted Tukey's contributions to EDA, and the advantages of using graphical data analysis over conventional methods. During this transition, we reviewed literature to examine graphical methods and contacted various areas of the Census Bureau to discover what kinds of graphical work was being done.

The software package ArcJmp was completed by Chris North (Assistant Professor of Computer Science at Virginia Tech) and given to several interested areas within the Census Bureau for testing. This software links the graphical abilities of JMP with the mapping displays of ESRI's MapObjects software.

Staff: Todd Williams (x4863), Ruben Mera, Pam Ferrari

E. 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 FY2002, staff continued to provide ongoing Hot-Line support for VPLX to the program directorates at the Census Bureau. Long-term support included research supporting multiple CPS supplements, the SIPP Variance System, the Census Long Form variances, variances for the Accuracy and Coverage Evaluation Survey, and the variances for the Survey of Construction. Staff provided VPLX support to multiple Census evaluation projects and made data files available for the development of the March 2001 CPS March Supplement variance program.

Staff began 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. The improved website will be easier for users to interface with each other. It will also be possible to run VPLX programs directly from the website. Enhancements to the documentation include a one-page, top-level summary of the features of VPLX in addition to a set of "Quick Start" documentation pages designed for the new user.

Staff: Aref Dajani (x4991), Sam Highsmith, Elizabeth Huang, Mary Ann Scaggs, Bob Fay (M&S)

F. Metadata Systems Research

(See M&S/CASRO Project 4100 B.)

G. SRD Portal Development

The purpose of this project is to conduct research into the usage of a web-based portal product to provide user controlled knowledge management. The scope of this research is yet to be determined.

A portal is a single page where all information is assembled, and access to both internal and external data are gathered. It provides a single point of interaction for users by utilizing a common interface. Portals will provide access to a wide range of web-enabled applications, bring together diverse data sources, and 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 FY2002, staff installed and configured the Oracle 9iAS portal product on an SRD system, with much appreciated assistance from the Oracle support branch of the Systems Support Division.

Staff constructed several unpublished web pages. These pages include all organization information including the division mission statement, organizational chart, staff and telephone numbers, Annual and Quarterly Reports, SRD Research Reports and Study Series.

Staff is creating the VPLX content area and the Record Linkage content area. The Usability staff is looking at setting up a knowledge management portal for usability projects. The development staff constructed a web page for public use throughout the Division. Further enhancements are being done on the VPLX content area and the Record Linkage content area.

The SRD Page was scheduled to be launched as an Intranet page on October 10, 2002.

Staff will continue developing forms, content areas, and pages for our Portal page. This will be an ongoing process.

Staff: Mary Ann Scaggs (x4966), Aref Dajani, Ned Porter, Sam Highsmith, Todd Williams

H. 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 FY2002, staff initiated research to investigate alternative computational methods for the solution of very large linear statistical models of wages versus person and firm effects. These models may be based upon tens of millions of records with a million person effects and one hundred thousand firm effects. Usual computational methods for solving the full model are not practical, so alternative iterative or sparse computational methods will be investigated. Parallel computing methods will be used as needed.

New work was initiated on the mixed model. After some research on existing algorithmic approaches, a new method was developed that is similar to the Expectation, Conditional Maximization (ECM) algorithm. Preliminary software provides estimates that match the estimates from the best known available mixed model software, ASREML, to 4 decimal places, while executing at least 17 times faster.

Staff: Rob Creecy (x4972)

I. ADDITIONAL TOPICS

I.1. Metadata Standards Research

The purpose of this project is the development of metadata standards.

The Census Metadata Standard was completed.

Staff: Bill LaPlant (x4887)

I.2. IT Accommodation Research

This research is intended to support either direct compliance by the U.S. Census Bureau with Section 508 of the Rehabilitation Act of 1973 as amended by the Workforce Investment Act of 1998. (PL 105-220) (codified as 29 U.S.C. 794) or to support the ability of Electronic & Information Technology (E&IT) suppliers and manufacturers to provide goods and services that comply with current or future versions of the Section 508 regulations. All of these activities continued through FY2002. During FY2002, staff continued to serve as Chairman of NCITS V2 (the IT Access Interfaces Technical Committee of the National Committee for IT Standards). NCITS V2 is developing standards that will support the development of Assistive Technology (AT) and 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.

The division agreed to provide staff to the GSA Center for IT accommodations (CITA) as a technical consultant on IT accessibility, standardization, and related issues for FY2002. Staff provided advice to CITA on IT accommodation research, AT/E&IT interoperability standardization, and the development of accessibility measurement methods. In this capacity, staff advises on the management of the Accessibility Forum, a GSA sponsored, industry led activity intended to make it easier for suppliers to provide government purchasers with 508 compliant E&IT equipment as soon as possible.

Staff spoke widely on the importance of standardization to the success of efforts to provide IT accommodation support for people with profound disabilities.

Staff: Bill LaPlant (x4887)

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 low-cost collaboration to any Census Bureau project 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: During FY2002, the Usability Lab sponsored four seminars for Census Bureau staff on usability and accessibility topics. Staff redesigned and printed a new brochure describing the lab's services throughout the user-centered design cycle. Staff presented talks, tours, and lab demonstrations to groups from within the Census Bureau and from other agencies (NASS, NSF, NCHS). We provided usability testing to domains which had added innovative new features to their Web sites and facilitated free uses of lab facilities by other government agencies. We developed a collaboration with Towson University Computer Science Program Masters-level students on two projects concerning the usability Web site. Staff produced a video about the usability and accessibility lab for Census IPTV.

Contracts: During FY2002, staff established or renewed contracts for research on information visualization and questionnaire navigation; usability testing and usability engineering services; applications development; specialized hardware and software support and eye tracking equipment. Staff terminated one contract for knowledge base development and one for programming support.

Facilities: During FY2002, staff opened the Accessibility Testing Facility to test electronic applications for usability by disabled users; opened two additional remote sites (Los Angeles and San Francisco) for usability testing; installed a video editing system to replace the Speed Razor system; created and printed a brochure to inform volunteer testers about usability testing; designed a database system to capture tester responses to questionnaires at remote sites; acquired new software: Camtasia, to create animated help applications for interfaces; Dreamweaver for Web site information about Census Bureau employees who have volunteered to serve as usability testers. We obtained new personal computing equipment for the fixed lab testing suites. We added push-to-talk headsets for test administrators to reduce the amount of stray noise when conducting simultaneous lab sessions. We acquired

new video cameras for evaluating computing devices used in the field. Staff created new documentation of usability procedures: a testing report template and some guides to card sorting, usability testing, logging and voice editing. Semi-annual maintenance of the fixed laboratory equipment was conducted by a contractor in April and September.

Staffing and *Training:* During FY2002, we mourned the loss of David Mingay, a key member who died this spring. Staff retained one outstanding intern from the summer program. We hired four other interns throughout the year, three of whom left by the end of the fiscal year. Staff attended seminars, courses or tutorials on Web design, information architecture, usable technical writing, program management, accessibility testing, mobile device testing, remote testing, and survey sampling.

Staff: Kent Marquis (x4719), Cyntrica Eaton, Joyce Farmer, Brian Greenbaum, Jon Jezyk, Larry Malakhoff, Safia Hamid, David Mingay, Betty Murphy, Erica Olmsted, Lorraine Randall

B. Usability Collaboration Projects

B.1. Project Management Repository and Quality Management Repository

The purpose of this joint Computer Assisted Survey and Research Office (CASRO) and Human Resources Division (HRD) project was to evaluate the usability of two Census Bureau Intranet sites prior to rollout and to monitor usability after the site is launched on the Census Bureau Intranet. These sites are intended to enable internal users to share, manage, and disseminate the Census Bureau's best practices, standards, and guidelines in both project and quality management. Usability is key to actual usage and user-satisfaction levels.

There was no testing activity this year, but CASRO has inquired about conducting a focused usability test on a new QMR functionality. Testing may occur in the first quarter of FY2003.

Staff: Betty Murphy (x4858), Kent Marquis

B.2. Foreign Born Webpage Usability

The purpose of this study is to evaluate the usability of the Foreign Born domain of the Census.gov Website. This will include establishing usability goals, assessing users' efficiency, accuracy, and subjective satisfaction with the site, diagnosing, and prioritizing usability problems, and making recommendations for how to make the site better. This will be the first study to collect data from a West Coast remote testing site.

During FY2002, the usability staff collaborated with Foreign born domain staff to create the preliminary documents for the usability study. These included a list of common user tasks, the usability goals, time and accuracy requirements for the tasks, and the testing script. Usability staff worked with two remote usability testing sites Baltimore and Los Angeles, (there were unforeseen technical difficulties with San Francisco) to recruit users and facilitate the usability test sessions. Usability staff completed five usability testing sessions with Los Angeles and Baltimore. Staff then wrote the final reports which included a list of the usability violations, recommendations for change, and video clip highlights. The project has been completed.

Staff: Erica Olmsted (x4893), Brian Greenbaum, Kent Marquis

B.3. 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 1) Automated Listing and Mapping Instrument (ALMI) and 2) the Group Quarters Automated Instrument for Listing (GAIL). Usability evaluation activities may include usability goal setting, task analysis, prototype evaluations, and usability testing in the field. Before the ALMI project is started, Lab staff will familiarize themselves with the issues by collaborating with the Computer Assisted Survey and Research Office (CASRO) on pen-based, mobile computing testing.

During FY2002, Pen instrument interviews filmed at Suitland House were edited and burned to CD for use by CASRO for demonstrating pen-based interviews. Lab staff worked out a usability evaluation project plan and schedule with the Demographic Statistical Methods Division. They specified a usability engineering contract that would cover participation in planning, data collection, and report writing for this project in addition to usability engineering services for other projects. They selected the contractor, did informal task flow analysis for the main procedures, conducted a dry-run of their data collection procedures in Lancaster, PA, collected usability data at 15 sites throughout the United States (mainly observing Locate Address procedures), and began data analysis.

Staff: Kent Marquis (x4719), Betty Murphy, Larry Malakhoff, Erica Olmsted, UserWorks, Inc.

B.4. 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 FY2002, usability staff worked with ACSD staff to create a questionnaire for ACSD branch chiefs and administrative staff to gather content information for the site. Usability and ACSD staff collaborated in a card sorting exercise with 9 users to determine the organization of the high level topics for the main page of the Web site. Usability staff analyzed the card sorting data using cluster analysis procedures and interpreted the results for the client. ACSD took the card sorting results and our recommendations and created a low fidelity version of the site. Usability staff performed two iterations of low fidelity usability tests with nine different users. Recommendations for improvement, along with explicit examples on how to improve the content's usability were delivered to the development team. The team made changes to the site that we recommended and returned for final exploratory usability testing of the working site. For the final usability test, we worked with twelve users. The usability report is forthcoming. Usability staff also worked with ACSD staff to evaluate the accessibility of the developing Intranet site using Insight/Infocus and Jaws software.

Staff: Erica Olmsted (x4893), Brian Greenbaum, Kent Marquis

B.5. Decennial Management (DMD) Intranet Site Accessibility Evaluation

DMD wishes to test its new Intranet site with a screen-reader (JAWS) and to evaluate the site for accessibility using In-Focus evaluation software.

During FY2002, staff assisted DMD in two sessions to evaluate its site and provided a printout of the evaluation results. The project has been completed.

Staff: Larry Malakhoff (x32688)

B.6. Field Directorate Intranet Site Usability Testing

The Field Division (FLD) at headquarters and the Technologies Management Office (TMO) wish to conduct a usability evaluation of a new Website for the Field Directorate, based on the oracle Portal architecture. The Usability Lab will provide usabilitiy testing and a final report including remote tests at regional offices and telephone centers, if feasible.

During FY2002, usability staff worked with Field Division and TMO to profile the main user classes and tasks each group will come to the Intranet site to accomplish. Usability staff, with a contractor from UserWorks, conducted a usability evaluation of the Field Directorate's redesign of its Intranet site, using seven volunteer users from headquarters. We analyzed the results of the study and wrote a report. The report included quantitative data describing how user performance compared to the usability performance goals. It also included an analysis of usabilitiy violations, recommendations for change, and video clip highlights. The report was sent to the client on July 1, 2002. The Field Directorate has made changes to the site and has requested a second round of usability testing in October, at headquarters and selected regional offices.

Staff: Erica Olmsted (x4893), Kent Marquis, UserWorks, Inc.

B.7. 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 Website. A goal of the project is to establish a blueprint for incorporating usability and user-centered design into all the HHES sites that are scheduled to be redesigned.

During FY2002, usability staff worked with HHES staff to create card sorting topics for the poverty and disability sites. Usability staff established two different procedures for the card sorting activity: creating card sorting topics from content labels or titles and creating card sorting topics from user questions, in a question format.

Usability staff analyzed the card sorting results of the poverty and disability sites and gave recommendations to HHES about the organization of higher-level information for their websites. Usability staff is currently working with HHES staff to prepare low fidelity prototypes of the disability and poverty sites. Low fidelity prototype testing is scheduled for October.

Staff: Erica Olmsted (x4893), Kent Marquis, Betty Murphy

B.8. Information Technology (IT) Intranet Site Usability Testing

The IT Intranet development team wants to ensure that its redesign of the existing site is usable. Through usability testing, we hope to discover usability problems and make recommendations to correct these problems.

During FY2002, staff conducted preliminary information architecture activities, identified two distinct user groups for the new site, drafted a set of

tasks for each user set, established usability goals for accuracy, efficiency, and satisfaction. Staff designed and conducted a usability study with ten participants in late April, issued a report with correction recommendations in mid-May, and provided a more detailed, final report in early June that included video highlights. The IT Internet development team implemented many of the recommendations and rolled out the redesigned site. This project is complete.

Staff: Brian Greenbaum (x4893), Erica Olmsted, Kent Marquis

B.9. Mobile Computing Devices - Map Usability Testing

The purpose of this project is to test map-display software for feasibility on two different mobile computing devices (MCDs). Initial testing occurred in the field, with experienced field representatives and former Census 2000 enumerators serving as test users. They were given a list of addresses to verify, and used the map displays as navigational aids. Our division's role involves consulting on the design of the test, development of feedback questionnaires for test users and observers, and analysis of the training, performance and debriefing data collected by video. This test of the MCDs and map displays evaluated their feasibility of use by enumerators in Census 2010. Further testing is planned.

During FY2002, staff took the lead in developing and documenting the test strategy, and contributed to development of other testing materials. Staff worked with Field Division personnel in refining procedures for the user's manual, which was used as the basis for training. Staff participated in both the dry run with experienced field representatives, and in the week of testing, with the less experienced address verifiers. Staff served as observers and videographers, along with other Census Bureau staff from Suitland and Charlotte, VA. Following the field testing in Cloucester County, VA, staff assisted the Decennial Management Division (DMD) in the analysis of observer feedback and prepared a report on successes and issues based on the videotaped observations. This report was incorporated into DMD's final report to the MCD working group. The project is completed.

Staff: Betty Murphy (x4858), Kent Marquis, Larry Malakhoff, Erica Olmsted

B.10. 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 will propose testing strategies to identify the best approach to encouraging citizen response in the 2010 Census.

During FY2002, led by staff from the Decennial Systems and Contracts Management Office, the working group developed functional and data/reporting requirements for both the IVR option and the Internet option for the 2003 National Census Test, and focused on selecting contractors to support development of the IVR and Internet options. Division staff participated in the work of both the IVR and Internet sub-groups, contributing previous experience in the use of IVR and the Internet in Census 2000. The group developed an experimental design for the test. Staff also worked on language of the cover letters, reminder postcards, reminder phone calls, and IVR guide. We arranged for cognitive testing of the mailings developed for the five panels designed to test the Internet option. Staff wrote the IVR Census short form script using comments and requirements from the Population Division and others, and delivered it in July. Staff participated in the preparation and review of material for the OMB package prior to its submission for Census Bureau approval.

Staff: Betty Murphy (x4858), Larry Malakhoff, Kent Norman (University of Maryland)

B.11. Electronic Response Options (ERO) Subgroup

The charter and deliverable for this working group was to make recommendations for various electronic options for reporting census short form data to the Response Options Subgroup (ROS). Staff was tasked with the Telephone Self-Response Options section of the report due in mid-April.

During FY2002, staff worked closely with staff from the Decennial Systems and Contracts Management Office to make recommendations for an IVR system with CATI backup for the census pretest in 2004. Staff began modifying the IVR script used in 2000 based on the suggestions given by Westat, Synectics, and results of the customer satisfaction survey. This script will be provided to SpeechWorks which will program the IVR. This project has been completed.

Staff: Larry Malakhoff (x3688)

B.12. Foreign Trade Division (FTD) Internet Site Accessibility Evaluation

FTD wishes to test its Internet site with InFocus accessibility evaluation software the JAWS

screenreader for accessibility by blind users.

During FY2002, staff from FTD used our InFocus software and JAWS to evaluate the accessibility of the FTD Web site.

Staff: Larry Malakhoff (x3688)

B.13. American Community Survey (ACS) Tutorial Accessibility Evaluation

A contractor, developing a Web-based tutorial for users of the ACS's Website, needs to assure that the Website conforms to Section 508 guidelines for accessability by disabled users. As part of this assessment, they will evaluate their tutorial using the Census Bureau's accessibility evaluation facility.

During FY2002, staff from the University of Baltimore visited the lab to test the accessibility of the online ACS tutorial. The application had problems with "non-skippable links," which means a non-sighted user might have to listen to a screen reader announce navigation links that may not be useful. Staff from the Systems Support Division assisted in resolving this issue by providing a software "fix."

Staff: Larry Malakhoff (x3688)

B.14. Census 2000 Gateway Testing

The Census 2000 Gateway is one of three Web sites within www.census.gov that contains new features and that was nominated by the Census Bureau's Website Transition Team for free usability 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. All testing will be conducted using the lab's remote-testing capabilities.

During FY2002, working with the contractor, UserWorks Inc., staff identified tasks for test users to perform. We developed a screening instrument to be used by staff at the remote-testing sites (Baltimore, San Francisco, and Los Angeles) for recruiting testers. Testing with both novice and expert data users began in mid-July 2002 and continued through the end of FY 2002. A report will be prepared early in FY2003.

Staff: Betty Murphy (x4858), Erica Olmsted, UserWorks, Inc.

B.15. Population Projections and Estimates Web Page Usability Evaluation

One of three Websites within www.census.gov that contains new features and that was nominated by the Census Bureau Website Transition Team for free usability testing. The site was developed by an outside contractor. Currently, it is managed by an entirely new development team which is interested in learning about the site's usability. The Usability Lab will use the remote facilities managed by the Census Information Centers to evaluate the site's usability.

During FY2002, usability staff worked with Population Division (POP) staff to set usability goals for users of the site, to define users and typical tasks, to recruit users' representative of the non-expert group at the remote sites, and to define representative tasks for the usability study. Usability staff, with staff from UserWorks, Inc., designed and conducted the usability study. We analyzed qualitative and quantitative data. Lab and contractor staff prepared a preliminary report of results for the POP staff, and began preparing the final report which includes a list of usability problems, a section on how to resolve the problems, and video clip highlights.

Staff: Erica Olmsted (x4893), UserWorks, Inc.

B.16. Heuristic Review of the M3 Electronic Questionnaire

The M3 is one of the first economic surveys to use the new web questionnaire authoring system developed by the Census Bureau's Systems Support Division, without the benefit of user-centered design processes or usability testing. Early in the process, Usability Lab staff conducted one expert review and suggested that additional reviews would be beneficial. At the request of the Manufacturing and Construction Division, the Usability Lab agreed to arrange for an outside evaluation of the questionnaire.

During FY2002, UserWorks conducted two reviews of the much revised application and submitted a report based on a consensus of the reviewer's evaluations and recommendations. This project is complete.

Staff: Beth Nichols (x4865), Kent Marquis, UserWorks, Inc.

C. Usability Laboratory Research

C.1. Remote Usability Testing Methods

The Customer Liaison Office has encouraged the Usability Lab to develop methods equipment, software, and contractual relationships to conduct usability tests using participants at remote Census Information Centers (CICs). The staff of the Usability Lab and the Telecommunications Office are collaborating to develop the remote testing capability and to evaluate it.

In FY2002, we installed equipment in two new locations, one in Los Angeles at the William C. Velasquez Institute and one in San Francisco at the Chinese American Voter Education Committee, and

began usability testing. We continued remote testing operations at the Baltimore Urban League location.

Staff: Kent Marquis (x4719), Erica Olmsted, Safiya Hamid, Chad Russell

C.2. User-Interface Standards and Guidelines

The purpose of this activity is to develop user-interface design standards and guidelines to foster usability, consistency, and accessibility across Windows-based and Web-based applications/sites for data collection, data analysis, data dissemination, and internal Census Bureau functions. Additional objectives are: 1) to link these guidelines to other IT standards and style guides developed for user-interface designers and implementers; 2) to put them in a form most useful for users.

During FY2002, following a Census Bureau-wide forum and review, staff worked on resolving issues and comments received from reviewers. Staff then submitted a revised version of the draft standard to the Standards Management Team. The Information Technology (IT) Directorate issued IT Standard 20.0, *Design Guidelines and Standards for Web-Based User Interfaces.* The first round of revisions is scheduled to begin in October, 2002.

Staff: Betty Murphy (x4988), Kent Marquis, Bill LaPlant

C.3. Visualization Research

The goal of this research is to explore possibilities for developing and applying flexible data-visualization techniques to Census Bureau data for both internal use and external data dissemination in collaboration with the University of Maryland's Human Computer Interaction Laboratory (HCIL) and Professor Ben Shneiderman. The task for FY2002 is to complete earlier exploratory work and then develop a prototype of a dynamic query interface for the Web, using Java.

During FY2002, the HCIL completed the prototype DHTML version of DynaMaps. Professors Golub and Shneiderman wrote a paper summarizing the approach and submitted it for publication. HCIL improved the JavaScript version with color-coding that corresponds to the selected and unselected areas on the sliders. In addition, a number of design flaws were corrected as a function of expert review and usability testing by the Usability Lab at the Census Bureau. A graduate student, Haixia Zhao, continued working on a Java implementation of DynaMaps. Displaying the details for the 3100 counties proved to be problematic in the web environment, especially if zooming was to be supported. She successfully solved basic performance problems on the county level maps. The project will continue into FY2003.

Staff: Kent Marquis (x4719), Human-Computer Interaction Lab at the University of Maryland, College Park, MD.

C.4. Electronic Questionnaire Navigation Research

This is long-term research with the University of Maryland, directed by Professor Kent Norman. The purpose is to understand principles and identify guidelines for designing electronic questionnaires for self-administration by establishment and household respondents. Topics include page vs. scroll formatting, how to manage conditional branching (skip patterns), the presentation of edit error messages, the coordination of information in company on-line databases with questionnaire demands, and the effect of one-time vs repeated surveys on performance.

During FY2002, Professor Norman analyzed data from two prior experiments on conditional jumps and questionnaire edits (error messages) and wrote two technical reports. The paper on conditional jumps was accepted into the *2002 Proceedings of the Human Factors and Ergonomics Society* for presentation at the October 2002 meeting. The research in questionnaire edits supported the design guideline to make the user aware of the need for answer corrections immediately rather than in a batch at the end of the questionnaire. In addition, Professor Norman wrote a technical report on the level of automaticity used in surveys with conditional jumps. This report outlines the tradeoffs between the respondent's need for control and contextual information and the efficiency provided by automatic conditional jumps. This research has helped to determine guidelines for designing computerized self-administered questionnaires and is currently being used to design optimal approaches for on-line surveys.

Staff: Kent Marquis (x4719), Human-Computer Lab and the Lab for Automation Psychology, University of Maryland, College Park, MD

C.5. Usability Intranet Site

With contractor support, staff will create and install an Intranet site to disseminate knowledge to Census employees about user-centered design and about tools useful in the design process. Staff will evaluate the site.

During FY2002, the contractor completed work and submitted a prototype version of the site for Census Bureau staff to modify and publish. Staff made some modifications and updates but did not actually publish the site. We arranged for a masters-level class in usability testing at Towson University to 1) identify the users, user tasks, and usability goals for the site and 2) to conduct an expert review of the site. The students completed their projects and submitted reports.

Staff: Kent Marquis (x4719), Larry Malakhoff, Betty Murphy, Humans and Computers, Inc., and UserWorks, Inc.

C.6. DynaMaps Card Sorting

The goal of this project is to improve the users' ability to find variables within the list of criteria fields that were used in the DynaMaps application. They were originally not in any order and provided frustration to users when they tried to search for specific ones. Card sorting by users will be employed to suggest how to organize the fields in a way that was easily understood by the users.

During FY2002, three successful card sorting sessions were completed, and the results were analyzed using a clustering algorithm. A grouping strategy was derived from the results and implemented in the application in late March for the next iteration of testing. A report was drafted in late March.

Staff: Erica Olmsted (x4893), Jon Jezyk

C.7. DynaMaps for CD-Roms

The purpose of this study is to refine a concept developed by the University of Maryland into a usable dynamic map interface for a CD-ROM that allows users to interact with data in a graphical format. The intent is to display Census Data on maps, scatter plots in tables so that novice users will quickly learn and use the application to meet their data retrieval goals. An objective of the developers is to create an interface that keeps the needs of all users in the forefront of all design decisions. Through iterative design and testing, the usability staff and the developers will uncover problems and improve the user interface. Lessons learned will be applied to both CD-ROM and Web versions of the DynaMaps application.

During FY2002, staff reviewed tapes from the last usability test and made additional design modifications. Informal subsequent testing revealed that additional modifications were needed. These modifications were specified to the programming contractor.

Staff defined and implemented several changes to the interface, including a reordering of the labels of the sliders, a deletion of a number of functions that distract new users, and the placement of help in the center of the screen. In preparation for usability testing, staff developed two different versions of how the user adds sliders to the display window; these versions will be tested separately from the DynaMaps application and compiled after the preferred version is identified. We contracted with Professor Chris North at Virginia Polytechnic Institute to develop a beta version for the Counties CD over the summer, and to create a final product in the fall. Staff reviewed the beta version and recommended modifications to the product before usability testing begins.

Staff: Erica Olmsted (x4893), Brian Greenbaum, Jon Jezyk, Kent Marquis, Tom Petkunas

C.8. DynaMaps Help Design and Testing

In previous iterations of the DynaMaps application, we observed 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.

During FY2002, a new help system was developed using centrally located colored buttons to attract attention, and video demonstrations of common functions. Five usability tests were completed using the new help system, and the initial results suggests that learning was improved.

In the second iteration of the DynaMaps help system, usability staff conducted one usability test which showed us the direction we need to move for redesigning the help. We have met several times with the HCIL lab at the University of Maryland to discuss a third iteration of a DynaMaps help system. Based on a prototype we developed at the Census Bureau, the HCIL Lab developed a working proof-of-concept demo using "sticky notes" to direct the user on how to use the application's various functions. Usability staff has worked on the actual wording of the help. We are in the process of preparing a low-fidelity prototype as a basis for further validating the wording and positioning of the sticky notes.

Staff: Erica Olmsted (x4893), Brian Greenbaum, Tom Petkunas, Kent Marquis, University of Maryland Human-Computer Interaction Lab

C.9. DynaMaps Behavior Coding

The goal of the behavior coding project for DynaMaps is to evaluate whether the users of the revised application were able to learn more about how to use the main functions than users of the original version.

During FY2002, staff developed behavior codes and rules. They began using the new Observer 4.0 logging software to do the behavior coding of the explore phases of the first two iterations of DynaMaps testing. Coding staff achieved a .84 inter-rater reliability correlation on the coding of a trial session. Results, using either original or revised coding rules appear to demonstrate substantial improvements in learning.

Staff: Jon Jezyk (x4893), Brian Greenbaum, Erica Olmsted, Kent Marquis

C.10. Questionnaire Database

The goal of this project is to provide a method of gathering and storing data from questionnaires that are administered during remote testing. The MS Access program will display the questionnaires as well as store all of the data in one place, which will aid in analysis.

During FY2002, staff created the database and a prototype questionnaire and installed required software. Initial tests revealed a number of technical and usability issues. The concept appears sound, but further implementation will be postponed until resources are available to overcome the problems. Meanwhile, we will continue to ask remote users to fill out paper questionnaires and ask the remote site personnel to fax them to us after each remote testing session.

Staff: Jon Jezyk (x4893), Eileen O'Brien

C.11. 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 FY2002, the University of Maryland designed an experiment using DynaMaps for acquiring information about statistical attributes of the states. The purpose of this study is to verify the utility of displaying data in a spatial rather than in tabular form. Two hypothetical databases were constructed with a number of variables. Some of the variables were spatially related (e.g., East-West) and some were intercorrelated with each other. Participants were asked to use the DynaMaps interface to identify states with values in target ranges. A series of guided, detailed instructions were used for the initial queries followed by short problem statements for later complex queries. Time and accuracy were recorded. Thirty subjects were run in the experiment. The results indicated that subjects were fairly fast and accurate with their answers. Results showed that subjects were able to learn how to find information when given explicit instructions at the beginning. In addition, after the series of problems, they were tested on implicitly acquired knowledge about interrelationships between
the attributes. As expected, attributes that were regionally oriented were easier to understand and related to other variables than attributes that were independent of geographic location.

Staff: Kent Marquis (x4719), University of Maryland Human-Computer Interaction Lab and Laboratory for Automation Psychology

C.12. Accessibility Testing Facility

The accessibility testing facility was created for Census web developers and contractors to test the accessibility of their web and computer applications to users with disabilities. Developers will be able to test the accessibility of their web and computer applications with both assistive and diagnostic software, on different platforms, on different operating systems, and with slow modem Internet connections. The facility will have special input devices for examination by all Census Bureau employees who are or have become disabled. These devices include a head tracking mouse pointer, a customizable keyboard, and joystick pointers.

During FY2002, we purchased and installed the planned equipment and software. This facility was publicized in a broadcast message in January. As a result of this broadcast message, we received a number of requests to use the facility. Staff created an administrator's guide with screen shots about how to log on and set up the accessibility testing environment for Census Bureau clients. Staff subsequently upgraded the screen reading software (JAWS) and added software to test the accessibility of non-Web applications for people with disabilities.

Staff: Larry Malakhoff (x3688)

C.13. Census Web Site Evaluation Project

Using existing studies of 5 key parts of the Census Bureau's Web site, the aim of this project is to uncover the major usability problems that are common across the domains. In FY2001, interns and Census Bureau staff analyzed 5 existing domains, created typical user and task profiles, elicited usability goals for user accuracy, efficiency, and subjective satisfaction, and conducted usability tests by videotaping each usability session, logging detailed time-based notes, having users perform typical tasks, and obtaining user satisfaction ratings. During FY2002, we will complete any projects the interns left unfinished, and examine their results for common usability shortcomings across the domains.

Our reexamination of the results involved calculating average levels of user accuracy and efficiency across the sites and content analyzing the specific usability findings for common themes. We found uniformly low levels of user accuracy and efficiency in all the studies. The content analysis showed that user problems were due to several common design problems: inadequate navigation schemes, excessive use of jargon, and problems with the way information is displayed in tables. We presented the results to the Census Internet Management Team, the

Census Internet Coordinators Group, the Census Web site Transition Team, and in a paper at the 2002 meetings of the American Statistical Association.

Staff: Erica Olmsted, (x4893), Kent Marquis

C.14. 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. In laboratory testing, test respondents will work with mocked-up survey prototypes incorporating different design options for the issues of interest. Staff and management from the economic directorate have been involved in identifying the issues and crafting the test plan. The planned laboratory research is 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 FY2002, nine electronic form layout issues were identified as needing empirical research for business surveys. During the laboratory phase, we plan to use a split panel design to study speed and accuracy while using different question-and-answer layouts for a business survey. Using current Census Bureau business surveys and censuses as our models, we crafted a fictitious electronic business survey with 29 items. We also created electronic records to be consulted to complete the survey. Each item in the survey had one of the 9 layout issues of interest. With advice from Dr. Kent Norman of the University of Maryland's Laboratory for Automation Psychology and Decision Processes, two different layouts were created for each of the 29 items. Using a fractional factorial design, 8 surveys were created with different permutations of the layouts for the 29 items. Pre- and a -post - questionnaires were created to collect demographics and skill level of the participants, and opinion data on different layouts presented. Five cognitive tests were performed on the survey. The survey versions and records were modified and debugged accordingly. The laboratory phase is to commence in the fall of 2003 using University of Maryland students as our subjects.

Staff: Beth Nichols (x4865), Betty Murphy, Cyntrica Eaton, Kent Marquis, University of Maryland Laboratory for Automation Psychology.

C.15. Usability Knowledge Base

The usability lab will create a repository of content pertaining to usability terms, methods, and material. This repository will be known as the usability knowledge base. It will be stored and maintained in an Oracle database. All Census Bureau staff will have access to portions of its content via the usability Intranet site, the usability standards, usability style guides, and through any tutorials or courses written by the lab. The goal is to create content useful for developers and managers who lack a usability background and which is consistent across the dissemination media, consistent with current literature and practice, and easy to find, understand, and apply to real-life projects.

During FY2002, a contractor began to outline the elements the usability knowledge base would contain. She conducted a brainstorming session with usability staff and used her own expertise in the field to provide a listing of Web links pertaining to usability and usability and user-centered design issues. Usability staff refined the list and assigned each element to a lab staff member who will create the first draft of the content.

Staff: Beth Nichols (x4865), Betty Murphy, Erica Olmsted, Larry Malakhoff, Kent Marquis, Humans and Computers, Inc.

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.

During FY2002, staff completed and distributed a research report documenting the results of cognitive research on the American Community Survey (ACS) related to "seasonality" questions for people who have multiple residences.

We completed and distributed a research report documenting the results of cognitive interviews with current and former members of the military to test proposed changes to the veterans' questions for the American Community Survey. Recommended changes were incorporated into the ACS.

We reviewed the Public-Police Contact Supplement to the National Crime Victimization Survey, and held numerous meetings with staff from the Demographic Surveys Division and the Bureau of Justice Statistics to discuss and revise the questionnaire. Questions for the supplement were finalized and a report is forthcoming. Recommended question wording changes were incorporated into the revised questionnaire.

We completed and distributed a research report documenting the results of cognitive research on a

newly-proposed question on cash balance plans for the SIPP Retirement and Pension Plan Coverage Topical Module. Recommendations are being incorporated into the module.

Staff: Terry DeMaio (x4894), Ashley Landreth, Kristen Hughes, Jenny Hunter, David Mingay, Lorraine Randall

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 FY2002, staff monitored the generic clearance, consulted with staff from other areas of the Census Bureau wishing to use the clearance, and kept OMB informed of all pretesting activities. Twenty-nine letters were submitted to OMB for work done under the generic clearance, with a total of 4,380.3 burden hours. Staff gave two presentations on questionnaire pretesting as part of Demographic Directorate Training on demographic surveys. Staff also chaired an interdivisional group, including participation by the Economic Statistical Methods and Programming Division, the Decennial Systems and Contracts Management Office, and the Demographic Surveys Division (DSD), which was charged by the Methodology and Standards Council with developing corporate standards for pretesting censuses and surveys. Several drafts of the standards were developed and reviewed, and the standards are close to being finalized.

Staff: Terry DeMaio (x4894), Kristen Hughes

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

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

Using QDERS 1999 data, staff published an article in *Public Opinion Quarterly* examining person-level vs. household-level questionnaire design.

Staff: Jennifer Rothgeb (x4968), Joanne Pascale, Ashley Landreth, 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 will be conducted from the Hagerstown Telephone Center in September 1999, using an RDD sample. Five researchers conducted questionnaire design experiments and two researchers conducted interviewer training experiments.

During FY2002, staff analyzed QDERS 2000 data and presented results of their experiments at domestic and international conferences.

Staff: Jennifer Rothgeb (x4968), Joanne Pascale, Eileen O'Brien, Tom Mayer

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 May/June 2002, using an RDD sample. Three researchers conducting questionnaire design and survey methods experiments are participating.

During FY2002, staff developed instrument specifications, case management specifications, and interviewer training. To facilitate the necessary experimental design, staff conducted four different training sessions during each of the four data collection periods with the Tucson Interviewer staff. The survey closed out on schedule. The raw data will be processed through DSD's TRANSCASES CATI processing program during FY2003.

Staff: Jennifer Rothgeb (x4986), Joanne Pascale, Ashley Landreth, Terry DeMaio

I. 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.

In FY2002, we continued to investigate the effects of public attitudes and perceptions about privacy and confidentiality, and to research methods to address respondent concerns and discourage nonresponse behavior by chairing and participating in an Interagency Household Survey Nonresponse (IHSNG) Group sub-group on Privacy and Confidentiality. The sub-group provided a forum to discuss related research with which various agencies are involved. We provided information and suggestions to other researchers and,

where appropriate, suggested possible opportunities to collaborate. We discussed the possibilities to organize and coordinate a cooperative, integrated research agenda that will identify and propose new research that the IHSNG might sponsor.

Staff: Tom Mayer (x4930), Eleanor Gerber, Melinda Crowley

J. 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 FY2002, staff advised and responded to the Government Accounting Office's (GAO) survey on the use of ethnographic methods in federal evaluation practice and advised GAO's formulation of a pilot ethnographic GAO evaluation of another government agency.

Staff attended the quarterly interagency meetings on migrant and seasonal farm workers.

Staff submitted reports for three Decennial ethnographic evaluations, "Ethnographic Social Network Tracing in the Context of Census 2000," "Comparative Ethnographic Research on Mobile Populations," and "Complex Households and Relationships in the Decennial Census and in Ethnographic Studies," and for two experiments, "Generation X Speaks Out on Censuses, Surveys, and Civic Engagement: An Ethnographic Approach," and "Protecting Privacy: The Ethnography of Personal Information Management." Staff prepared a draft report concerning the "Enumeration Barriers Specific to *Colonias*." (For more details, see Decennial Projects.)

Staff: Leslie Brownrigg (x4995), Manuel de la Puente, Eleanor Gerber, Laurie Schwede, Melinda Crowley, Kristen Hughes

K. ADDITIONAL TOPICS

K.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 to researchers to enable them to improve the questions.

During FY2002, two versions of questions on telephone expenses, transportation, and environmental issues were included in QDERS 2000 to facilitate our ongoing research examining the effectiveness of different pretesting methods. The independent outcome measures we used to evaluate the results of the QDERS data included behavior coding, interviewer ratings of perceived interviewer problems and interviewer ratings of respondent problems, as well as item nonresponse. Our findings indicate that our earlier pretesting results (using expert review, question appraisal and cognitive interviews) did predict actual problems as evidenced by our independent outcome measures from QDERS. Of interest was also whether the pretesting techniques served to identify improved survey outcomes. We obtained mixed results. We saw improved item nonresponse in the experimental questions and improvements in interviewer ratings of respondent problems but we had little evidence that the experimental questions demonstrated improvements from the point of view of the interviewer. This finding may not be totally surprising, given that our pretest methods were less effective for detecting interviewer problems and consequently, the revised questions did not serve to eliminate interviewer problems. Details of this research will be included in a chapter of the QDET monograph.

Staff: Jennifer Rothgeb (x4968), Gordon Willis (NCI), Barbara Forsyth (Westat)

K.2. Visual CASI

This project explores the use of images or graphics to help convey the meaning of questions to respondents and to enhance listing in CAPI and Internet instruments.

During FY2002, staff circulated information on the visual design program, Metis/Computas (from Norway) with direct links and mines data from legacy data in most formats.

Staff: Leslie A. Brownrigg (x4995)

K.3. 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. There are two parts to the project. The first is to conduct a "survey" of major survey research organizations to find out exactly how they conduct their cognitive interviews. The second is to conduct an experiment to evaluate the analytic procedures used by different organizations in terms of the completeness of the results obtained and the quality of 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 FY2002, data collection was completed by contractors and staff using three alternative cognitive interview methods. Telephone debriefing of the contractors was also completed. Expert reviews were conducted by staff from the Census Bureau and other Federal agencies to provide independent measures of questionnaire problems. Cognitive interviews were conducted by a contractor to evaluate the questionnaire recommendations that resulted from the application of the three alternative cognitive interview methods. A telephone debriefing of the contractor was also conducted. Experimental treatment reports and expert review reports were coded and the data were keyed into a database. Staff developed a coding scheme to capture details of the three experimental protocols and the way they were administrated during the interviews. They used the scheme to code the tape-recorded interviews. Coded data were also keyed into a database. Data analysis was completed and a paper was written that will be presented at the QDET conference, and included in the monograph of invited papers.

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. Staff developed CATI instrument programming specifications and training materials for the four experimental treatments, and trained survey interviewers at the TTC. Staff developed behavior coding materials and trained coders at the HTC. Behavior coding for this project has been completed.

Staff: Terry DeMaio (x4894), Ashley Landreth, Cathy Keeley, Jennifer Rothgeb, Eleanor Gerber, Tina Arbogast, Lorraine Randall, Jenny Hunter

K.4. Interviewer Refusal Aversion Training

Initial nonresponse rates for continuing Federal household surveys are increasing (e.g., Atrostic, Bates, Burt, Silberstein, & Winters, 1999). Interviewer administered surveys allow interviewers to influence respondent participation because they can generate person-level, customized appeals. Interviewer training, however, is often inadequate in developing skills to effectively engage respondents in this way. Interviewers do not feel prepared in answering respondents' questions, communicating the purpose of the survey, and establishing and maintaining rapport with the respondent (e.g., Doughty et. al., 2000). Based on the concepts of tailoring and maintaining interaction (Groves and Couper, 1998), Groves and McGonagle (in press), have recently examined a theory guided training protocol designed to enhance interviewers' skills in avoiding refusal. Results from testing the protocol omnibus RDD research survey in 2000 were presented to the Census Advisory Committee of Professional Associations Meeting in October, 2001. Research was then expanded to the National Health Interview Survey (NHIS) interviewers. Using NHIS CAPI call history records (July-December, 2001, New York and Dallas Regions), staff found a positive effect for the refusal aversion training regimen in a face-to-face demographic survey environment.

During FY2002, staff worked closely with the Joint Program in Survey Methodology, the Decennial Management and Field Divisions, and NCHS to study and implement the training in a Florida test of the re-engineered NHIS in June 2003 and nationally in January 2004. Preliminary discussions for its application to the Survey of Income and Program Participation and the Consumer Expenditure Survey during 2004 have also begun. Staff continued to collaborate with the Research Triangle Institute, Nielson Media Research, Arbitron, Statistics Canada, and the Survey Research Center of the Institute for Social Research at the University of Michigan during this period. As a result, these institutions are testing or implementing the aversion training regimen in their own surveys. Future analyses are planned to better

understand the contribution of the interviewer on respondent cooperation versus PSU level effects, refine the interviewer evaluation protocol, and to study evaluator inter-coder reliability.

Staff: Eileen O'Brien (x2695), Tom Mayer

K.5. Research Method Comparison for Computer Crime Questions

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

During FY2002, staff analyzed results of respondent debriefing questions administered in the field between September and December 2001. Staff prepared a paper comparing the results of respondent debriefing with previously-conducted cognitive interviews, and presented it at the Annual Meetings of the American Association for Public Opinion Research.

Staff conducted an additional round of cognitive interviews using the revised questions adopted for the field survey. Staff oversaw tape recording of field interviews between September and December 2001, developed behavior coding materials and trained coders in HTC to code the taped interviews. The behavior coding has been completed.

Staff: Kristen Hughes (x8438), Terry DeMaio

K.6. Assisted Living Situations

This project involves conducting research on the varied situations that comprise the assisted living facilities that house senior citizens. Unstructured interviews and possibly cognitive interviews will be used to investigate the categorization of these units either in the census or in household surveys.

During FY2002, staff began this project. We conducted unstructured interviews with staff from federal agencies, national housing organizations, and assisted living facilities.

Staff: Jenny Hunter (x4927)

Research and 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 for the next five years. 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 or have arrangements with outside experts/consultants to broaden their ability to meet all of the potential needs of the Census Bureau. These five-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.

R&D 2002 Contracts

Multiple contracts were awarded in five technical areas: technology services; assessment, planing, and analysis; statistical analysis; methodological research; and minority-focused and special population research.

During FY2002, five new task orders were awarded, 71 modifications were made to the task orders, and seventeen task orders were completed. To date, 101 tasks with a value of over $39 million have been awarded. Currently, there are 27 active tasks. Many custom reports were prepared for various purposes and various program managers that summarized the task orders by type of research, cost, task order and researcher, etc. Monthly progress reports and year end accruals were prepared.

Staff: Ann Dimler (x4996)

R&D 2002 Contracts Task Orders

Staff provided technical management of two task orders under the R&D 2002 Contracts. One task order is with the National Opinion Research Center (NORC) on sampling and estimation methods. The NORC contract was subcontracted to DataMetrics Inc. The other task order is with the Research Triangle Institute on Incentives Use by Federal Agencies.

During FY2002, for the NORC task order, several DataMetrics researchers gave on-site presentations on their research. An all-day meeting was held at the Census Bureau to share research and to provide a forum for discussing the research being conducted on multiple-system estimation methods. The annual meeting of the NORC/DataMetrics researchers and Census Bureau staff was held at the Joint Statistical Meetings in August 2002.

The Research Triangle Institute task order was completed. The research result was the final report: "Use of Incentives in Household Establishment Surveys, Annotated Bibliography."

Staff: Ann Dimler (x4996)

R&D 2007 Contracts

Multiple contracts were awarded in six technical areas: assessment, planning and analysis; statistical analysis; methodological research; survey engineering; subpopulation research; and data analysis and

dissemination.

During FY2002, the process included several stages, and ran from January 2001 to August 2002. The final Request for Proposal was approved and advertised in December 2001. Over fifty written proposals were received from interested contractors in January 2002. Six teams of technical persons with expertise in the relevant technical areas were formed to evaluate the proposals. The evaluation was composed of three parts: evaluation of the written proposals; evaluation of the cost proposals; and evaluation of the oral presentations. This stage was from January to May 2002. The final stage of the process included reaching consensus, documentation and recommendation, Census Bureau and Department of Commerce, EEO approval and subcontracting plan approval and award of contracts. This stage was from June through August, 2002. Final awards and approvals required documentation of every step of the process. In addition to documentation, numerous meetings were scheduled and held to promote a thorough understanding of the process, to provide the justifications for award and to get the necessary approvals.

The Census Bureau made awards for the R&D 2007 contracts on September 30, 2002, to twenty-five contractors.

Assessment, Planning and Analysis: This technical area focuses on the assessment, planning, and analysis of Census Bureau programs and the way the Census Bureau conducts business.

Data Analysis and Dissemination: This technical area focuses on subject matter research. Examples included report guidelines; demographic analysis; population estimates and projections; data analysis; data dissemination and presentation; survey indicators/measures; cross-tabulations; policy; and time series analysis.

Methodological Research: This technical area focuses on survey methodology such as questionnaire design, public cooperation and participation, and improvement of enumeration methods and procedures for surveys of individuals, households, and establishments. Examples included questionnaire design, response burden studies, cooperation and participation, privacy issues, data collection techniques, technology use, and human factors and usability research.

Statistical Analysis: This technical area focuses on research and development of statistical theory and methodologies for the design and production of estimates for censuses and surveys. Examples included sample design and estimation, missing data techniques, total error models, coverage measurement, disclosure avoidance, experiments and simulation studies, administrative records research, statistical computing, and other statistical fields such as census and survey evaluations, experimental design, regression analysis, nonparametric measures, statistical modeling, survey methodology, variance estimation for complex estimators, quality assurance, stochastic processes, statistical graphics, survival analysis and latent class analysis, statistical standards, evaluations, small area estimation, and time series analysis.

Subpopulations Research: This technical area focuses on research and development of methods for the collection of information for racial and ethnic groups and other populations requiring special enumeration procedures. Examples included questionnaire design, enumerations modes, administrative records, hard-to-enumerate methods, ethnography, and non-English language data collection instruments.

Survey Engineering: This technical area focuses on the application of scientific and mathematical principles to practice survey problems such as the design and operation of efficient and economical surveys, machines, processes, and systems. Examples included MAF/TIGER enhancements, data collection and capture, data tabulation and processing, systems design and integration, software enhancements, electronic printing and release, information technology accessibility, statistical programming, data security, other services such as training on the use and application of high performance computers (standard course offerings or specially designed programs), technical assistance for geo-processing, computer science, computer operations, statistics, configuration management software evaluation, or other relevant fields in the high performance computing environment, development of software prototypes to address the areas discussed above, and computer services such as CPU time and telecommunications for work done in this area.

Staff: Ann Dimler (x4996), Tina Arbogast

Research Assistance

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

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

3. PUBLICATIONS

3.1 JOURNAL ARTICLES, PUBLICATIONS

Aston, J., Cunningham, V., Asselin, M., Hammers, A., Evans, A., and Gunn, R. (In Press). "Position Emission Tomography Partial Volume Correction: Estimation and Algorithms, *"Journal of Cerebral Blood Flow and Metabolism.*

Botkin, J., Smith, K., Croyle, R., Baty, B., Wylie, J., Dutson, D., Chan, A., Hamann, H., Lerman, C., McDonald, J., Venne, V., Ward, J., and Lyon, E., (In Press). "Genetic Testing for a BRCA1 Mutation: Prophylactic Surgery and Screening Behavior in Women 2 years Post Testing." *American Journal of Medical Genetics.*

Causey, B.D. (2002). "Parametric Estimation of the Number of Classes in a Population," *Journal of Applied Statistics, Vol. 29, No. 6,* 925-934.

Findley, D., Poetscher, B., and Wei, C-Z. (In Press). "Modeling of Time Series Arrays by Multistep Prediction or Likelihood Methods," *Journal of Econometrics.*

Hess, J., Moore, J., Pascale, J., Rothgeb, J., and Keeley, C. (2001). "The Effects of Person-Level Versus Household-Level Questionnaire Design on Survey Estimates and Data Quality," *Public Opinion Quarterly, Vol. 65, No. 4,* 574-584.

Legler, J., Breen, N., Meissner, H., Malec, D., and Coyne, C. (2002). "Predicting Patterns of Mammography Use: A Geographical Perspective on National Needs for Intervention Research," *Health Services Research., 37, 4*, 929-947.

Martin, D.E. (In Press). "An Algorithm to Compute the Probability of a Run in a Binary Fourth-Order Markovian Trial," *Computer and Operations Research..*

Murphy, E., Marquis, K., Nichols, E., Kennedy, K., and Mingay, D. (2001). "Refining Electronic Data-Collection Instruments and Data-Collection Tools through Usabilitiy Testing," *Research in Official Statistics, Vol 4, #2,* 23-33.

Press, James, S. and Marquis, K.H. (2001). "Bayesian Estimation in a U.S. Census Bureau Survey of Income Recall Using Respondent-Generated Intervals," *Research in Official Statistics, 1,* 151-168.

Strange, M., Ezzell, J., and Dajani, A. (2002). "Initial Public Offerings: The Case of Health Maintenance Organizations (HMOs)," *International Business & Economic Research Journal, Vol. 1, No. 6,* 85-95.

Thibaudeau, Y. (In Press). "Model Explicit Item Imputation for Demographic Categories," *Survey Methodology.*

3.2 BOOKS/BOOK CHAPTERS

Doyle, P., Lane, J., Theeuwes, J., and Zayatz, L. (Eds). (2001). *"Confidentiality, Disclosure, and Data Access: Theory and Practical Application for Statistical Agencies,* Amsterdam, The Netherlands: Elsevier Science BV.

Gerber, E. (2001). "The Privacy Context of Survey Response: An Ethnographic Account," in Doyle, Lane, Theeuwes, and Zayatz (Eds), *Confidentiality, Disclosure, and Data Access: Theory and Practical Application for Statistical Agencies,* Amsterdam, The Netherlands: Elsevier Science BV., 371-294.

Yancey, W.E., Winkler, W.E., and Creecy, R.H. (2002). "Disclosure Risk Assessment in Perturbative Microdata Protection," in *Inference Control in Statistical Databases,* (J. Domingo-Ferrer, ed.) *Lecture Notes in Computer Science,* New York: Springer, 135-152.

Zayatz, L. (2002). "SDC in the 2000 U.S. Decennial Census," in *Inference Control in Statistical Databases* (J. Domingo-Ferrer, ed.), *Lecture Notes in Computer Science,* New York: Springer, 193-202.

3.3 PROCEEDINGS PAPERS

Accompanying Measure to Research and Development in Official Statistics (AMRAD), Luxembourg, December 13-14, 2001.

- Zayatz, Laura (To Appear). "SDC in the 2000 U.S. Decennial Census, *"Proceedings of Statistical Data Protection, 2001, Eurostat.*

Quest, 2001, Washington, D.C., October 24-25, 2001.

- DeMaio,T. and Landreth, A. (2001). "Experimentation with Alternative Methods of Conducting Cognitive Interviews," *Proceedings of QUEST 2001,* 64-66.

American Association for Public Opinion Research, Montreal, Canada, May 17-20, 2001.

- Nichols, E., Marquis, K., and Hoffman III, R. (2001) "The Effect of Motivational Messaging on Mode

Choice and Response Rates in the Library Media Center Survey," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.

- Pascale, J. (2001). "Measuring Private and Public Health Coverage: Results from a Split-Ballot Experiment on Order Effects," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Redline, C., Lankford, D., and Christopher, P. (2001). "Eye-Movement Analysis: A New Tool for Evaluating the Design of Visually Administered Instruments," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Rothgeb, J., Willis, G., and Forsyth, B. (2001). "Questionnaire Pretesting Methods: Do Different Techniques and Different Organizations Produce Similar Results," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Willimack, Diane and Nichols, E. (2001). "Building an Alternative Response Process Model for Business Surveys," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.

2001 Joint Statistical Meetings (American Statistical Association), Atlanta, GA, August 5-9, 2001.

- Chen, B.-C., Winkler, W., and Hemmig, R. (2001). "Using the DISCRETE Edit System for ACS Surveys," *2001 Proceedings of the Section on Computing (CD-ROM),* American Statistical Association.
- Doyle, P. and Moore, J. 2001. "Methods Panel to Improve Income Measurement: Analysis of an Experimental SIPP Instrument," *2001 Proceedings of the Section on Government CD-ROM),* American Statistical Association.
- Farooque, G. and Findley, D. (2001). "Comparing the Automatic ARIMA Model Selection Procedures of TRAMO and X-12-ARIMA 0.3," *2001 Proceedings of the Section on Business and Economics (CD-ROM),* American Statistical Association.
- Findley, D. (2001), "Asymptotic Stationarity Properties of Out-of -Sample Forecast Errors of Misspecified RegARIMA Models," *2001 Proceedings of the Section on Business and Economics (CD-ROM),* American Statistical Association.
- Griffen, R. and Malec, D., (2001). "Census 2000 Accuracy and Coverage Evaluation: Assessment of Synthetic Assumption," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Hawala, S. (2001). "Enhancing the '100,000 Rule' on the Variation of the Per Cent of Uniques in a Microdata Sample and the Geographic Area Size Identified on the File," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Hood, C. and Findley, D. (2001). "Comparing Direct and Indirect Seasonal Adjustments of Aggregate Series," *2001 Proceedings of the Section on Business and Economics (CD-ROM),* American Statistical Association.
- Ikeda, M. and Cantwell, P. (2001). "Missing Data Procedures for the Census 2000 Accuracy and Coverage Evaluation Sample," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Kim, J. and Winkler, W. (2001). "Multiplicative Noise for Masking Continuous Data," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Kirk, M., Buckles, E. Mims, W., Appel, M., and Johnson, P. (2001). "Preliminary Results from the Census 2000 Industry and Occupation Coding," *2001 Proceedings of the Section on Government (CD-ROM),* American Statistical Association.
- Klein-Griffiths, J. (2001). "Household Income Screening Procedures in the Survey of Income and Program Participation," *2001 Proceedings of the Section on Government (CD-ROM),* American Statistical Association.
- Martin, E., de la Puente, M., and Bennett, C. (2001). "The Effects of Questionnaire and Content Changes on Responses to Race and Hispanic Origin Items: Results of a Replication of the 1990 Census Short Form in Census 2000," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.

- Martin, D. and Findley, D. (2001). "Spectral Properties of Linear Concurrent and Symmetric Seasonal Adjustment Filters of SEATS and X-11/12 ARIMA for Short and Moderate-Length Time Series," *2001*

Proceedings of the Section on Business and Economics (CD-ROM), American Statistical Association.

- Massell, P. (2001). "Using Linear Programming for Cell Suppression in Statistical Tables: Theory and Practice," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Mayer, T. and O'Brien, E. (2001), "Interviewer Refusal Aversion Training to Increase Survey Participation," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Mulry, M. 2001. Discussion of Session on Census Methodology, *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Nguyen, D., Nganha, T., and Cantwell, P. (2001). "Missing Data Results for the Census 2000 Accuracy and Coverage Evaluation," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- O'Brien, E. (2001). "Application of Cognitive Methods to an Establishment Survey: A Demonstration Using the Current Employment Statistics Survey," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Pascale, J. (2001). "Labor Force Participation in the Survey of Income and Program Participation," *2001 Proceedings of the Section on Government (CD-ROM),* American Statistical Association.
- Redline, C., Dillman, D., Dajani, A., and Scaggs, M.A. (2001). "The Effects of Altering the Design of Branching Instructions on Navigational Performance in Census 2000," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Rowland, S. and Zayatz, L. (2001). "Automating Access with Confidentiality Protection: The American FactFinder," *2001 Proceedings of the Section on Social Statistics, (CD-ROM)* American Statistical Association.
- Slud, E. 2001. "Order Selection, Random Effects, and Multilevel Predictors in Modeling Decennial Census Response," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Steel, P. and Sperling, J. (2001). "The Impact of Multiple Geographies and Geographic Detail on Disclosure Risk: Interactions between Census Tract and ZIP Code Tabulation Geography," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Thibaudeau, Y. (2001). Discussion of Databases of Interest to Environment and Natural Resource Researchers, *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Williams, T. (2001). "Flexible Matching Imputation: Combining Hot-Deck Imputation with Model-Based Methodology," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Winkler, W. and Chen, B-C. (2001). "Extending the Fellegi-Holt Model of Statistical Editing," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.
- Zayatz, L. and Mackun, P. (2001). "Disclosure Limitation for Census 2000 Public Use Microdata Files," *2001 Proceedings of the Section on Social Statistics, (CD-ROM)* American Statistical Association.
- Zelenak, M. F., McGrath, D., Nguyen, N., and Cantwell, P. (2001). "Missing Data for the Census 2000 Accuracy and Coverage Evaluation," *2001 Proceedings of the Section on Survey Research Methods (CD-ROM),* American Statistical Association.

3.4 STATISTICAL RESEARCH DIVISION RESEARCH REPORTS

RR (**Survey Methodology #2001-01**), Jeffrey Moore, "Asset Ownership Screening Procedures in the Survey of Income and Program Participation (SIPP Methods Panel)," October 3, 2001.

RR (Survey Methodology #2001-02), Tom Mayer, "Interviewer Attitudes about Privacy and Confidentiality," November 20, 2001.

RR (Survey Methodology #2001-03), Jeffrey Moore and Laura Loomis, "Using Alternate Question Strategies to Reduce

Income Nonresponse," November 29, 2001.

RR (Statistics #2002-01), William Yancey, William Winkler, and Robert Creecy, "Disclosure Risk Assessment in Perturbative Microdata Protection," January 31, 2002.

RR (Statistics #2002-02), William Winkler and Bor-Chung Chen, "Extending the Fellegi-Holt Model of Statistical Data Editing," February 1, 2002.

RR (Survey Methodology #2002-01), Thomas Mayer, "Privacy and Confidentiality Research at the U.S. Census Bureau: Recommendations Based on a Review of the Literature," February 7, 2002.

RR (Survey Methodology #2002-02), Cleo Redline, Don Dillman, Aref Dajani, and Mary Ann Scaggs, "The Effects of Altering the Design of Branching Instructions on Navigational Performance in Census 2000," February 11, 2002.

RR (Survey Methodology #2002-03), Jeffrey Moore and Laureen Moyer," Questionnaire Design Effects on Interview Outcomes," February 15, 2002.

RR (Survey Methodology #2002-04), Jeffrey Moore and Laureen Moyer, "ACS/CATI Person-Based Field Experiment," February 19, 2002.

RR (Computing #2002-01), William Yancey, "BigMatch: A Program for Extracting Probable Matches from a Large File for Record Linkage," March 6, 2002.

RR (Survey Methodology #2002-05), Jennifer Hess, Jeffrey Moore, Joanne Pascale, Jennifer Rothgeb, and Catherine Keeley, "The Effects of Person-level vs. Household-level Questionnaire Design on Survey Estimates and Data Quality," March 11, 2002.

RR (Survey Methodology #2002-06), Jeffrey Moore and Laura Loomis, "Reducing Income Nonresponse in a Topic-Based Interview," March 11, 2002.

RR (Statistics #2002-03), William Yancey, "Working Papers for Mixture Model Additive Noise for Microdata Masking," May 28, 2002.

RR (Computing #2002-02), Bor-Chung Chen and William Winkler, "An Efficient Formulation of Age Comparison in the DISCRETE Edit System," June 10, 2002.

RR (Statistics #2002-04), Bill Bell and Don Martin, "Computation of Asymmetric Signal Extraction Filters and Mean Squared Error for ARIMA Component Models," August 26, 2002.

3.5 OTHER REPORTS

- Rick Griffin and Don Malec (2001). "ESCAP II: Sensitivity Analysis for the Assessment of the Synthetic Assumption." *Executive Steering Committee for A.C.E. Policy II, Report 23.*

- Don Malec (2001). "Unit Level Models for Small Area Estimation: Applications to Census Adjustment of Small Areas and Small Area Estimation for the American Community Survey." Prepared for: *Census Advisory Committee of Professional Associations Meeting,* October 18-19, 2001.

- Elizabeth Nichols, "Results from the Applied Behavior Analysis of American Community Survey Nonrespondents, *Human-Computer Interaction Memorandum Series #50*, August 7, 2001.

- David Mingay and Gautam Rao, "Report on Usability Testing of Census Bureau's DynaMaps CD-ROM Product," *Human-Computer Interaction Memorandum Series #51,* October 16, 2001.
- Erica Olmsted, "A Usability Evaluation of the 2002 Foreign-Born Web Domain," *Human-Computer Interaction Memorandum Series* #52, May 17, 2002.

- Brian Greenbaum, "Usability, Evaluation of the Proposed Information Technology Directorate Intranet Web Site," *Human-Computer Interaction Memorandum Series* #53, June 7, 2002.

- Amy Thruston and Dick Horst, UserWorks, Inc., "Heuristic Usability Review of the Web-enabled Manufacturers' Shipments, Inventories, and Orders (M3) Survey," *Human-Computer Interaction Memorandum Series* #54. June 6, 2002.

- Bill LaPlant, "Setting Standards for IT Accommodation & Universal Access," *Third Quarter Issue, Newsletter of the Rehabilitation Engineering and Assistive Society of North America.*

- Erica Olmsted and Amy Thruston (UserWorks, Inc.), "A Usability Evaluation of Field Directorate Intranet Site," *Human-Computer Interaction Memorandum Series #54a,* June 6, 2002.

4. TALKS AND PRESENTATIONS

Association of Public Data Users Conference, Arlington, VA, October 15-17, 2001.

- Laura Zayatz, "Disclosure Limitation Practices at the Census Bureau."

Census Advisory Committee of Professional Associations, Washington DC, October 18-19, 2001.

- Donald Malec, "Unit Level Models for Small Area Estimation: Application to Census Adjustment of Small Areas and Small Area Estimation for the American Community Survey."
- Tom Mayer, "Interviewer Refusal Aversion Training."

Telecommunication Rehabilitative Engineering Research Center State of the Science Conference, Washington, DC, October 19, 2001.

- Bill LaPlant, "The Impact of Standards on 'Compatibility'."

American Public Health Association, Atlanta, GA, October 21-25, 2001.

- Joanne Pascale, "Health Insurance Measurement Methodologies: A Data Quality Assessment."

Quest 2001 Conference, Washington, DC, October 24-25, 2001.

- Terry DeMaio, "Experimentation with Alternative Methods of Conducting Cognitive Interviews."
- Jennifer Rothgeb, "Challenges and Strategies in Gaining Acceptance of Research Results from Cognitive Questionnaire Testing."

Department of Statistics and Computer Science, Virginia Tech University, Blacksburg, VA, November 1, 2001.

- Bill Winkler, "Record Linkage and Machine Learning."

ASA/SRM SIPP Advisory Group, Alexandria, VA, November 9, 2001.

- Pat Doyle and Jeff Moore, "SIPP Methods Panel: Results and Future Plans."

Federal Committee on Statistical Methodology Research Conference, Arlington, VA, November 14, 2001.

- Don Malec, "Small Area Estimates from the American Community Survey Using a Housing Unit Model." (Available on the FCSM Website).
- Elizabeth Murphy, Elizabeth Nichols, A. Anderson, M.D. Harley, and K. Pressley, "Building Usability into Electronic Data-Collection Forms for Economic Censuses and Surveys," *Statistical Policy Working Paper 34, Part 4 of 5,* 113-122.
- Elizabeth Nichols and Diane Willimack, "Balancing Confidentiality and Burden Concerns in Censuses and Surveys of Large Businesses," *Statistical Policy Working Paper 34, Part 3 of 5,* 20-29.

American Anthropological Association, Washington, DC, December 1, 2001.

- Leslie Brownrigg, Chair and Discussant, "On the Move: The Social Organization of Highly Mobile People."
- Laurie Schwede and Anna Chan, "Socio-Cultural Factors and Their Possible Effects on the Census."

Statistics Netherlands, Voorburg, Netherlands, December 10, 2001.

- Bill Winkler, "Methods for Using and Analyzing Text."

Statistical Disclosure Protection 2001, Luxembourg, December 13-14, 2001.

- Phil Steel, "SDC in the 2000 U.S. Decennial Census."

Statistical Data Confidentiality: From Theory to Application, Luxembourg, December 13, 2001.

- Bill Winkler, "Disclosure Risk in Perturbative Microdata Protection."

Council of Professional Associations on Federal Statistics, Washington, DC, December 14, 2001.

- Laura Zayatz, "Disclosure Limitation Practices at the Census Bureau."

Bank of Thailand, Bangkok, Thailand, January 22-24, 2002.

- David Findley, Course on "Seasonal Adjustment with X-12-ARIMA."

Computational Sciences and Statistics Seminar, George Mason University, Fairfax, VA, January 25, 2002.

- William Winkler, "Machine Learning Methods for Text Classification."

Third International Symposium on Frontiers of Time Series Modeling: Modeling Seasonality and Periodicity, Tokyo, Japan, January 31-February 1, 2002.

- David Findley, "Spectral Properties of Linear Concurrent and Symmetric Seasonal Adjustment Filters of SEATS and X-11/12-ARIMA for Short and Moderate Length Time Series."
- Brian Monsell, "An Update on the Development of the X-12-ARIMA Seasonal Adjustment Program."

Statistics Department Seminar, University of California, Riverside, CA, February 12, 2002.

- Bill Winkler, "Disclosure Risk Assessment in Perturbative Microdata Protection via Record Linkage."

Response Rate Summit, Crystal City, VA, February 21-22-2002.

- Eileen O'Brien, "Refusal Aversion Training Research at the Census Bureau."

Statistics Seminar, Department of Mathematics and Statistics, University of Maryland, Baltimore County, MD, March 1, 2002.

- Bill Winkler, "Machine Learning Methods for Using and Analyzing Text."

25th Annual Meeting of the Society of Ethnobiology, Storrs, CT, March 6-9, 2002.

- Leslie Brownrigg, "Inca Rank Terms."

Statistical Engineering Division Seminar, National Institute of Standards and Technology, Gaithersburg, MD, February 13, 2002.

- Don Malec, "A Closer Look at Combining Data Among a Small Number of Binomial Experiments."

Capital PC User Group and the Washington Institute for Operations Research and the Management Sciences, Washington, DC, March 13, 2002.

- Brian Monsell, a Presentation and Demonstration of the R Statistical Package.

JPSM Brown Bag, University of Maryland, College Park, MD, March 14, 2002.

- Don Malec, "The American Community Survey as a Tool for Small Area Estimation."

Center for Survey Research and Methodology (ZUMA), Mannheim, Germany, March 14-15, 2002.

- Manuel de la Puente, "Translation Guidelines and Best Practices at the U.S. Census Bureau: Needs, Research, and Development."
- Eleanor Gerber, "The Ethnographic Context of Surveys."

International Biometric Society, Eastern North American Region Spring Meeting, Arlington, VA, March 18, 2002.

- Donald Malec (Invited paper), "Complete Bayesian Small Area Estimates from the American Community Survey Using a Housing Unit Model.

Association of American Geographers Annual Meeting, Los Angeles, CA, March 19-23, 2002.

- Manual de la Puente, "Border Communities on the US/Mexico Border: Lessons Learned from Census 2000."

Statistics Department Seminar, Virginia Tech University, Blacksburg, VA, March 21, 2002.

- David Findley, "Selecting Between Non-Nested Statistical Models: Examples and Methods."

Sri Lanka Department of Census and Surveys, Combo, Sri Lanka, February 5-8, 2002.

- Brian Monsell, Course on "Seasonal Adjustment with X-12-ARIMA."

MINI STATFest: A Conference for Undergraduates, Hampton University, Hampton, VA, April 13, 2002.

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

Census Advisory Committee of Professional Associations, Arlington, VA, April 18-19, 2002.

- Manuel de la Puente, "Overview of the Census Bureau's Language Program."

Computational Sciences and Statistics Seminar, George Mason University, Fairfax, VA, May 3, 2002.

- Bill Winkler, "The Fellegi-Holt Model of Statistical Data Editing: Computational Algorithms and Research Problems."

Statistics Department, University of Maryland, College Park, MD, May 9, 2002.

- John Aston, "Statistical Methods for Population Inference in Spatio-Temporal Data Sets."

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

- Eleanor Gerber, "Images of the Survey: Sociocultural Factors in Respondent Behavior."
- Kristen Hughes and Terry DeMaio, "Does This Question Work? Evaluating Cognitive Interview Results Using Respondent Debriefing Questions."
- Joanne Pascale, "Data Quality of Health Insurance Surveys."
- Joanne Pascale, "Measuring and Estimating Health Insurance Status."
- Anna Chan, Julia Klein Griffiths, Jeff Moore, Joanne Pascale, Nancy Bates, and Pat Doyle, roundtable discussion, "Improving Income Measurement."
- Theresa Leslie, D. Raglin, and Laurie Schwede, "Understanding the Effects of Interviewer Behavior in the Collection of Race Data."
- Manuel de la Puente, "A Blueprint for Obtaining High Quality Data from Non-English Speaking Households: Translation Guidelines, Pretesting Standards, and Related Research."

Field Directors' Conference, Clearwater, FL, May 19-22, 2002.

- Larry Malakhoff, "Selected RM&IE Results."

U.S. Census Bureau/Statistics Canada Interchange, Ottawa, Canada, June 27, 2002.

- Tom Mayer, "The Use of Cognitive Research in Developing Privacy and Confidentiality Notices."
- Brian Monsell, Discussion, "Issues Concerning the Conversion of Statistics Canada's Seasonal Adjustment Processing from X-11-ARIMA/2000 to X-12-ARIMA."

XV World Congress of Sociology, Brisbane, Australia, July 7-13, 2002.

- Jennifer Rothgeb, "Questionnaire Pretesting Methods: Do Different Techniques and Different Organizations Produce Similar Results."

50th Anniversary Meeting of the Society for Industrial and Applied Mathematics, Philadelphia, PA, July 8-12, 2002.

- Paul Massell, "Protecting Suppressed Cells in Statistical Tables: An Optimization Problem."

Washington Statistical Society Seminar, Washington, DC, July 16, 2002.

- Paul Massell, discussant, "Confidentiality Audit on Suppressed Entries in Multi-Dimensional Contingency Tables."

Neuroreceptor Mapping 2002, Oxford, UK, July 19-21, 2002.

- John Aston, "Partial Volume Correction: A Recipe."
- John Aston, "Image Space Variance from Wavelet Transforms."

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

- John Aston, "Statistical Based PET Partial Volume Correction."
- Pat Cantwell, "Accounting for Imputation When Estimating Variances in the Economic Surveys at the Census Bureau."

- Sam Hawala, "Simulation Study of the Effectiveness of Masking Microdata with Mixtures of Multivariate Normal Distributions."
- Jay Kim, Danielle Corteville, and Patrick Flanagan, "Maximizing Retention of Primary Sampling Units (PSUs) in a Two-PSU Per Stratum Design."
- Donald Martin, David Findley, and Kellie Wills, "On a Generalization of the Box-Jenkins Airline Model."
- Paul Massell, "Optimization Models and Programs for Cell Suppression in Statistical Tables."
- Eileen O'Brien, "Interviewer Training to Increase Survey Participation."
- Laurie Schwede, Theresa Leslie, and Deborah Griffin, "Interviewers' Reported Behaviors in Collecting Race and Hispanic Origin Data."
- Eric Slud, "Model-assisted Weighting for Surveys with Multiple Response Mode."
- Thuy Trang Nguyen, William Bell, and James Gomish, "Investigating Model-Based Time Series Methods to Improve Unadjusted and Seasonally Adjusted Estimates from Monthly Value of Construction Put-in-Place Surveys."
- Todd Williams, "Flexible Matching Imputation in the Manufactured Homes Survey."
- Laura Zayatz, "Have You Done Enough to Disclose-Proof Your Public Use File?"

International Conference on Improving Surveys, Copenhagen, Denmark, August, 2002.

- Joanne Pascale, "A Quantitative and Qualitative Assessment of the Data Quality of Health Insurance Measurement Methodologies."

Statistics Seminar Series, George Mason University Fairfax, VA, August 30, 2002.

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

Statistics Department Seminar, University of Tennessee, Knoxville, TN, September 17, 2002.

- Phil Steel, "The Impact of Geographic Detail on Disclosure Risk for Two Commonly Used Statistical Geographies."

Statistics Colloquia, George Washington University Washington, DC, September 20, 2002.

- John Aston, "Partial Volume Correction for Neuroimaging using Tensor Based Statistical Algorithms."

Washington Statistical Society Seminar, Washington, DC, September 26, 2002.

- John Aston and Siem Jan Koopman, "Robust Seasonal Adjustment Using Heavy-Tailed Distributions."

5. STATISTICAL RESEARCH DIVISION SEMINAR SERIES

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

Lawrence H. Cox, National Center for Health Statistics, "Properties of Multi-Dimensional Statistical Tables and the Problems They Pose for Statistical Disclosure Limitation," October 31, 2001.

John Aston, Imperial College (London) and McGill University (Montreal), "Statistical Methods for Population Inference in Spatio-Temporal Data Sets," November 19, 2001.

Nathaniel Schenker, National Center for Health Statistics, "From Single-Race Reporting to Multiple-Race Reporting: Using Imputation Methods to Bridge the Transition," November 20, 2001.

James Filliben, National Institute of Standards and Technology, "An Exploratory Data Analysis Retrospective," November 27, 2001.

Jason Schacter, Population Division, Bureau of the Census, "What We Know about Residential Mobility from Surveys," Leslie Brownrigg, SRD, Bureau of the Census, "The General Method of the Ethnographic Social Network Tracing Evaluation," L Marcelin, School of Medicine, University of Miami, "Haitian Migrant and Seasonal Farm Workers Based in South Florida," Alicia Chaviara-Prado, University of Illinois, "Mexican Former Migrant Workers in the Midwest," Nancy Murray, Hatpin Communications, "Young Adult Seasonal Park Workers," Kathi Kitner, National Marine Fisheries Service, NOAA, "Commercial Fishermen Docking on the South Atlantic," Brian Gilley, North Central College, Naperville, IL, "An American Indian Men's Society in Oklahoma," Susan Lobo, Intertribal Friendship House, University of Arizona, "Homeless American Indians in the San Francisco Bay Area," November 28, 2001.

Milo Schield, Augsburg College, Minneapolis, MN, "Student Difficulties in Reading Census Tables of Rates and Percentages," December 3, 2001.

Doug Wakefield, U.S. Access Board's Office of Technical and Information Services, "Taking the Guesswork Out of 508," December 18, 2001.

Jae-Kwang Kim, WESTAT, "Replication Variance Estimation for Fractional Hot Deck Imputation and Nearest Neighbor Imputation," January 15, 2002.

Yuling Pan, Georgetown University, "Politeness and Cross-cultural Communication," January 23, 2002.

William Winkler, SRD, Census Bureau, "Record Linkage and Machine Learning," January 29, 2002.

Patricia Hu, Bruce Peterson, and Demin Xiong, Oak Ridge National Laboratory, "GIS and Image Based Approaches to TIGER Enhancement," February 6, 2002.

Brandon Whitcher, National Center for Atmospheric Research, Boulder, CO, "Wavelet-Based Estimation Procedures for Seasonal Long-Memory Models," February 13, 2002.

William Winkler, SRD, Census Bureau, "Masking and Re-Identification Methods for Public-Use Microdata," February 21, 2002.

William Winkler, SRD, Census Bureau, "Overview of the Fellegi-Holt Model of Statistical Data Editing: Current Methods and Research Problems," March 13, 2002.

William Winkler, SRD, Census Bureau, "Machine Learning Methods for Text Classification," March 27, 2002.

Chris North, Virginia Tech University and Dave Desjardins, Census Bureau (former employee) "SNAP EDA Mapping Software," April 8, 2002.

Michelle AuCoin McGuire, University of Chicago, "A Linguistic Analysis of Early African American English," April 15, 2002.

Kent Norman, University of Maryland at College Park, "Web-Based Surveys: Questions, Answers, and Designs," April 16, 2002.

Yuling Pan, Georgetown University, "Sociolinguistic Study of Professional Communication Across Cultures," April 16, 2002.

P. Lahiri, University of Nebraska, Lincoln & University of Maryland, College Park, "A Weighted Jackknife Method for the Fay-Herriot Model with an Application in the SAIPE Program," April 23, 2002.

Li-chiung Yang, Georgetown University, "Interpreting Contextual Meaning in Spoken Discourse," April 24, 2002.

Jurgen Doornik, Nuffield College, "A Demonstration of the GiveWin Version of X-12-ARIMA, A Windows Version of X-12-ARIMA," April 30, 2002.

Jonathan Lazer, Towson University, "YOU ARE HERE: Information and Architecture and Web Navigation," May 8, 2002.

Kirsh Muralidhar, University of Kentucky and Rathindra Sarathy, Oklahoma State University, "An 'Optimal' Data Swapping Procedure," May 20, 2002.

Laura Zayatz, SRD, Census Bureau, "Statistical Disclosure Control for Census 2000," May 21, 2002.

Patrick Cantwell, SRD, Census Bureau, "Some Recent Developments in Composite Estimation in Sample Surveys," May 22, 2002.

John T. Wulu, Jr., Department of Health and Human Services, "Regression Analysis of Count Data," May 23, 2002.

Dean Judson, PRED, Census Bureau, "Multiple System Estimation with Erroneous Enumerations: An Experimental Bayesian Latent Class Approach," May 29, 2002.

Ellen Hertz, National Highway Traffic Safety Administration, "National Estimates of an Audited Variable with a Large Number of Missing Values," May 30, 2002.

Ben Shneiderman, University of Maryland, College Park, "Leonardo's Laptop: Human Needs and the New Computing Technologies," June 20, 2002.

Tony Hak, Faculteit Bedrijfskunde/Rotterdam School of Management, Erasmus University, Rotterdam, The Netherlands, "The Response Process in Establishment Surveys," July 10, 2002.

Jerry Maples, Department of Statistics, Pennsylvania State University, "A Semiparametric Mixed Effects Model with Marginal Means Structure," July 24, 2002.

Danny Pfeffermann, Hebrew University and University of Southampton, "Bootstrap Approximation to Prediction MSE for State-Space Models with Estimated Parameters, August 7, 2002.

Thomas Ryan, Consultant, "Parameter Estimation of Logistic Regression - Not an Easy Matter," August 19, 2002.

Kenneth R. W. Brewer, Australia National University, "Combined Survey Sampling Inference: Compromise or Consummation," Part 1, "The Why and the How," Part 2 "Some Simple Variance Formulas and Estimators," August 20, 2002.

6. PERSONNEL ITEMS

6.1 HONORS/AWARDS/SPECIAL RECOGNITION

Bronze Medal Award, Bureau of the Census

- *Elizabeth Huang,* for survey design research and application of innovative statistical methodology. Her work encompassed multiple frames research included administrative records, decennial census undercount adjustment, composite estimation and empirical Bayes estimation with regard to demographic surveys, and nonresponse adjustment for economic area surveys.

Customer Service Award

- *Aref Dajani* - for effectively connecting with others at the Census Bureau while providing outstanding statistical and computing assistance on many projects for variance estimation and analysis of data from experiments, tests, and evaluations (awarded October 8, 2002).

6.2 SIGNIFICANT SERVICE TO PROFESSION

John Aston
- Refereed a paper for *The American Statistician.*
- Refereed a book for *Wiley Probability and Statistics Section.*

Pat Cantwell
- Associate Editor, *Journal of Official Statistics.*
- Program Chair, *Section on Survey Research Methods, American Statistical Association.*
- Refereed a paper for the *Journal of the American Statistical Association.*

Bev Causey
- Refereed a paper for *The American Statistician.*

Manuel de la Puente
- President-Elect, District of Columbia Sociological Society.
- Member, the American Sociological Association Working Group on the Definition of Race.
- Refereed a paper for *Latino Studies Journal.*
- Invited Discussant, 2002 Meeting of American Association for Public Opinion Research.

Terry DeMaio
- Refereed papers for *Public Opinion Quarterly.*
- Secretary, Washington/Baltimore Chapter, American Association for Public Opinion Research.

David Findley
- Refereed a paper for *Taiwan Economic Forecasting and Policy* and *Economic Theory.*
- Chair, Business and Economic Statistics Section, ASA.
- Evaluator of a Candidate for Promotion to Full Professor, Department of Statistics, Colorado State University.

Sam Hawala
- Member, Confidentiality and Data Access Committee.
- Member, National Center for Education Statistics Disclosure Review Board.

Bill LaPlant
- Member, Federal Universal Access Working Group.
- Member, Federal CIO Council XML Working Group.

Don Malec

- Refereed a paper and manuscripts for *The American Statistician, Survey Methodology, Communications in Statistics, Statistics in Medicine.*

Donald Martin

- Refereed papers for *Computers & Operations Research, The American Statistician.*
- Member, Committee on Minorities in Statistics of ASA.

Paul Massell

- Member, Confidentiality and Data Access Committee.
- Refereed a paper for the *Journal of Official Statistics.*

Brian Monsell

- Secretary-Treasurer, Business and Economic Statistics Section of ASA.
- Online Assistant Editor, Business and Economic Statistics Section of ASA.

Mary Mulry

- Associate Editor, *The American Statistician.*
- Chair, Fellows Committee, Survey Research Methods Section of ASA.
- Refereed a paper for *Journal of Official Statistics.*

Eileen O'Brien

- Member, American Statistical Association's Survey Review Committee (2001-2003).
- Member, Interagency Household Survey Nonresponse Group.
- Member, Interagency Group on Establishment Nonresponse.

Joanne Pascale

- Editor, *Apha Statistics Section's Winter 2001 Newsletter,* issued in December, 2001.

Jennifer Rothgeb

- Produced *Proceedings of the 2001 QUEST International Workshop.*
- Chair, Organizing Committee, International Conference on Questionnaire Development, Evaluation, and Testing Methods (QDET).

Laurie Schwede

- Session Chair, "Theory in Anthropology," American Anthropological Association Annual Meeting.
- Reviewed a National Science Foundation proposal.
- Liaison, Testing and Experimentation Program Steering Committee.

David Smith

- Refereed a paper for *The American Statistician.*.

Phil Steel

- Member, Confidentiality and Data Access Committee.
- Member, American Statistical Association's Committee on Privacy and Confidentiality.

Bill Winkler

- Member, Organizing Committee, Work Session on Statistical Data Editing, U.N. Economic Commission for Europe, Helsinki, Finland.
- Refereed papers for *Journal of Official Statistics, Survey Methodology, Information Sciences,* and *Journal of the American Statistical Association.*
- Member, Eurostat-sponsored Data-Clean Project.
- Lead discussant, Invited papers under the topic *Neural Networks* at DataClean.Networkshop.
- Program Committee, Data Quality in Cooperative Information Systems.

Tommy Wright

- Refereed papers for *The American Statistician*.
- Associate Editor, *The American Statistician.*
- Member, Editorial Board, *American Journal of Mathematical and Management Sciences.*
- Member, Editorial Board, *Journal of Transportation and Statistics.*
- Member, Mathematics Department Advisory Council, The University of Tennessee, Knoxville, TN.

Laura Zayatz

- Member, Confidentiality and Access Committee.
- Member, National Center for Education Statistics Disclosure Review Board.
- Reviewed papers for the *Journal of Official Statistics* and *Research in Official Statistics.*
- Reviewed, for disclosure problems, two microdata files that the Social Security Administration wishes to publicly release.

3. PERSONNEL NOTES

David Mingay's passing saddened us all. David was born and raised in England and was educated at Cambridge University where he received a Ph.D. in cognitive psychology. He came to the United States in 1985, and worked at the National Center for Health Statistics, the National Opinion Research Center, and the University of Chicago. David was a pioneer in the use of cognitive interviewing for questionnaire development and in the use of automated systems and data collection methods. He joined the Statistical Research Division in November 2000.

RETIREMENTS from the Census Bureau: Marty Appel, David DesJardins, Chris Dyke, Easley Hoy, Cathy Keeley, and Matt Salo.

Jon Jezyk, an undergraduate senior in Psychology at Catholic University, joined the Usability Staff as an academic year intern.

Eric Slud (Mathematics Professor at the University of Maryland) joined our division on a Schedule A appointment as a member of the Statistical Estimation and Analysis Research Group.

John Aston joined the Time Series Research Group for two years as a National Institute of Statistical Science Postdoctoral Researcher.

Kellie Wills joined the Time Series Research Group.

Cleo Redline accepted a position with the National Science Foundation.

Pam McGovern joined our division on a detail from the Demographic Statistical Methods Division as part of the Census Bureau's Mid-Career Enrichment Program to work on development of a Language Program.

Brian Greenbaum, a senior computer science major at the University of Maryland, College Park, joined our Human Factors and Usability Research Group.

David Findley accepted an appointment at the Census Bureau as Senior Mathematical Statistician for Time Series.

Jennifer Hunter joined our Questionnaire Pretesting for Household Surveys Group.

Yuling Pan joined our Questionnaire Design and Measurement Research (2) Group.

Summer Visitors:

- Cyntrica Eaton, a graduate student in computer science at the University of Maryland, College Park, joined the Usability Lab as a summer intern.

- Susie McFadden-Resper, a graduate student in anthropology at American University, joined the Questionnaire Design & Measurement Research (2) Group as a summer intern.
- Michelle Jiles, a Ph.D. student in sociology at Catholic University, joined the Questionnaire Design & Measurement Research (2) Group as a summer intern.
- Luke Shepard, an undergraduate computer science major at the University of Chicago, joined the Disclosure Limitation Research Group as a summer intern.
- Lesley Smalls (B.S. in computer science from Frostburg State University) joined the Disclosure Limitation Research Group as a summer intern.
- Eric Lee, a graduate student in computer science at the University of Maryland, College Park, joined the Disclosure Limitation Research Group as a summer intern.

David Smith accepted a faculty position in statistics at Tennessee Technological University.

Thuy Trang Nguyen returned to the Manufacturing and Construction Division after a detail assignment to the Time Series Research Group.

José Dulá (Associate Professor at the University of Mississippi) joined our division on a Schedule A appointment as a member of the Computing Research Group through June 2003.

Tom Reynolds (graduate student in psychology at George Mason University) joined the Usability Lab for the 2003 academic year.

In October 2002, our colleague Bev Causey passed away.

APPENDIX A **Statistical Research Division's FY 2002 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		
5604	Special Populations		
6121	Statistical Requirements		
6352	Coverage Measurement A.C.E.		
6367	Census 2000 Evaluation Projects		
6812	Data Product Definition, Creation, Production, and Dissemination		
	1. Decennial Coverage Research	Don Malec	Rick Griffin
	2. Alternative Questionnaire Experiment (AQE2000)	Eleanor Gerber	Ed Byerly
	3. Response Mode & Incentive Experiment (RM&IE)	Larry Malakhoff	Violeta Vazquez
	4. A.C.E. Missing Data Research and Development	Mike Ikeda	Michael Beaghen
	5. Privacy Implications of the Decennial Census	Eleanor Gerber	Jerry Gates
	6. Complex Households and Relationships in the Decennial Census and in Ethnographic Studies	Laurie Schwede	Randall Neugebauer
	7. Generation X Speaks Out on Censuses, Surveys, & Civic Engagement: An Ethnographic Approach	Melinda Crowley	Kevin A. Shaw
	8. Enumeration Barriers Specific to Colonias Along the US/Mexico Border in the Context of Census 2000	Manuel de la Puente	Randall Neugebauer
	9. Comparative Ethnographic Research on Mobile Populations	Jennifer Hunter	Alice Green
	10. Evaluation of the Decennial Frame of Group Quarters and Sources	Leslie Brownrigg	Violeta Vazquez
	11. Administrative Records Linkage Support	Ned Porter	Kimball T. Jonas
	12. Decennial Privacy Research	Tom Mayer	Jerry Gates
	13. Research on Accuracy of the Census and the A.C.E. Estimates	Mary Mulry	Rita Petroni
	14. Behavior Coding Evaluation of the Census Quality Survey	Eleanor Gerber	Dave Hubble
	DEMOGRAPHIC		
0906	Current Population Survey (CPS)		
	15. Tobacco Use Supplement to the 2003 CPS: Smoking Cessation	Melinda Crowley	Gordon Willis
1461	*16. SIPP Methods Panel*	Jeff Moore	Pat Doyle
1465	Survey of Income and Program Participation (SIPP) Research		
	17. Continuous Instrument Improvement Group (CIIG)	Jeff Moore	Bob Kominski
	18. Adapting Standard Analytical Procedures to the Complex Sampling Structure of SIPP	Ruben Mera	Stephen Mack
	Quick Turnaround Pretesting of Household Surveys		
	19. National Crime Victimization Survey	Terry DeMaio	Marilyn Monahan
	20. SIPP Cash Balance Question	Terry DeMaio	Gordon Lester
	21. SIPP Advance Letter Research	Terry DeMaio	Pat Doyle
	22. ACS Veteran's Status Question	Terry DeMaio	Don Fischer
1467	*23. Survey of Program Dynamics*	Jennifer Rothgeb	Patrick Benton
7455	*24. American Housing Survey*	Ruben Mera	Dennis Schwanz
4000	*25. 2000 Sample Redesign Research*	Jay Kim	Danielle Corteville
4200	Continuous Measurement/American Community Survey		
	26. Questionnaire Design & Measurement: Race and Ethnicity Research	Laurie Schwede	Debbie Griffin
	27. Questionnaire Design & Measurement: Border Communities (Colonia) Research	Manuel de la Puente	Debbie Griffin
	28. ACS Language Research: Focus Groups with FRs and SFRs and Cognitive Testing of CAPI/CATI Spanish Language Instrument	Manuel de la Puente	Debbie Griffin
	29. ACS Small Area Estimation Research	Don Malec	Freddie Navarro
	30. Nonresponse Weighting Adjustments for ACS Application	Eric Slud	Larry Cahoon
7165	*31. Research for Small Area Income and Poverty Estimates (SAIPE)*	Elizabeth Huang	Dave Waddington
	ECONOMIC		
3320051	32. *Disclosure Limitation Methods*	Laura Zayatz	Nash Monsour
3320052	Time Series Research		
	33. Seasonal Adjustment Support	Brian Monsell	Anne Russell
	34. X-12-ARIMA Development and Evaluation	Brian Monsell	Catherine Hood
	35. Research on Seasonal Time Series - Modeling and Adjustment Issues	Don Martin	Catherine Hood
3320054	*36. Implicit Ratio Edit Generation*	Maria Garcia	Katherine J. Thompson
TBA	*37. Flexible Matching Imputation Research*	Todd Williams	Doug Bond

	METHODOLOGY AND STANDARDS	
7676	*38. Statistical Consulting/Postal Rate Commission*	Leroy Bailey . . . . . Bob Cohen (USPS)
8083	*39. IT & AT for the Disabled/General Services Administration*	Bill LaPlant . . . . . Ed Reniker (GSA)
7225	*40. U.S. Department of Health & Human Services/Testing Improvements*	Joanne Pascale . . . . Robert Stewart (HHS)
8863	*41. NIST - Bayesian Statistical Methodology*	Don Malec . . . . . Nell Sedransk (NIST)
4100	CASIC	
	42. Metadata Systems Research	Sam Highsmith . . . . . Cheryl Landman
Other	*Usability Laboratory*	
	43. Foreign Born Webpage Usability	Erica Olmsted . . . . . Laura Yax
	44. Administrative and Customer Services Intranet Site Development & Testing	Erica Olmsted . . . . . Susan Boyer
	45. Decennial Management Division (DMD) Intranet Site Accessibility Evaluation	Larry Malakhoff . . . . . Linda Baez
	46. Field Directorate Intranet Site Usability Testing	Erica Olmsted . . . . . Dennis Van Langen
	47. Housing and Household Economic Statistics Information Architecture	Erica Olmsted . . . . . Joe Dalaker
	48. Information Technology Intranet Site Usability Testing	Brian Greenbaum . . . . . Carol Bateman
	49. Mobile Computing Devices – Map Usability Testing	Betty Murphy . . . . . Karen Medina
	50. Response Options Strategies Working Group	Betty Murphy . . . . . Suzanne Fratino
	51. Electronic Response Options Subgroup	Larry Malakhoff . . . . . Suzanne Fratino
	52. Foreign Trade Division (FTD) Internet Site Accessibility Evaluation	Larry Malakhoff . . . . . Blake Sanders
	53. Census 2000 Gateway Testing	Betty Murphy . . . . . Marian Brady
Other	*54. Population Projections and Estimates Web Page Usability Evaluation*	Erica Olmsted . . . . . Todd Gardner
	55. Expert Heuristic Review of the M3 Electronic Questionnaire	Beth Nichols . . . . . Mary C. Potter
	56. Interviewer Refusal Aversion Training	Eileen O Brien . . . . . Adrienne Oneto

FY 2002 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 2002. For FY 2002, the *measures of performance* for our division are:

Measure 1. *Overall, Work Met Expectations*
Percent of FY 2002 Program Sponsored Projects/Subprojects where sponsors reported that work met their expectations.

Measure 2. *Established Major Deadlines Met*
Percent of FY 2002 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 2002 Program Sponsored Projects/Subprojects reporting at least one improved method, techniques developed, solution, or new insight.

Measure 3b. *Plans for Implementation*
Of the FY 2002 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 2002 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 2002.

Measure 6. *Proceedings Publications*
Number of proceedings publications documenting research that appeared in FY 2002.

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 1-14, 2002 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 _____

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 Date
Statistical Research Division

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

Brief Description of Results/Products from FY 2002 ***(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 2002 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 ☐ No ☐
b. _____ Yes ☐ No ☐
c. _____ Yes ☐ No ☐

Comments **(Sponsor)**: _____

_____

_____

_____

_____

_____

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 2002 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 2002.

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 2002 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):** _____

OVERALL: Expectations Met/Improving Future Communications

4. Overall, work on this project or subproject by SRD staff during FY 2002 met expectations. **(Sponsor)**

   ☐ Strongly Agree ☐ Agree ☐ Disagree ☐ Strongly Disagree

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

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

   Suggestions **(Sponsor)** _____

(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
Tom Petkunas
VACANT

Statistical Computing Research
Bill Winkler
Bor-Chung Chen
Maria Garcia
Judi Norvell
Yves Thibaudeau
William Yancey

Computing Applications
Sam Highsmith
Aref Dajani
Ned Porter
Mary Ann Scaggs
Todd Williams

Computing Research
Carol Corby (Acting)
José Dulá (U. Mississippi)
Bill LaPlant
Nita Rasmann
VACANT
VACANT

Assistant Division Chief for Mathematical Statistics

Pat Cantwell
Alice Bell

Sampling Research
Lynn Weidman
Ann Dimler
Mike Ikeda
Jay Kim
Mary Mulry
Gloria Prout
Julie Tsay

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

Time Series Research
Brian Monsell (Acting)
John Aston (NISS Post doc)
Don Martin (Howard U)
Kellie Wills
VACANT

Assistant Division Chief for Survey Methodology

Center for Survey Methods Research

Manuel de la Puente
Maria Cantwell

Questionnaire Design & Measurement Research -1
Jeff Moore
Anna Chan
Julia Klein-Griffiths
Tom Mayer
Beth Nichols
Joanne Pascale
Jennifer Rothgeb

Questionnaire Design & Measurement Research -2
Eleanor Gerber
Melinda Crowley
Safiya Hamid
Pam McGovern*
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
Larry Malakhoff
Betty Murphy
Erica Olmsted
Thomas Reynolds (S)
VACANT

Office of the Chief

Tommy Wright
Hazel Beaton

(S) - Student *Detail from DSMD

November 2002

U.S. Census Bureau
Statistical Research Division
Federal Building 4
Washington, DC 20233
301-457-1030