Annual Report *of the* Statistical Research Division

Fiscal Year 2010

Decennial Directorate Customers

Demographic Directorate Customers

Economic Directorate Customers

Field Directorate Customers

Other Internal and External Customers

Estimation	Ethnography	Time Series
Usability	Edit and Imputation	Language and Sociolinguistics
Questionnaire Design		Small Area Estimation
Record Linkage	**SRD Expertise for Collaboration and Research**	Statistical Computing
Sampling	Disclosure Avoidance	Modeling
Measurement	Seasonal Adjustment	Cognitive Testing

*S*ince August 1, 1933—

"... As the major figures from the American Statistical Association (ASA), Social Science Research Council, and new Roosevelt academic advisors discussed the statistical needs of the nation in the spring of 1933, it became clear that the new programs—in particular the National Recovery Administration—would require substantial amounts of data and coordination among statistical programs. Thus in June of 1933, the ASA and the Social Science Research Council officially created the Committee on Government Statistics and Information Services (COGSIS) to serve the statistical needs of the Agriculture, Commerce, Labor, and Interior departments ... COGSIS set ... goals in the field of federal statistics ... (It) wanted new statistical programs—for example, to measure unemployment and address the needs of the unemployed ... (It) wanted a coordinating agency to oversee all statistical programs, and (it) wanted to see statistical research and experimentation organized within the federal government ... In August 1933 Stuart A. Rice, President of the ASA and acting chair of COGSIS, ... (became) assistant director of the (Census) Bureau. Joseph Hill (who had been at the Census Bureau since 1900 and who provided the concepts and early theory for what is now the methodology for apportioning the seats in the U.S. House of Representatives) ... became the head of the new Division of Statistical Research ... Hill could use his considerable expertise to achieve (a) COGSIS goal: the creation of a research arm within the Bureau ..."

Source: Anderson, M. (1988), *The American Census: A Social History*, New Haven: Yale University Press.

Among others and since August 1, 1933, the Statistical Research Division has been a key catalyst for improvements in census taking and sample survey methodology through research at the U.S. Census Bureau. The introduction of major themes for some of this methodological research and development where staff of the Statistical Research Division[1] played significant roles began roughly as noted—

- **Early Years (1933–1960s):** *sampling (measurement of unemployment and 1940 Census); probability sampling theory; nonsampling error research; computing; and data capture.*

- **1960s–1980s:** *self-enumeration; social and behavioral sciences (questionnaire design, measurement error, interviewer selection and training, nonresponse, etc.); undercount measurement, especially at small levels of geography; time series; and seasonal adjustment.*

- **1980s–Early 1990s:** *undercount measurement and adjustment; ethnography; record linkage; and confidentiality and disclosure avoidance.*

- **Mid 1990s–Present:** *small area estimation; missing data and imputation; usability (human-computer interaction); and linguistics, languages, and translations.*

At the beginning of FY 2011, most of the Statistical Research Division became known as the Center for Statistical Research and Methodology. In particular, with the establishment of the Research and Methodology Directorate, the Center for Survey Measurement and the Center for Disclosure Avoidance Research were separated from the Statistical Research Division, and the remaining unit's name became the Center for Statistical Research and Methodology.

[1]The Research Center for Measurement Methods joined the Statistical Research Division in 1980. In addition to a strong interest in sampling and estimation methodology, research largely carried out by mathematical statisticians, the division also has a long tradition of nonsampling error research, largely led by social scientists. Until the late 1970s, research in this domain (e.g., questionnaire design, measurement error, interviewer selection and training, nonresponse, etc.) was carried out in the division's Response Research Staff. Around 1979 this staff split off from the division and became the Center for Human Factors Research. The new center underwent two name changes—first, to the Center for Social Science Research in 1980, and then, in 1983, to the Center for Survey Methods Research before rejoining the division in 1994.

U.S. Census Bureau
Statistical Research Division
Room 5K108
4600 Silver Hill Road
Washington, DC 20233
301-763-1702

We help the Census Bureau improve its processes and products. For fiscal year 2010, this report is an accounting of our work and our results.

Statistical Research Division

At the beginning of FY 2011, most of the Statistical Research Division became known as the Center for Statistical Research and Methodology. In particular, with the establishment of the Research and Methodology Directorate, the Center for Survey Measurement and the Center for Disclosure Avoidance Research were separated from the Statistical Research Division, and the remaining unit's name became the Center for Statistical Research and Methodology.

Highlights of What We Did...

As a technical resource for the Census Bureau, each researcher in 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 and/or subprojects.

Highlights of a selected sampling of the many activities and results in which Statistical Research Division staff members made contributions during FY 2010 follow, and more details are provided within subsequent pages of this report:

– developed a model and associated estimation methodology based on certain moment conditions without fully specifying parametric distribution to estimate proportion or erroneous enumeration in a decennial census; developed a new model for analyzing categorized data with possibly non-ignorable non-response for coverage component estimation; examined the error structure in the estimates of immigration which are a component of the forthcoming 2010 Demographic Analysis estimates.

– released an updated version of *X-12-ARIMA*, Build 188 of version 0.3; continued empirical studies of model-based seasonal adjustment diagnostics, developed algorithms and code for estimation of seasonal long memory models, studied Markov Chain Monte Carlo techniques for Bayesian estimation of seasonal time series models, studied how seasonal adjustment methods facilitate short-term forecasting of cyclical and trend dynamics, and implemented a method to fit time series models by minimizing multi-step ahead forecasting error.

– finalized and sent Disclosure Avoidance Specifications for Census 2010; investigated an alternative method of data swapping (cyclical rather than pairwise) with improved performance; developed, programmed, tested, and evaluated different methods for repair of the age variable in previously released Census 2000 Public Use Microdata Samples (PUMS); worked contractors on the development of disclosure avoidance techniques and software for the Census Transportation Planning Package ACS special tabulation.

– developed software for applying synthetic data procedures to the edited Group Quarters 2010 Census data; generated a new package in R, tentatively called PEP, to generate edit-compliant synthetic data for ACS.

– conducted cognitive pretesting of five alternative versions of the race and ethnicity questions for the 2010 Census Program for Evaluations and Experiments; pretested eleven different Alternative Questionnaire Experiment (AQE) race and ethnicity question panels that are translated into Spanish; conducted ad hoc review of Spanish and Chinese 2010 Census promotion materials.

– completed work with focus groups across the United States with members of cohabitating couples to collect qualitative information about alternative terms, definitions, categories, on questions that most accurately measure relationship status and partnership situations; coordinated the successful OMB approval of 32 pretesting of questionnaire requests involving 3,388.5 respondent burden hours across all program areas.

– performed three rounds of usability testing for the development of an online instrument for the Decennial Reinterview Project; performed accessibility testing of various websites, applications, and documents, including the Business Help Site, Census 2010 Web Site, Census In the Schools documents, and Web TA.

How Did We Do...

For the 12th year, we received feedback from our sponsors. Near the end of fiscal year 2010, our efforts on fifty-five of our program (Decennial, Demographic, Economic, External) sponsored projects/subprojects with substantial activity and progress and sponsor feedback (Appendix A) were measured by use of a Project Performance Measurement Questionnaire (Appendix B). Responses to all fifty-five questionnaires were obtained with the following results (The graph associated with each measure shows the performance measure over the last 12 fiscal years):

Measure 1. Overall, Work Met Expectations

Percent of FY2010 Program Sponsored Projects/Subprojects where sponsors reported that overall work met their expectations (agree or strongly agree) (54 out of 55).......98%
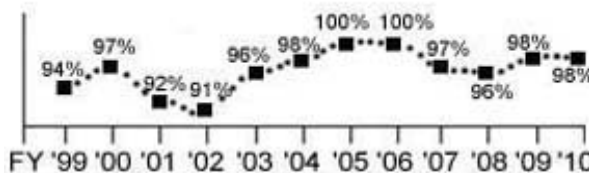
Measure 2. Established Major Deadlines Met

Percent of FY2010 Program Sponsored Projects/Subprojects where sponsors reported that all established major deadlines were met (43 out of 48 responses) 90%
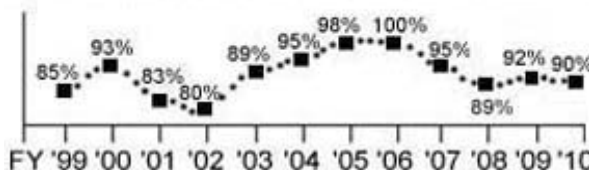
Measure 3a. At Least One Improved Method, Developed Technique , Solution, or New Insight

Percent of FY2010 Program Sponsored Projects/Subprojects reporting at least one improved method, developed technique, solution, or new insight (48 out of 55 responses).............87%
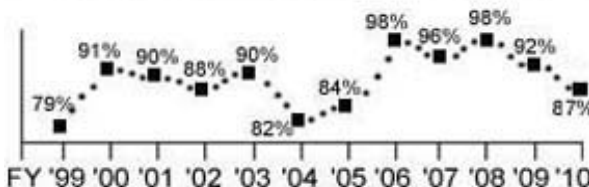
Measure 3b. Plans for Implementation

Of these FY2010 Program Sponsored Projects/Subprojects reporting at least one improved method, technique developed, solution, or new insight, the percent with plans for implementation (38 out of 48 responses) 79%
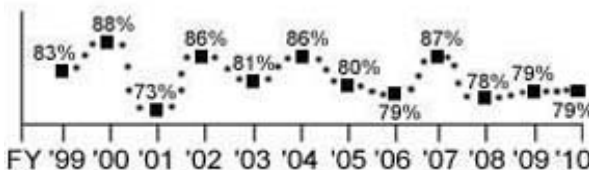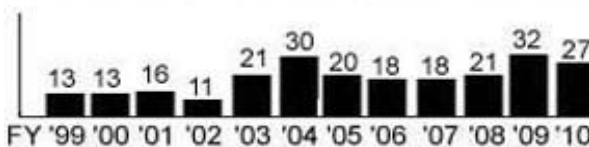
Measure 4. Predict Cost Efficiencies

Number of FY2010 Program Sponsored Projects/Subprojects reporting at least one "predicted cost efficiency" 18
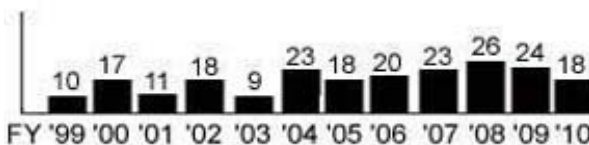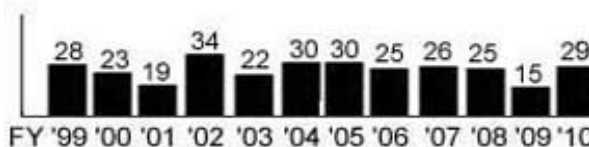
From Section 3 of this ANNUAL REPORT, we also have:

Measure 5. Journal Articles, Publications

Number of peer reviewed journal publications documenting research that appeared (14) or were accepted (13) in FY2010 27

Measure 6. Proceedings, Publications

Number of proceedings publications documenting research that appeared in FY2010 29

Measure 7. Division Research Reports/Studies, Publications

Number of division research reports/studies publications documenting research that appeared in FY2010 46
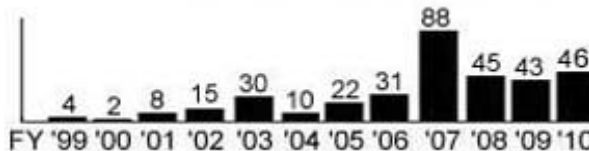
Each completed questionnaire is shared with appropriate staff to help improve our future efforts.

TABLE OF CONTENTS

1. COLLABORATION

1.1 FORMS DEVELOPMENT (DECENNIAL PROJECT 5210001)

A. Census Questionnaire Design Features (Other than Race and Ethnicity)

Description: This project involves participation in decennial content team meetings, including the Content and Forms Design Integrated Product Team, the Housing Unit Operational Integration Team, the Nonresponse Followup Instrument Subteam, the Mode Consistency Subteam, and the Census Program for Evaluations and Experiments (CPEX) Implementations Team. It also includes cognitive pretesting of census questionnaires and census experimental questionnaires.

Highlights: During FY 2010, staff assisted the Field Division with updating training on how to deal with erasures on the Enumerator Questionnaire forms. Staff also provided expert guidance on the development of and testing plan for the experimental Internet census form for the 2010 Census Quality Survey.

Staff: Jennifer Hunter Childs (x34927), Nathan Jurgenson, George Higbie, Anissa Sorokin, Matthew Clifton, Lorraine Randall

B. Development of Race and Ethnicity Questions
Description: This project involved the cognitive pretesting of five alternative versions of the race and ethnicity questions used in the Decennial Census for the 2010 Census Program for Evaluations and Experiments (CPEX) panels. It also involved cognitive pretesting of race and ethnicity questions used in a re-interview which will be conducted in the 2010 Census.

Highlights: During FY 2010, staff conducted cognitive pretesting of five alternative versions of the race and ethnicity questions used in the Decennial Census for the 2010 Census Program for Evaluations and Experiments (CPEX) panels.

Staff also conducted cognitive pretesting of race and ethnicity questions used in a reinterview which will be conducted in the 2010 Census. Staff revised the reintervew cognitive pretesting report and conducted more analysis of reinterview cognitive pretesting data.

Staff also worked with the Decennial Management Division and the Population Division to pretest 11 different Alternative Questionnaire Experiment (AQE) race and ethnicity question panels that are translated into Spanish. Staff revised pretesting protocols for the 11 question panels to address issues relevant to Spanish monolinguals, and then used a committee approach to translate these protocols into Spanish. Staff developed strategies for recruiting Spanish monolinguals of various nationalities, races, and ethnicities. Staff recruited respondents, conducted cognitive interviews and presented preliminary results to the Population Division and Decennial Management Division.

Staff: Rodney Terry (x35475), Jennifer Hunter Childs, Patricia Goerman, Terry DeMaio, Yuling Pan, Matthew Clifton, George Higbie, Amelia Tseng, Marissa Fond, Nathan Jurgenson

1.2 LANGUAGE PLANNING AND DEVELOPMENT (Decennial Project 5210003)

Description: Staff members participate in the inter-divisional Decennial Task Force, or language team, which focuses on developing and planning the Language Program for the 2010 Census, pre-census tests, and the Dress Rehearsal. In addition, staff members in our division provide consultation and technical support in the design, development and conduct of research for Decennial language-related projects.

Highlights: During FY 2010, staff conducted ad hoc review of Spanish and Chinese 2010 Census promotion materials. Staff participated in and contributed to a series of Lessons Learned meetings regarding the 2010 Decennial Census language program. Staff provided verbal and written comments to the Decennial Management Division.

Staff: Patricia Goerman (x31819), Yuling Pan

1.3 DATA COLLECTION PLANNING AND DEVELOPMENT (Decennial Project 5310001)

A. Accessible Web Surveys (Research)
Description: There is much for Web survey designers to keep in mind when designing surveys to conform to *Section 508* regulations. The regulations require persons with disabilities to have access comparable to the access available to others. This means individuals with visual deficits who use a screen-reader to read text must have the same visual sequence of questions, answer choices, skip patterns, and instructions.

Highlights: Providers of Web survey tools claim to conform to *Section 508*, but are the applications actually accessible? During FY 2010, staff corresponded with software developers at the University of Michigan, ISR,

Indiana University, Research Triangle Institute, Washington State University, and the Bureau of Labor Statistics to obtain sample surveys with common questionnaire elements and performed a comparison of the accessibility of question grids, radio button groups, check-boxes, logical keyboard navigation with the presence of visual focus, and data-entry field labels. The products evaluated include Microsoft .NET 3.5, Illume 4.7.0.46, Blaise IS 4.8.1.1460, Hatteras 3.0, SurveyMonkey (2010), and Jdeveloper & HomeSite (J&H). These packages were evaluated with Internet Explorer 7 and the Job Access With Speech (JAWS) 11 screen-reader software.   The evaluation revealed Illume and Blaise IS developers need to inform screen-reader users about table commands to make their question grids accessible. The only web survey to provide visual focus when tabbing with JAWS 11 running was developed with J&H. Microsoft .NET developers need to provide an alternative to color to inform respondents about missing data. JAWS users can read J&H and SurveyMonkey screens from top to bottom without back-tracking to locate the question. Vocalized labels did match text visible on the screen, except for Blaise IS and Hatteras. SurveyMonkey developers should correct display issues before posting a survey. The evaluation criteria ranked the Web survey development tools in this order (most to least accessible) J&H, SurveyMonkey, Hatteras, Illume, Blaise IS, and Microsoft .NET.

Staff: Lawrence Malakhoff (x33688), Temika Holland, Andrew Zukerburg (NCES)

B. Decennial Reinterview Internet Testing – Usability Input
Description: The usability team works with the cognitive team on the development of an online instrument for the Decennial Reinterview Project. The project consists of two rounds of usability and cognitive testing of prototypes of the internet instrument and the letters associated with the paper forms. Staff provides test plans and formal reports for each round of testing as well as provides input at regular team meetings during the development of the ACS online instrument and its subsequent field testing.

Highlights: During FY 2010, staff attended regular meetings and assisted the sponsor team in developing wireframe screens for the Internet data collection interface. Also, two rounds of usability and cognitive testing were planned and are expected to take place later in FY 2010.

Staff have performed three rounds of usability testing and are in preparation for testing of the final instrument in FY 2011.

Staff: Kathleen Ashenfelter (x34922), Temika Holland, Victor Quach, Elizabeth May Nichols

1.4 SPECIAL PLACE/
GROUP QUARTERS (GQ) PLANNING AND DEVELOPMENT
(Decennial Project 5310008)
[See Projects 5610005 and 5610006 (F).]

1.5 STATISTICAL DESIGN AND ESTIMATION
(Decennial Project 5610002)

A. Decennial Record Linkage
[See Projects 0351000 and 1871000 (A), General Research - Statistical Computing Methodology]

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

Highlights: During FY 2010, staff reviewed swapping methods in preparation for the upcoming running of the official swapping program, developed by disclosure avoidance staff, on Decennial 2010 state files. This review involved discussion of issues such as swapping rate and imputation rate that affect program output. Some documents on these topics were sent to decennial staff. An official confidential memo was sent from the disclosure avoidance staff to Dennis W. Stoudt, Decennial Systems and Processing Office (DPSO), Michael Clark (DSPO), and several others in DSPO, that has, as an attachment, the Disclosure Avoidance Specifications for Census 2010. This confidential attachment describes the way swapping is implemented, e.g., it describes the input and output files, the way the SAS program should be run, the way the results should be analyzed, and lists the code for the program.

Staff investigated an alternative method of data swapping (cyclical rather than pairwise). The new method performs better in terms of both data protection and data quality. Documentation on this new method has been drafted.

Staff developed, programmed, tested, and evaluated different methods such as rank swapping, rank swapping with the perturbation rule, rank swapping with stratification, synthetic data, and swapping based on the empirical conditional distributions as possibilities for the repair of the age variable in previously released Census 2000 Public Use Microdata Samples (PUMS).   Staff

developed software for the repair of the previously released PUMS files and the findings were documented.

Staff developed software for data swapping for the Island Areas Census 2010. The program was tested using the decennial 2000 data. The findings were documented.

Staff worked with contractors from WESTAT on the development of disclosure avoidance techniques and software for the Census Transportation Planning Package ACS special tabulation.

Staff developed software for applying synthetic data procedures to the edited Group Quarters 2010 Census data. The software protects person level-sensitive demographic attributes such as age, gender, and ethnicity at various geographic levels such as tract, block, county, PUMAs, and state. Staff implemented a test-run exercise using states Ohio and Alabama to evaluate the quality of the software and synthetic output. Results of the test run were reviewed to develop the specifications for production.

The Disclosure Review Board continued work on developing disclosure rules for ACS five-year base tables.

Staff: Laura Zayatz (x34955), Asoka Ramanayake, Jason Lucero, Paul Massell, Julie Tsay, Michael DePersio

C. Census Unduplication Research
Description: The goal of this project is to conduct research to guide the development and assessment of methods for conducting nationwide matching and unduplication in the 2010 Decennial Census and future Censuses.   One of the major problems is how to incorporate the effects of name frequency into the unduplication procedures.   Our staff also provides assistance in specifying and reviewing output from the matching and unduplication procedures for test Censuses and for Census 2010. We began this project in May of 2004.

Highlights: During FY 2010, staff continued examining results from running the four-pass matching system on the Census 2000 data. Based on this examination, some revisions were made to the list of invalid names for matching in the 2010 Decennial Census. Two related specifications have been revised: "2010 Decennial Census Coverage Followup and Census Coverage Measurement Match Modeling Software Requirements Specification" (procedures for evaluating person links) and "2010 Decennial Census Census Coverage Followup and Census Coverage Measurement Person Matching Parameter Software Requirements Specification" (parameter settings for matching). Staff are co-authors for both specifications.   Staff attended multiple sessions starting in early April 2010 and extending to early August 2010 to review data and set cutoffs for within-response (multiple links between a pair of returns) links from the

2010 Census Duplicate Person Identification (DPI) process. The review covered ninety-three separate runs of the DPI system. Staff produced first drafts of three memoranda for the record: one provided a rough comparison of results between the 2000 research and the 2010 DPI, one listed various situations encountered in either the 2000 research or the 2010 DPI and their resolutions, and one documented the rules we tended to follow in setting cutoffs for within-response links.

Staff: Michael Ikeda (x31756), Ned Porter

D. Statistical Design for Experiments and Evaluations
Description: We will provide statistical expertise in the experimental design for decennial-related experiments and evaluations.   We will also investigate novel methodologies for analyzing top coded income data that follows a lognormal distribution.

Highlights: For applying latent class analysis to survey data, the survey conditions need not be identical for the different interviews under the Internet Reinterview Evaluation Study Plan of the 2010 Census. During FY 2010, we planned a latent class analysis that permits the estimation of simple response variances under assumptions that are more general than those made in traditional analysis.

Staff: Thomas Mathew (x35337)

1.6 COVERAGE MEASUREMENT PLANNING AND DEVELOPMENT (Decennial Project 5610003)

A. Coverage Measurement Research
Description: Staff members conduct research on model-based small area estimation of census coverage, and they consult and collaborate on modeling census coverage measurement (CCM).

Logistic Mixed Models for Small Domain Estimation
During FY 2010, an evaluation to include random effects in the logistic coverage models as a way to safeguard against model error (due to type II error) was conducted using Census 2000 revision II data on the correct enumeration rate The aim of the study is to determine the possible gains in using random effects in the coverage measurement context and to determine the usefulness of available software. The resulting effect estimates demonstrate the typical compromise between their corresponding fixed effect estimate and zero. If only one random effect distribution is used for a large number of effects (many which may be zero), the added random effects contribute little due to a small estimated variance component.   If however, subsets have separate random effects distributions, some effects make a larger contribution. Using a large, realistic model based on Decennial Statistical Studies Division (DSSD) coverage

research, standard SAS procedures often never converged. Both an implementation of Breslow and Lin's approximation in PROC IML and the Pinero and Bates procedure, implemented in R, performed well. Work in developing data-based procedures to select multiple random effect models and better ways to handle the "design effects" in the data are needed. It is suggested that the techniques could be used, as is, as part of a sensitivity test.

Small Area Coverage Estimation

Hierarchical Bayes Small Area Estimation: During FY 2010, small area estimates using data from the 2006 Travis County test site were made under a variety of assumptions. A draft report was disseminated to the DSSD Small Area Estimation Group.

Estimation of Synthetic Estimators in the CCM Program: The CCM Group is interested in evaluating the performance of a synthetic estimator based on a logistic model for small areas (e.g., counties, places, etc.). While standard design-based methods can be applied to estimate the design-based variance of the synthetic estimator, it is either impossible or difficult to produce a reliable estimate of design-based bias of the synthetic estimator of an individual small area. A reasonable solution to this problem, as discussed in the CCM Group meeting, is to evaluate the synthetic estimator using the average design-based mean square error (MSE) criterion, where the average is taken over small areas in a given group defined by certain criteria (e.g., size). During FY 2010, the CCM Group tried out the average design-based MSE estimator proposed earlier by Gonzales and Waksberg on a test Census dataset, but for many groups the method produced negative average MSE estimates, which motivated staff to consider a simple average MSE estimator that will always yield strictly positive MSE estimates while maintaining good design-based properties. The method-of-moments type estimator of average MSE seems to produce more reasonable results than Gonzales-Waksberg method in the preliminary data analyses done by the CCM Group.

Empirical Best Prediction (EBP) Method to Estimate the Proportion of Erroneous Enumeration for the 2006 Travis County, Texas Test Data:
The project is very much related to the Census Bureau's effort to understand how future censuses can be improved. The main goal of this project is to develop a robust statistical method to estimate the proportion of erroneous census enumeration, an important component of possible census error that the CCM group is studying. During FY 2010, a model and the associated estimation methodology were developed based on certain moment conditions without fully specifying parametric distributions. The method will be tested on the 2006 Travis County, Texas test data. Staff worked on the development of a small area Census Coverage model for race/origin, age/sex, and tenure covariates, under a two-

stage sampling design developed by Partha Lahiri. Staff found variance estimates and design effects for all P-Sample and E-Sample clusters for 2006 Census Coverage and researched methods of calculating simultaneous equations in SAS and incorporating the design components.

Non-ignorable Models for Coverage Component Estimation
During FY 2010, staff developed a new model for analyzing categorical data with possibly non-ignorable non-response. The model construction is based on partitioning techniques and model averaging. This new model was compared to non-ignorable models published in the statistical literature through data analysis of unresolved correct enumeration status and simulation studies. An over-parameterized model that specifies that all response rates and outcome rates are, a priori, uniformly distributed was also included in the analysis. Results suggest that both a basic ignorable model and a hierarchical non-ignorable model are not very robust. The partitioning model is also non robust but more robust than the basic models. The uniform over-parameterized model appears robust at the expense of a large variance. This study should be a useful tool for analyzing the sensitivity of the estimates to different model assumptions. A report documenting this work is being prepared.

Coverage Measurement Planning and Development:
During FY 2010, staff used data from Census 2000 and Census Coverage Estimation to develop race-based, data-driven models for Match rates and Correct Enumeration rates. These models were used to evaluate the viability of producing multiracial estimates, the success of the current race/origin domain framework, and the validity of modeling different racial subpopulations (among other questions). Staff presented recommendations several times to Decennial Statistical Studies Division stakeholders, in meetings of the Component Missing Data team, and the CCM Estimation team. Staff is submitting a research report for review based upon this work. Staff has also developed a SAS program to automate the recursive partitioning procedure, based on previous partitioning work.

Staff attend the regular DSSD group meetings: the CCM Estimation Team Meeting, the Component Missing Data Meetings and the Small Area Estimation Group Meetings.

Staff: Don Malec (x31718), Aaron Gilary, Ryan Janicki, Jerry Maples

B. Accuracy of Coverage Measurement
Description: 2010 Census Coverage Measurement (CCM) Research conducts the research necessary to develop methodology for evaluating the coverage of the

2010 Census. This includes planning, designing, and conducting the research, as well as analyzing and synthesizing the results to evaluate their accuracy and quality. The focus is on the design of the Census Coverage Measurement survey and estimation of components of coverage error with secondary emphasis on the estimation of net coverage error. The estimation of overcount and undercount separately has not been done for previous censuses because of the difficulty of obtaining adequate data for unbiased estimates.

Highlights: During FY 2010, staff provided technical expertise and experience in the planning and implementation of coverage measurement research for the 2010 Census. This included serving on teams formed to conduct research in preparation for planning and implementing CCM.

Work continued on designs for some of the CCM evaluation studies in the 2010 Census Program for Evaluations and Experiments (CPEX). The combination of CCM and CPEX projects is designed to provide information about the basic types of errors that may affect the CCM implementation.

Our staff completed a comprehensive document that describes the work to date on the error structure for estimates of components of census coverage error. The nonsampling error components are defined in a manner that facilitates modeling with the evaluation data that will be available from the CCM CPEX studies. In addition, a second draft of the study plan for the CPEX study "Developing an Error Structure in Components of Census Coverage Error" was prepared for a presentation to the Census Integration Team in October 2010.

In addition, staff examined the error structure in the estimates of immigration which are a component of the forthcoming 2010 Demographic Analysis estimates. The Demographic Analysis estimates are used to evaluate census coverage at the national level. Also, Demographic Analysis sex ratios are used in constructing an adjustment for correlation bias in the dual system estimator used by CCM for estimating census net coverage error.

Staff: Mary Mulry (x31759)

C. Questionnaire Wording and Automation Team
Description: The purpose of this project is to design the coverage measurement survey instruments for the 2010 Census. These instruments will gather enough data to measure both person and household coverage of the 2010 Census. In preparation for 2010, there will be a 2006 test of the coverage measurement operation in specific sites in conjunction with the 2006 Census Test. For 2006, there will be automated person interview (PI) collecting an independent roster of people living at pre-selected sample addresses in the sites and their residency. There will also be a paper-based person followup (PFU) questionnaire

which collects additional residency information about some people collected in the census or the independent roster, but for whom we did not collect enough residency information to determine where they should have been counted for the census. Both these instruments will be used to measure person coverage. Our immediate goals are to create and test these two instruments given requirements from other teams working on coverage measurement planning. This team is further tasked with developing the independent housing unit listing booklet (ILB), and housing unit followup (IHUFU) forms in order to measure housing unit coverage in 2008/2010.

Highlights: During FY 2010, the 2010 CCM operations of listing, housing unit follow-up and person interviewing occurred.

Staff observed the 2010 Census Coverage Measurement listing operation. The observation is documented in Nichols, B. (2009). "Observation of the 2010 Census Coverage Measurement Independent Listing Operation in Prince Georges County, Maryland on October 27, 2009." *DSSD 2010 Census Coverage Measurement Memorandum Series #2010-F-14*, November 17, 2009.

Staff observed the 2010 Census Coverage Measurement Initial Housing Unit Followup operation. The observation is documented in Nichols, B. (2010). "Observation of the 2010 Census Coverage Measurement Initial Housing Unit Followup (IHUFU) Operation in Maryland on March 9, 2010," *DSSD 2010 Census Coverage Measurement Memorandum Series #2010-F-18*, March 23, 2010.

Staff provided comments on the CCM Person Interview training and then observed PI interviews in the field for the respondent debriefing project.

Staff prepared for the Person Followup (PFU) interview which will occur in FY 2011. Staff provided comments on the CCM PFU training. Staff analyzed and reported on the respondent debriefings conducted as part of the two Person Followup operational tests. The field tests were conducted to test modifications to the form to assist interviewers in finding the most knowledgeable respondent for the interview. (This proved problematic in recent Census Tests.)

Results of the first field test showed that the interviewers did not ask the scripted questions to determine whether the respondent was knowledgeable about where the followup person lived in 2008. Thus, the questions needed more revision. Following changes to the content and layout of the questions and also to the training of the interviewers, a second field test was conducted in August, 2009, again with expired CPS sample. Interviews were conducted in New York, NY and several counties in the Washington, DC metropolitan area. The revised questions were more successful. Interviewers asked them correctly and navigated properly through the form. The debriefing questions showed that proxy respondents did not have a

good idea of the kinds of questions that would be asked. This occurred in instances where the proxy said they had heard of the follow-up person but did not know them well enough to answer questions; it also occurred when the proxy said they did know them well enough and then did not have enough information to answer the questions. The results of these field tests will be incorporated in the 2010 PFU operation, which will be conducted in January 2011. Findings are documented in Nichols, B., and Childs, J. (2009). "2008 and 2009 PFU Operational Tests Respondent Debriefing Results." *DSSD 2010 Census Coverage Measurement Memorandum Series #2008-D7-17*, October 27, 2009.

Staff: Elizabeth Nichols (x31724), Jennifer Hunter Childs, Terry DeMaio, Nathan Jurgenson

1.7-1.8 COVERAGE IMPROVEMENT PLANNING AND DEVELOPMENT/ EVALUATION PLANNING COORDINATION (DECENNIAL PROJECTS 5610005 AND 5610006)

A. Development of Questionnaires for Decennial Coverage Improvement
Description: Staff consults on and participates in the development of questions and questionnaires designed to improve within household coverage in the Decennial Census. We participate in the development and pretesting of household and individual-level coverage questions in the decennial short form and the Coverage Followup (CFU) reinterview instrument as well as experimental versions of both.

Highlights: During FY 2010, staff assisted in the development of and completed cognitive testing on an alternative version of the Coverage Followup (CFU) instrument, called the Targeted CFU (TCFU), which will be subjected to a large qualitative test in 2010 using cases that have been identified as duplicates in the 2010 Census. Staff presented results of the cognitive testing and corresponding recommendations to team. Based on these recommendations, the TCFU questionnaire was revised and improved. Staff also assisted the Decennial Statistical Studies Division in planning a large-scale cognitive test and qualitative study of the TCFU and duplicates in the 2010 Census. This study will be conducted under contract but with significant technical oversight by staff in our division. For this contract, staff coordinated operations with and prepared training materials for the contractor and began to pilot test the materials.

Staff: Jennifer Hunter Childs (x34927), Anissa Sorokin, George Higbie, Matthew Clifton, Nathan Jurgenson, Lorraine Randall

B. 2010 CPEX Experimental Overcount Booklet
Description: The purpose of this project is to develop and test an alternative mailout census booklet with special coverage questions to compare to the standard census form in terms of coverage in the Census 2010 Alternative Questionnaire Experiment split-panel test. Both forms include a question asking whether each person in the household sometimes lives or stays somewhere else and for what reason. On the standard census form, this question functions as a flag for later phone followup to get more complete coverage data. The alternative mailout booklet converts this question into a screener for a new set of questions on the mailout form itself to identify respondents' alternative addresses and where to count that respondent. If it works, the alternative approach has the potential to improve coverage as well as cut the costs and time involved in conducting followup operations.

Highlights: During FY 2010, we sent the revised final report to the sponsor for final approval. It is entitled: "Cognitive Testing of the Census 2010 Experimental Overcount Questions for the Census 2010 Alternative Questionnaire Experiment: Final Project Report." In this project, we identified serious problems with the skip and the other address question and noted that the sequence may not work for tenuously attached persons with more than two addresses. We recommended format and wording changes, some of which were accepted by the Decennial Statistical Studies Division. We also documented that the experimental overcount question sequence we developed did succeed in identifying, from information on the census form itself, the correct residence for 75% of the persons in the household who sometimes live or stay somewhere else. This suggests that the experimental form may work well in a live census to resolve overcount coverage questions from data on the form itself, thereby reducing the costs and time of later phone follow-ups. The revised booklet will be sent out as an experimental panel in the 2010 Census "Avoid Followup Experiment."

At the invitation of the sponsor, one staff member served as a critical reviewer of the Decennial Statistical Studies Division's 2010 CPEX Experiment Study Plan: "The Avoid Followup Experiment Study Plan" and submitted an extended list of comments and suggestions on additional research questions that would strengthen this study plan.

Staff: Laurie Schwede (x32611), Anissa Sorokin, Virginia Wake Yelei.

C. Evaluations, Experiments, and Assessments Operational Integration Team (EEA OIT)
Description: The purpose of the EEA OIT is to facilitate planning and timely implementation of 2008 Census Dress Rehearsal and 2010 Census evaluations, experiments, and assessments. The group guides and monitors the development, implementation, and reporting of the 2010 evaluations, experiments and assessments. It

ensures that program integration and implementation of the 2010 Census Program of Evaluations and Experiments (CPEX) meets the guidance provided by the Census Integration Group and prepares and monitors the 2010 Census Program for Evaluations and Experiments Master Plan.

Highlights: During FY 2010, staff served as an advisor to the Evaluations, Experiments, and Assessments OIT and as a Co-Advocate for Coverage Improvement Evaluations and Experiments by attending meetings, giving feedback on other research related to CPEX evaluations, and keeping researchers in our division informed of DMD/DSSD updates. Staff keeps the DMD and DSSD managers informed of progress on our division's CPEX evaluations. At the request of the General Accountability Office, staff submitted a copy of the draft study plan for the 2010 CPEX evaluation, "Comparative Ethnographic Studies of Enumeration Methods and Coverage." Additionally, staff served as an official critical reviewer for the Decennial Statistical Studies Division Avoid Followup CPEX experiment study plan and as a pre-reviewer for related CPEX study plans.

Staff: Laurie Schwede (x32611)

D. Evaluation of CCM Interviews
Description: The 2010 Census Program for Evaluations and Experiments (CPEX) includes studies that focus on the quality of the data collected in Census Coverage Measurement Program (CCM). In particular, the focus is on two CCM interviews, the Person Interview (PI) and the Person Followup (PFU) in 2010. The primary methodologies used to evaluate the PI and PFU are respondent debriefing studies and recall bias studies. These studies will provide information about how well the CCM instruments capture the members of the household at each housing unit on CCM interview day and the usual residence of each household member and/or followup person on Census Day. The recall bias study also investigates the quality of the reporting of dates that respondents moved and the reporting regarding previous residents of the housing units. Additionally, these studies will highlight the causes and possible remedies within the questionnaire for any errors of usual residence and household membership.

Highlights: During FY 2010, staff finalized the Census Program for Evaluation and Experiments (CPEX) study plan on the planned respondent debriefings for the CCM Person Interview and Person Followup. Staff presented the study plan to the Census Integration Group (CIG) on April 7, 2010. Staff prepared and trained nine headquarters staff to be observers for the Person Interview (PI) respondent debriefings. The observers were to observe, tape record, and conduct respondent debriefings for the CCM Person Interview. Staff in our division accomplished their goal of seeing at least 16 PI interviews.

Staff assisted Decennial Statistical Studies Division (DSSD) in acquiring the 2006 Questionnaire Design and Experimental Research Survey (QDERS) specification, output programs and training. DSSD will use a modified 2006 QDERS instrument to conduct a recall bias study for CCM.

Staff drafted a Request for Information (RFI) to independently study memory of landmark events. This will be a separate project not connected with the CCM recall bias study. Staff sent out the RFI to the 2014 R&D contractors and the general contracting office. Staff wrote a Request for Proposal (RFP) for NORC to study memory recall of migration dates in the National Longitudinal Survey of Youth. NORC will perform this work in fiscal year 2011.

Staff: Elizabeth Nichols (x31724), Mary Mulry, Jennifer Hunter Childs

E. Investigation of Study Methods for the Census Coverage Measurement (CCM) on Group Quarters (GQ) Population
Description: This project undertakes research and studies before and during the 2010 Census to ultimately develop potential methods for assessing the group quarters population coverage accuracy in the 2020 CCM program. Study methods for the 2010 research includes field observations, in-depth interviews, focus groups, cognitive pretesting, ethnography, respondents debriefings, and a pilot small scale post-enumeration CCM-like survey with student population residing at university housing in 2010. Staff will document the success and difficulties for conducting a 2010 ethnographic study on the coverage measurement evaluation of each of the eight broad types of group quarters populations and a pilot field test of a CCM-like survey with the student population.

Highlights: During the first quarter of FY 2010, staff completed 15 in-depth interviews with GQ administrators and the 2010 GQ project study plan. The plan was presented to and approved by the Census Integration Group. Staff recruited six university researchers to conduct the ethnography study of group quarters in the 2010 Census. Each researcher will be studying one specific broad type of group quarters. During the second quarter, staff prepared and finalized a statement of work, budget, and all other necessary documents to secure six contracts with seven ethnographers for the GQ study. Staff handled all the initiation and followup procedures for background checks, Title 13 training and security badges for the contractors. Staff organized a two-day orientation and training conference for the ethnographers who will be conducting the ethnography study of group quarters in 2010. Staff reviewed and provided comments to the final proposals submitted by each of the ethnographers. Staff monitored and provided continual guidance to the ethnographers regarding their field activities throughout the quarter. In February, the lead researcher contacted and met with administrative staff of

one of the study sites to prepare for the CCM-like study. Staff designed and pre-tested a new questionnaire to conduct a CCM-like study for the university housing population. Staff submitted the final instrument to Administrative & Customer Services Division for imaging. Staff submitted an Office of Management and Budget clearance package to obtain clearance for all post-enumeration interviewing and focus group activities. During the third quarter, Staff contacted respective universities' points of contact in April and May and arranged to conduct the CCM-like study using a mailout/mailback self administered survey mode. Questionnaires were printed by the Administrative Customer Services Division and assembled in-house by staff. Approximately 470 questionnaires were mailed to Site One and questionnaires were distributed by the university residence hall staff. Staff made arrangements to visit two Greek houses in May to deliver the remaining 40 questionnaires. Staff completed data collection. Staff recruited 10 students at the first research site to participate in a post-survey debriefing focus group, which was conducted on site in May. The audio recording of the focus group was transcribed and a focus group report was drafted. The lead researcher met with all the ethnographers three times via telephone conference in April, May, and July and began reviewing American Community Survey observation reports ethnography field note reports and decennial enumeration observation reports. During the last quarter, staff began to code data collected from the pilot CCM test and to review draft observations, interview, focus groups and final reports submitted by ethnographers.

Staff: Anna Chan (x38462), Matthew Gore, George Higbie, Temika Holland, Stephen Lubkemann

F. 2010 Census Language Study (CPEX)

Description: We will conduct systematic/structured observations of Nonresponse Followup (NRFU) interviews in areas with heavy concentration of linguistically isolated (LI) households from various national origins. The aim of this research is to: (a) observe how enumerators in the 2010 Census environment approach LI households, (b) observe what measures are taken by enumerators to collect the required census data from these households, and (c) based on our observations, determine what changes, if any, are needed to improve the conduct of in-person interviews with LI households.

Highlights: During FY 2010, we developed the study plan for this research, worked on study design and theoretical framework to guide the research. We conducted research to identify ethnographers to lead seven language teams to carry out the observation. We successfully assembled a multilingual research team with seven sub-teams consisting of 23 ethnographers in seven languages (Spanish, Chinese, Korea, Russian, Vietnamese, Arabic, and Portuguese). We provided two-day training for the contracted ethnographers and

coordinated with the Field Division to make arrangements for the ethnographers to observe NRFU interviews in the target languages. Ethnographers in all seven languages observed NRFU interviews for two weeks in the designated areas of high concentration of speakers of the target languages. The research sites include California, Illinois, Ohio, Michigan, Massachusetts, New York City, and Greater Washington DC area. Each language team observed 70-170 interviews in the target language as well as in English. They also conducted debriefing interviews with respondents and enumerators to understand how they perceived the U.S. Census and to obtain information of the efficiency of the 2010 Census language materials. The contracted ethnographers are working on final reports to document main findings of the study.

Through this ethnographic study, we learned that there are linguistic as well sociocultural barriers to enumerate non-English speaking populations. Linguistic barriers include: 1) a lack of standard translation of the NRFU questionnaire. Enumerators had to translate the census questions on-the-fly which posed a thread to data completeness and data quality; 2) a lack of systematic assessment of bilingual enumerators' language proficiency in both speaking and writing. While some bilingual enumerators were effective in interacting with non-English-speaking respondents, some were not due to the low level of proficiency in target languages. There is a need for a standardized procedure to assess and verify the language competence of bilingual enumerators. Sociocultural barriers include a lack of awareness of the census purpose and operation, fear and distrust of the government, low level of literacy, low education attainment, immigration status, gender, and identity issues. We are currently working on analyzing findings and developing recommendations for the 2020 Census.

Staff: Yuling Pan (x34950), Steven Lubkemann, Marissa Fond

G. 2010 Census Behavior Coding Evaluation

Description: In order to learn how well census enumerators/interviewers ask, and how well respondents answer, census questions, behavior coding studies will be conducted for all major interviewer-administered instruments (NRFU, CFU, CCM PI) in 2010. The purpose is to calibrate how well survey instruments are administered by interviewers and to identify problems with how interviewers ask and respondents answer questions. By conducting behavior coding for these three interviewer-administered instruments, this study will tell us whether census questions are being asked as intended and will identify problems with the questions and with interviewer training. This study can further help the Census Bureau interpret apparent disparities in data that may arise between different operations.

Highlights: During FY 2010, staff completed study plans and schedules for the 2010 Behavior Coding of NRFU,

CFU and CCM PI that were presented to and approved by senior staff. These plans were finalized and made public. Staff coordinated data collection and analysis with other divisions. Staff prepared training materials for the Nonresponse Followup (NRFU) behavior coding and conducted training at the Tucson Telephone Center. Coding was completed at the telephone center and analysis will begin early next fiscal year. Staff also verified the recordings data from CFU that will be used in coding next fiscal year.

Staff: Jennifer Hunter Childs (x34927), Nathan Jurgenson

H. Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups
Description: In this project, Census Bureau researchers are contracting with outside ethnographers to conduct comparative ethnographic research on enumeration methods and coverage in nine race/ethnic communities during Census 2010. The aim is to identify ways to improve census enumeration methods and coverage for race/ethnic populations, some of which have been categorized as hard-to-enumerate groups in previous censuses. We hope to contribute to knowledge on a critical census issue: persistent differential miscounts of some minority populations across censuses. This field study involves accompanying enumerators to observe, tape, and debrief respondents during three Census 2010 operations involving personal visit census data collection. We will identify and explore three sets of issues affecting the completeness and accuracy of the census: 1) enumeration methods, 2) questionnaire issues, and 3) residence rule/coverage issues. We will identify types and sources of possible coverage error, and characteristics of households and persons at risk of coverage error, as well as address questions of whether possible coverage errors are more commonly found in complex, than non-complex households, and explore similarities and differences across sites.

Highlights: During FY 2010, we developed the official project study plan in the fall of 2009 and revised it in response to reviews from our seven critical reviewers, and obtained approval from the Census Integration Group. We recruited ethnographers at professional association meetings for anthropologists, sociologists, and other social scientists and also on those associations' listserves, with special attention focused on researchers who had recent prior fieldwork experience working with our target race/ethnic groups in U.S. communities. We reviewed the qualifications of more than 120 potential ethnographers and identified 16 ethnographers matched to race/ethnic groups with which they had unique research experience.

We wrote contract documents and worked with Acquisitions, Security, ITSO, LTSO, and the Policy and Privacy Offices to get the contracts in place in time for

the researchers to observe the field data collection at the height of each operation.

We worked closely with Decennial Statistical Studies Division (DSSD) sampling experts to get sampling lists for use in research site selection. We also worked with special data from a prior communications study to identify tracts with high Planning Data Base hard-to-count scores and high proportions of our target race/ethnic groups to aid in selecting the race/ethnic research sites for this study. We worked with Field Division staff to get approvals from 8 Regional Census Centers Directors and relevant Local Census Offices to conduct the observations in their areas during each of the operations. We wrote contract documents and worked with Acquisitions, Security, ITSO, LTSO and the Privacy and Policy Offices to get all contracts awarded in time to observe the height of the three operations. We developed custom training materials, checklists, and enumerator handouts for each operation, as well as guidelines for telephone follow-ups, and standard household interview and coverage typologies and table shells to facilitate comparisons across sites. We conducted six training sessions. The team leader went out to observe interviews during two of the three operations. All researchers have finished their observations during Census field operations and are currently working on transcriptions and reports.

We worked with LTSO and ITSO to explore whether it would be possible to develop a new IT policy that would allow us to issue special long-term loaner laptops and flash drives to our contractors so that they could transcribe Title 13 Census Personally Identifiable Information (PII) from tapes while working in their homes (which is not permitted for regular Census Bureau employees). We did obtain approval to do this with specially configured laptops with minimal software and no internet connectivity that would be returned at the end of the project and wiped clean. We arranged for researchers to safely transmit Title 13 materials by downloading them onto a secure Safeboot flash drive and FEDEXing it to the Census Bureau where project staff can download it onto secure drives and FEDEX the flash drive back to the researcher. This agreement has now been written into the overall Cen08 IT Security Policy for the Census Bureau. It allows approved employees and contractors to use long-term loaner laptops to process and store Title 13 data and use secure encrypted Safeboot flash drives to transfer data from the loaner laptop to Census secure systems. In addition, we negotiated an agreement between LTSO and our division that opens up this opportunity for other research projects that involve collection of Title 13 PII data in the field. Researchers on two other projects have already made arrangements to obtain long-term loaner laptops under this new policy.

Staff: Laurie Schwede (x32611), Rodney Terry, Lorraine Randall, Matthew Clifton, Greg Bulmash, Christina Cooper.

I. Explaining How Census Errors Occur through Comparing Census Operations History with Census Coverage Measurement (CCM) Results

Description: The goal of this project is to help us understand what sorts of errors tend to be associated with the different Census operations, especially for persons and housing units removed from the census. We will compare Census files to the CCM results for a subsample of CCM areas. This comparison is intended to help find patterns of errors in Census operations and provide insights into ways to avoid these errors.

Highlights: During FY 2010, staff completed several drafts of the study plan and sent them to the Decennial Management Division (DMD). One draft was sent to DMD in response to a request from the General Accounting Office for study plans for several evaluations. The study plan has been reviewed by critical reviewers and some additional interested parties and has been approved by the Census Integration Group. Staff also prepared draft schedules for the evaluation and sent them to DMD, answered a DMD questionnaire on data sources for the evaluation, and continued outlining the file requirements for the evaluation.

Staff: Michael Ikeda (x31756), Mary Mulry

J. 2011 Relationship Survey

Description: Recent changes in the American legal and social landscape with respect to family composition, relationships, and same sex marriages have potential impacts on the content of the relationship and marital status questions in Census Bureau censuses and surveys. This project involves a program of research and testing that will guide the development of revised questions. Exploratory focus groups will be conducted across the United States with members of cohabiting couples to collect qualitative information about alternative terms, definitions, categories, and/or questions that most accurately measure relationship status and partnership situations. Cognitive interviews will be conducted to evaluate questions developed from the results of the focus groups. The revised questions will be subjected to further testing and evaluation before being implemented by the Census Bureau.

Highlights: During FY 2010, staff assisted in the planning, contracting, and observation of focus groups conducted nationwide with members of cohabiting same-sex and opposite-sex couples. The results of the focus groups showed that: 1) participants interpreted the request for relationship and marital status on a Federal form as requesting a legal status; 2) participants who had been married anywhere reported that they were "husband/wife" and "now married" regardless of whether the marriage was recognized in their state of residence; 3) in states where marriage was allowed, married and unmarried same-sex participants had no trouble in answering the relationship and marital status questions; and 4) very few of the participants reported that they were "husband/wife" or "now married" if they were not legally married.

Staff: Terry DeMaio (x34894), Nancy Bates (DIR)

1.9 AMERICAN COMMUNITY SURVEY(ACS) (Decennial Project 5385060)

A. ACS Missing Data and Imputation

Description: This project undertakes research and studies on missing data and imputation for the American Community Survey and aims to impute missing socioeconomic data in the National Assessment of Educational Progress (NAEP) data files using Census long form and American Community Survey (ACS) data.

Highlights: During FY 2010, we collaborated with National Center for Education Statistics staff on imputing students' socioeconomic data into the NAEP and ECLS data. We had previously provided files with Census and ACS data donor cells corresponding to each NCES recipient. Because donors must be used repeatedly to find matches for most sample NAEP and ECLS data, NCES decided to not used the matching scheme but to find donors using only geographical distance. We provided programming support for this approach and created the Census long form SEDF and ACS databases needed for estimating socioeconomic status index using the contractor's new requirements.

Staff: María García (x31703), Yves Thibaudeau

B. ACS Group Quarters (GQ) Item Imputation and Micro Data Disclosure Avoidance Research

Description: American Community Survey group quarters microdata and tabulations are protected from identity disclosures via synthetic data methods. This project coordinates staff in our division, Decennial Statistical Studies Division (DSSD), Population Division (POP), and Housing and Household Economic Statistics Division (HHES) to generate production code (in the R language) for this purpose. Staff will also ascertain the effectiveness of using synthetic data methods as an alternative to hot deck allocation in ACS group quarters.

Highlights: During FY 2010, staff generated a new package in R, tentatively called PEP, to generate edit-compliant synthetic data for American Community Survey. PEP integrates previously-used modeling strategies from past ACS years with additional software to improve workflow and insure edit consistency (notably we adapted Bill Winkler's DISCRETE program into C for this purpose). Having tested the software on both Statistical Research Division and American Community Survey Office production systems, we successfully used PEP for statistical disclosure control (via synthetic data methods) for the 2009 ACS group quarters sample. Future research will focus on integrating multiple

imputation methods into PEP and testing these methods on the ACS.

Staff: Laura Zayatz (x34955), Paul Massell, Rolando Rodríguez, Jason Lucero, Asoka Ramanayake, Lisa Singh, Bimal Sinah, Tapan Nayak

C. ACS Applications for Time Series Methods
Description: This project undertakes research and studies on applying time series methodology in support of the American Community Survey (ACS).

Highlights: During FY 2010, staff organized a new project to examine discrepancies between multi-year estimates and related single year quantities. Staff formulated a testing framework, implemented corresponding statistical procedures in SAS, and presented preliminary results to client. Upon receiving feedback, staff undertook additional computations and completed a first draft of project for the client. After receiving criticism on methodology, staff devised a new nonparametric methodology and presented results to the client, who was satisfied with the results.

Staff: Tucker McElroy (x33227), Natalya Titova, Chaitra Nagaraja

D. ACS Variances
Description: Work under this heading this year concerned two research projects: (i) Completion of a project providing design-based superpopulation consistency theory along with linearized variance formulas for variances of estimators of totals from complex surveys to which replication methods like BRR can be compared. (ii) Development of a method of simultaneous nonresponse adjustment, calibration to achieve population controls, and weight smoothing or truncation.

Highlights: Work under this heading during FY 2010, concerned two research projects: (i) Development of a method of simultaneous nonresponse adjustment, calibration to achieve population controls, and weight smoothing or truncation, and (ii) Comparison of alternative methods of estimating variances for complex survey estimates with a new method `hybrid' estimates of cross-classified population totals containing some design-based cell size estimates, for some cross-classified categories, along with model-based conditional-probability estimates within those catgories. Both projects were the culmination of 1-2 years' work, involving writing up theoretical research, developing computational illustrations on Census data, and revising technical report and journal article manuscripts.

Staff: Eric Slud (x34991), Yves Thibaudeau

E. ACS Small Area Estimation for Selected Characteristics
Description: This project aims to propose, develop, and evaluate small area estimation methodologies to produce ACS estimates for selected characteristics in geographies with small populations. The characteristics of initial interest are unemployment, income, and poverty.

Highlights: During FY 2010, an interdivisional group was formed to address this topic. Meetings were held to present and discuss background material. The group was divided into two subgroups due to the different requirements for small area estimates of Group Quarters and the total population. A staff member was selected as coordinator of the second group. That member reviewed documentation to determine what variables are available for use from the StARS and the LEHD administrative record systems. Only two meetings of the group were held due to demands of the following project.

Staff: Lynn Weidman (x34902)

F. ACS Small Area Estimation for Group Quarters (GQ)
Description: This project aims to propose, develop, and evaluate small area estimation methodologies to produce ACS estimates for the GQ population (totals and characteristics) for substate geographies, including counties, places, block groups, and tracts.

Highlights: During FY 2010, an interdivisional group was formed to investigate methods to be applied with ACS estimation processing in the spring of 2011. The goal of the methodology is to impute person records into enough GQ facilities not in sample so that: (a) the 1- and 3-year data files have GQ data present in every county with a GQ and (b) the 5-year data files have GQ data present in every tract with a GQ. Furthermore, the method must fit into the existing ACS system for producing estimates and variance estimates.

The initial testing and evaluation of the methodology was carried out using simulated samples based on the Census 2000 GQ records that contain only short form variables. A single GQ population was constructed as the sampling frame for each of 25 5-year samples and their imputation, so that the combined sampling and imputation variance could be estimated without being confounded with variability among different GQ populations. The sample simulation included assigning a number of people to each GQ in each year. These numbers were assigned randomly using a method that was based on the relationship between expected size and actual size determined from analysis of data on ACS GQ sampling universe and weighting files. The evaluation looked at the properties of estimates of demographic characteristics across these samples via various statistics. Similarly, the changing size of the GQ frame by major type across the five years was also simulated. The procedure for doing so was determined by investigating ACS sampling files, matching them as necessary across years, analyzing the results, and applying a function to select GQs for deletion.

Software for simulating the population was redesigned, rewritten, and tested, and the 25 samples were generated. As part of the population simulation, the changing size of the GQ frame by major type across the five years was also simulated. ACS sampling files were investigated, matched as necessary, and the results analyzed to determine the effective decrease in frame size, by identifying GQs deleted and added in each year. Then a procedure to select GQs for deletion to meet this changing size was developed and software written to apply it.

A staff member proposed the use of nearest neighbor hot-deck imputation and four alternatives for searching for donors, which varied by nearest neighbor definition and the lowest level of geography searched for donors, were proposed. Software was written to carry out these four proposed methods as well as to calculate the evaluation statistics. A project proposal was drafted, revised, and accepted by the ACS Research and Evaluation Team. Imputation by four methods was completed for data simulated from Census 2000 GQ population. Weighting was performed and estimates, standard errors, and graphs summarizing comparison statistics were produced for demographic characteristics by major GQ type by geography (state, county, tract) for the four imputation methods and the observed data only. These results were examined and summaries prepared, along with a division research report. One result was the discovery that, most notably for prisons, the correct proportion of males and females were sometimes not imputed. This is because the fact that most prisons contain persons of only one sex was not accounted for. This shortcoming was fixed for the next evaluation.

The second evaluation uses ACS data from years 2006-09 along with a derived 2010 sample data file. Many more variables than just demographics will be evaluated. Meetings were held to discuss and documentation drafted to describe files, records, and variables to be used in the imputation and weighting. For the year 2010, a relationship between maximum size from the Census 2010 GQ universe file and actual size as observed by ACS was modeled for the purpose of method evaluation. For each year in 2006-10, a relationship between the expected ('06-'09) and census observed ('10) and the ACS observed GQ sizes, including deleted GQs, was developed to apply to non-sampled GQs to adjust their expected populations. Imputation and related programs were modified to take account of the different data file format than was used for the simulation. Imputation was carried out and calculation of the evaluation statistics has begun.

Two methods were proposed to approximately estimate the effect of imputation on variances of GQ estimates. These two methods will be compared and evaluated.

Staff: Lynn Weidman (x34902), Chandra Erdman, Patrick Joyce, Chaitra Nagaraja

G. ACS Data Issues
Description: Various issues related to the quality and presentation of ACS estimates were discussed and investigated by small interdivisional teams or division staff. The goal of these investigations was to make recommendations to aid in resolving the issues.

Highlights: During FY 2010, data analysis was completed for the report "Developing Guidelines Based on CVs for when Three-Year Estimates Can Be Used Instead of Five-Year Estimates in the American Community Survey." This extends previous work which looked at ACS one-year estimates.

At the request of Decennial Statistical Studies Division American Community Survey staff, a staff member in our division reviewed material on statistical methods of multiple comparisons and drafted a technical document describing what procedures should be used for various comparisons and tests using ACS data. Several examples of use of the methods were prepared and others of interest are needed from subject matter experts.

The current rule developed by our division for identifying outliers in data review is based on a regression method which looks at each review measure in succession, using the remaining measures as predictors. However, there are two statistical issues with this approach: (a) the problem of "masking" (using outlier values to construct the regression model in order to find those outliers) and (b) ignoring the strong correlations among measures. Research has begun to develop an improved method of detecting outliers that addresses these issues. Literature on several methods has been investigated, including outlier detection using Mahalonobis distances and minimum ellipsoid volume methods. A staff member is currently investigating similar approaches which utilize measures from a geography as a single vector as opposed to a response measure and an explanatory measure vector.

Staff: Lynn Weidman (x34902), Chaitra Nagaraja

1.10 AMERICAN COMMUNITY SURVEY (ACS)/METHODS PANEL (Decennial Project 5385095)

A. ACS Language Research
Description: This project provides technical and research support for addressing language issues in American Community Survey (ACS) data collection instruments and supporting documents. Staff members serve on inter-divisional working groups and provide consultation and technical support in the design and development of language research for the ACS.

Highlights: During FY 2010, staff members continued to be active members of the ACS Language Team and the

ACS Content Council. We worked closely with the ACS Language Team to plan and review Statements of Work for three new projects: 1) Translation review and cognitive testing of translated ACS Language Assistance Guides in Chinese and Korean, 2) Project D, Part II, III and IV, which is the cognitive testing of the remaining half of the Spanish CATI/CAPI instrument, and 3) cognitive testing of the ACS Regional Office (RO) letters in five languages (Project C). Staff worked as Census Bureau research analysts, providing technical guidance for the contractor (RTI). Specifically, we helped review and provided comments on the recruitment plans, cognitive interview protocol development and translation, interview summaries, and alternative translations. In addition, staff translated and reviewed the Spanish version of the ACS internet instrument.

We also developed analytical frameworks encompassing theories in sociolinguistics and psychology to provide guidance for the contractor in summarizing interview data and reporting findings. Staff conducted training for the contractor's interviewers on the ACS Language Assistance Guides project and Project C. We developed two types of coding schemes for these two projects. One is a coding scheme using a sociolinguistic framework to classify translation issues identified in the cognitive testing process. This coding scheme enables systematic review of issues and provides possible solutions to those issues. The coding scheme has been implemented in the ACS Language Assistance Guide project. And the other is a coding scheme that uses principles of influence from a psychological perspective to code cognitive testing recruitment strategies and the interpretation of messages conveyed in ACS CAPI RO letters. Additionally, we have introduced a method (also a type of coding) of collecting respondents' survey experiences and reporting respondents' behaviors in the consent procedure as indicators and supporting evidence for analyzing respondents' verbal responses in the multilingual cognitive interviewing event.

Staff also provided technical consultation on a discourse approach to re-structuring survey letters in a culturally appropriately manner. Findings from the cognitive testing of ACS RO letters suggest that fixing translation problems at the lexical and sentence levels does not significantly improve the translation quality or enhance respondents' understanding of the key messages conveyed in the survey letters. Translated survey letters need to be structured according the discourse convention of the target language. This requires exploration of a new methodology that departs from the traditional translation approach. Staff innovatively designed a methodology incorporating discourse analysis methods to re-structure survey letters in Chinese for Round 2 testing of the ACS RO letters project. The Round 2 findings showed significant improvement of Chinese respondents' understanding of the letters. Due to staff recommendation, the ACS language team decided to conduct a new research project to adopt the method of

discourse approach to re-structuring ACS RO letters in four target languages (Chinese, Korean, Russian and Vietnamese) and further test it. Staff contributed to the Statement of Work for this new research.

In addition, staff worked to support the upcoming ACS Content Test, which will test Spanish and English question wording. During this fiscal year, staff worked to support the development of an automated CARI behavior coding system which will be used in the Content Test. Staff from our division contributed to contract requirements and planning, reviewed contractor deliverables, tested software and planned for the upcoming behavior coding project.

Staff: Yuling Pan (x34950), Patricia Goerman, Rodney Terry, George Higbie, Marissa Fond

B. ACS Data Reliability Indicator Project
Description: The usability team designed a series of usability evaluations of a new method of displaying the ACS data tables. The new feature to be tested was a color-coded indicator of the reliability of the data. The purpose of the testing was to examine how well the data-reliability indicator worked for users (especially as compared to the current ACS data tables without the indicator) and to identify any problems that actual users might have with the data tables. The data reliability indicator was based on the Coefficient of Variation (CV), which is defined as the standard error of an estimator divided by the estimate. Another purpose of this testing was to examine whether users would notice and use the Margin of Error (MOE) when answering questions about the estimates from the table. This second testing goal was based on the observation that although the MOE is currently provided with each estimate, the MOE is routinely ignored by ACS data users.

Highlights: During FY 2010, staff conducted a third round of usability testing, delivered a report of the result to the sponsors, and briefed the Methodology and Standards council on the results.

Staff: Kathleen Ashenfelter (x34922), Victor Quach

C. ACS Messaging Project
Description: The purpose of this project is to develop and test new messages on ACS letters and a brochure to alert ACS respondents in 2010 that they are required to respond to both ACS and census questionnaires. In 2000, ACS response rates were affected by the 2000 Census environment. Until March 2000, ACS response rates rose as a result of census publicity, but they fell for the rest of the year after respondents also received their census forms, particularly around Census Day. The aim of this project is to try to avoid these drops in response rates in 2010 by informing ACS respondents that they will be receiving both forms and need to complete both.

Highlights: During FY 2010, we finalized the Phase 3 project report, titled "Cognitive Testing Results for the ACS Field Flyer for Use in 2010: ACS Messaging Project: Phase 3 Final Report," sent it to the sponsor, and have had it approved for posting on our division's website. We are incorporating some additional text on cognitive testing of the letters into the near-final Phase 1/Phase 2 report to provide context for the results of the 2010 ACS split-panel test of experimental letters and envelopes that will be released early next year. Our report is entitled, "Cognitive Testing of Experimental American Community Survey Envelopes and Letters for 2010: ACS Messaging Project Phases 1 and 2 Report." We found clear and consistent patterns of respondent behavior across three rounds of observations across cognitive testing: 1) just 32 to 43% of all respondents found and appeared to read the package cover letter and 2) about three quarters of all respondents took the inserts out with the questionnaire on top, but were much less likely to find and read the letter, while those taking out another insert first had the opposite pattern. No such patterns were observed with the 2010 Census package. We suggested changes in the ordering and number of questionnaire package inserts that might increase the likelihood that respondents will read the cover letters, possibly increasing response rates. These suggestions could be applied beyond our original target year of 2010.

Staff: Laurie Schwede (x32611), Anissa Sorokin

D. ACS Internet Testing – Usability Input
Description: The usability team is leading our division's contribution to the development of an online instrument for the American Community Survey. The multi-year project will consist of several rounds of usability testing of prototypes of the internet instrument. Staff will provide test plans and formal reports for each round of testing as well as provide input at regular team meetings during the development of the ACS online instrument and its subsequent field testing.

Highlights: During FY 2010, staff conducted cognitive and usability testing on the mailing materials associated with this project (i.e., letters, forms, and postcards). Staff conducted three rounds of usability testing and are preparing for a fourth round in early FY2011. Staff are participating in the ACS Internet Testing Team, which will continue to meet into FY2011.

Staff: Kathleen Ashenfelter (x34922), Temika Holland, Victor Quach, Elizabeth Nichols, Matthew E. Jans, Sabin Lakhe

E. ACS Internet Testing – Cognitive Input
Description: The Cognitive Lab is participating in the development of an online instrument for the American Community Survey. The multi-year project will consist of questionnaire development and several rounds of cognitive and usability testing of prototypes of the internet instrument in both English and Spanish. Staff

will provide questionnaire development and mode consistency expertise at regular team meetings during the development of the ACS online instrument and its subsequent field testing. Staff will also participate in the joint cognitive/usability testing sessions led by the Usability Lab.

Highlights: During FY 2010, staff provided expert guidance on question wording and mode consistency at weekly development meetings. Staff prepared experimental versions of roster screens for the instrument for cognitive and usability testing and assisted preparing the cognitive and usability test plan. Staff drafted a CATI Reinterview survey called the Attitudes and Behavior Study (ABS). Questions in this survey will ask about Internet use/non-use, privacy concerns, and effectiveness of mailing materials. These questions will vary slightly across the three respondent groups: Internet respondents, mail respondents, and nonrespondents. In addition, staff prepared a Spanish translation of the Internet instrument and continues to provide guidance on question wording in both English and Spanish.

Staff: Jennifer Hunter Childs (x34927), Elizabeth Nichols, George Higbie

F. ACS Internet Test Experimental Design Team
Description: Staff is contributing methodological expertise and input to the development of the experimental design that will be used for the field testing of the ACS Internet form Census Quality Survey (CQS) that is currently being developed and is planned for early 2011. The design of this methodology includes such considerations as sampling, pre-notification letters, mailing schedule, panel design, and planned analysis of the results.

Highlights: During FY 2010, staff completed two rounds of usability testing on the internet instrument and the mailing materials associated with the CQS. The CQS was completed and went live during FY 2010. Staff will be performing an analysis of the paradata collection from this survey in FY2011.

Staff: Kathleen Ashenfelter (x34922), Temika Holland, Victor Quach, Elizabeth Nichols, Sabin Lakhe

G. Iterative Testing of the American Community Survey Web Site
Description: The American Community Survey (ACS) Web site is making major re-designs to its interface. The Usability Lab has been asked to participate in the iterative user-centered design development of the site. This effort encompasses the full spectrum of user-centered design activities, from iterative low-fidelity paper prototype testing to high-fidelity testing with a working prototype. The purpose of the testing is to identify usability issues. Recommendations made to resolve the issues are intended to improve the usability of the Web site for all users.

Highlights: During FY 2010, staff conducted iterative usability testing on the ACS Web site. Initial testing was on low-fidelity prototypes where novice and more expert users gave feedback on locating information on the new design of the home page. Staff identified areas of the interface that needed to be modified before the next round of testing took place and once ACS Office staff had developed the new screens, staff conducted two rounds of usability testing on a higher fidelity prototype with novice users. Our findings highlight some of the following: the main page was the dominant source for information and participants did not efficiently use the top navigation. We recommended that the labels presented on the top navigation be more applicable to common tasks novice users would come to the site to complete. In addition, most participants believed that the ACS form could be filled out online, which is not true. We recommended placing information on how to respond to the survey more prominently on the Web site so that it is easily recognized by users. Participants also commented that information about penalties should be grouped together with information about whether the survey was mandatory. Staff presented the findings and presented the quick report to the client and is finalizing the report for the division series.

Staff: Erica Olmsted-Hawala (x34893), Jennifer Chen, Temika Holland

H. ACS Content Test Pilot CARI Behavior Coding
Description: Staff began work leading a behavior coding project in support of the 2010 ACS Content Test. This project is also a pilot project that will utilize and test the Census Bureau's new CARI behavior coding system. The project involves the testing of both Spanish and English versions of new and/or revised content that is under consideration to add to the ACS. Staff provided support for the cognitive testing of the new content and is now supporting the next step in the project: the behavior coding of field test results.

Highlights: In FY 2010, staff participated in the development and testing of the software that will be used for the ACS pilot project. In addition, staff participated in bi-weekly ACS Content Test team meetings to assist in project planning and met with project sponsors to create a detailed project schedule.

Staff: Patti Goerman (x31819), Joanne Pascale

1.11 CURRENT POPULATION SURVEY (CPS) / ANNUAL SOCIAL AND ECONOMIC SUPPLEMENT (ASEC) TABLES (Demographic Project TBA)

Description: Staff provided technical consultation services and programming support for the redesign and

content of SAS programs that produce the table packages for the 2007 Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) that will feature information at the national and regional levels for special population/topics.

Highlights: During FY 2010, staff updated and tested table shell generating software in Visual BASIC. Staff coached personnel in the Population Division to generate more than one hundred new 2009 race, older, and gender tables accurately and efficiently. Completion will become a joint effort of staff in our division and staff in the Demographic Statistical Methods Division.

Staff: Aref Dajani (x31797), Tom Petkunas

1.12 DEMOGRAPHIC SURVEYS DIVISION (DSD) SPECIAL PROJECTS (Demographic Project 0906/7374)

A. Data Integration
Description: The purpose of this research is to identify microdata records at risk of disclosure due to publicly available databases. Microdata from all Census Bureau sample surveys and censuses will be examined. Potentially linkable data files will be identified. Disclosure avoidance procedures will be developed and applied to protect any records at risk of disclosure.

Highlights: During FY 2010, using the Matcher, staff linked files from the Publically Purchased Lists to the PUMS file (Public Use Microdata file from the ACS) to look for outliers based on similar blocking criteria. The Matcher was modified to produce outliers (block sizes were less than five) and not output every match. The outliers should in theory be easier to attack. Staff set up spider programs to gain additional knowledge of individuals in the PUMS file. Staff created a data warehouse and identified a series of attributes such as state, income, etc. After examining the distinguishing power of each variable, each attribute combination with fewer then 10 individuals in the ACS PUMS file was flagged. The records were matched to the wholesale data. Combinations of attributes that were linked in both files with fewer then five combinations are outliers, flagged for attack. These results are confidential and are documented in a memorandum.

Staff: Ned Porter (x31798), Lisa Singh, Rolando Rodríguez

B. Using Survey Paradata to Manage Surveys in the Field and Estimate Survey Error
Description: This project seeks to understand how paradata (survey process data) are currently used and where they are stored throughout the Census Bureau, particularly in the Demographic and Field Divisions. The broader goal of the project is to modernize project

management through the use of graphical representations of paradata that are displayed in interactive, real-time dashboards. This improvement in paradata access will allow managers to make quick and better project management decisions and will equip the Census Bureau with the ability to carry out responsive design.

Highlights: During FY 2010, staff worked with the Technology Management Office (TMO) to plan a process of receiving extracts of data from TMO (Regional Office Sample Control (ROSCO) and Cost and Reporting Management Network (CARMN) systems) so that we can reproduce and program automatic production of cost and progress reports currently created by hand in the Field Division. We plan to expand the use of these data to demonstrate the possibility of a flexible cost and reporting system that is usable for all stakeholders in the survey data collection process. As an extension of this work, two of the team members have been involved in the operational efficiency business case on the use of paradata to contain costs and improve data quality. Another team member made significant progress on the use of control charting and statistical process control principles in understanding paradata from the National Health Interview Survey (NHIS) Computer Assisted Personal Interviewing (CAPI) systems. Staff developed additional paradata research projects with National Crime Victimization Survey (NCVS) (screener paradata and victimization reports) and NHIS (CAPI paradata, primarily question timings; comparing Quality Control (QC) reinterview to the Performance and Data Analysis system (PANDA) for insuring data quality). As an outgrowth of this general paradata work, an SPC-oriented interest group (dubbed "Club Deming") has formed with members from the Demographic Studies Division, American Community Survey Office, Data Integration Division, and our division. This group plans to meet monthly and discuss uses of Statistical Process Control (SPC) around the Census Bureau. We continue to discuss long-term data transfer plans with TMO with the goal of creating a data warehouse for the storage of paradata and cost data. Sam Highsmith has lead this effort.

Staff helped coordinate a paradata workshop to be held at Census Bureau headquarters in October, 2010.

Staff: Matt Jans (x36724), Kathy Creighton (Contractor, DSD), Chris Laskey (DSD), Cheryl Landman (DSD), Chris Stringer (DSD), Ben Duffey (DSD)

1.13 QUICK TURNAROUND PRETESTING OF HOUSEHOLD SURVEYS
(Demographic Projects 1465001)

Description: 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.

A. Rental Housing Finance Survey
Description: This project involves cognitive testing of a new questionnaire that collects information from property owners and managers of multi-family housing units sponsored by the Department of Housing and Urban Development (HUD).

Highlights: During FY 2010, staff held several meetings with staff from Housing and Household Economic Statistics Division, Demographic Statistics Division, and HUD to discuss the scope of the project. Pretesting of the Rental Housing Finance Survey questionnaire was delayed by finalizing the Interagency Agreement and higher priority to the testing of the American Housing Survey questionnaire.

Staff: Terry DeMaio (x34894), Jennifer Beck

B. National Crime Victimization Survey (NCVS)
Description: The NCVS asks respondents to report the race of the offender for any crimes committed against them. Currently the response categories (White, Black, Other) do not match the categories included in Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. This project includes a review of survey methods and social scientific literature relevant to identification of the race of offenders, as well as designing and implementing experimental research on this topic.

Highlights: During FY 2010, staff worked with Demographic Surveys Division and Bureau of Justice Statistics (BJS) staff to develop a research plan to revise and evaluate new questions to identify the race of the offender of violent crimes. We wrote a review of social psychological literature focusing on memory and perception as they apply to the identification of proxy racial information. The review highlighted the difficulty of this task in terms of the accuracy of the identification. In addition, we developed a new series of questions, consistent with the OMB standards, to identify the race of some offender when the respondent reports that single and multiple offenders were involved in the crime. We also developed a plan to conduct respondent debriefings with respondents who are administered the new race-of-offender questions in outgoing NCVS rotation groups. This research plan, along with the literature review, will be submitted by BJS to OMB as documentation of their progress in attempting to meet the OMB standards. We also conducted research on the 2011 School Crime Supplement. We conducted a literature review on the topic of cyberbullying related to a series of survey questions. We also conducted cognitive pretesting of the questionnaire with respondents aged 12-17 in the Washington, DC metropolitan area. The results showed that respondents had several types of problems with a new question asking about the amount of time they spend

on school-sponsored extracurricular activities. First, they interpreted this phrase to include such activities as school fundraisers and community service activities. Second, some respondents included only things that were related to education, and as a result, did not include hours spent participating in sports activities. Third, some respondents included only after-school activities, while other included activities that took place during the day. In the end, responses captured information about a different range of activities than the specific school-sponsored activities included in the previous question. Testing of changes to the cyber-bullying questions revealed that respondents used and/or were familiar with newly-added social networking sites Formspring and Twitter, but were not familiar with Second Life. Alternative wordings "online communications" and "online activities" were tested to determine which was most appropriate for inclusion in the questionnaire. Respondents interpreted these terms quite differently. They interpreted the terms "online communication" to generally include socializing using social networking sites and instant messaging or chatting, while they defined "online activities" in terms of gaming. The sponsor, the National Center for Education Statistics, agreed to include most of the recommended revisions in the supplement when it is fielded in January 2011.

Staff: Terry DeMaio (x34894), Jennifer Beck

C. Development of the CARI Behavior Coding System
Description: This project involves consultation with staff from several divisions (DSD, DSMD, DSSD, TMO, FLD, ACSO, and other divisions) and RTI International, which is the contractor developing the behavior coding component of the CARI system. This program will enable interviews, or snippets of interviews, to be tape-recorded as they are conducted, and will allow behavior coders to listen to sound files and see the interviewer's entries as they code the interactions between the interviewer and respondent using an automated system. The system will also output data to SAS and Excel for behavior coding analysis.

Highlights: During FY 2010, staff participated in joint application design sessions for the behavior coding component, biweekly working group meetings, and a behavior coding orientation session, and provided comments on numerous versions of the user feedback component and e-training for the behavior coding component. We also participated in the Systems test, Verification Test, End-to-end Systems Test, Load Test, and Integrated Verification Test.

Staff: Terry DeMaio (x34894), Patricia Goerman, Joanne Pascale, Jennifer Beck, Matt Jans, Kathleen Ashenfelter

D. American Housing Survey Modules
Description: This project involves cognitive testing of two new AHS modules (Healthy Homes and Home Accessibility) and revisions to the AHS Mortgage Module. These modules are sponsored by the Department

of Housing and Urban Development and scheduled for implementation in 2011.

Highlights: During FY 2010, staff participated in development of the questionnaire and discussions required to get OMB approval for the testing. Cognitive interviewing for the project is currently underway.

Staff: Terry DeMaio (x34894), Kathleen Denny, Andrea Schwanz

E. Survey of International Air Travelers
Description: This project involves participation in pretesting activities for surveys that are not funded by survey-specific projects.

Highlights: During FY 2010, staff conducted an expert review of the questionnaire for the Survey of International Air Travelers, sponsored by the Department of Commerce. We also conducted an expert review of the Household Survey of the Unbanked and Underbanked, sponsored by the Federal Deposit Insurance Corporation. We also provided guidance on questions asked of respondents on cell phones in the Survey of Fishing, Hunting and Wildlife-Associated Recreation and TPOPS.

Staff: Terry DeMaio (x34894)

1.14 RE-ENGINEERED SURVEY OF INCOME AND PROGRAM PARTICIPATION RESEARCH
(Demographic Project 1465444)

A. Re-Engineered SIPP Methodological Research
Description: The re-engineered Survey of Income and Program Participation (SIPP) is scheduled to replace the current SIPP in 2013. This project conducts long-term methodological research to evaluate SIPP and to inform the design of re-engineered SIPP instruments and procedures, which are based on event history calendar (EHC) methodology.

Highlights: During FY 2010, the long-planned 2010 SIPP-EHC CAPI field test was implemented in January through mid-March in six Regional Offices resulting in 4,165 completed household interviews and a response rate of 82%. Research activities included finalizing training materials, observing training sessions and interviews, and implementing several evaluation methodologies. These included coordinating the recording and transcription of a small sample of interviews, collection of item-level Field Representatives (FR) notes concerning instrument issues and problems, collection of detailed cost data, coordination of FR debriefings (concerning both the instrument and training), training evaluations by trainers and respondent debriefings concerning EHC "landmark event"

procedures. Analysis of the field test data itself is in progress, and includes data analysis to compare estimates from the production survey with the re-engineered survey, a comparison of data quality for both surveys against administrative records, and analysis of the interview transcripts. Staff also began looking ahead to the 2011 field test, which will focus mainly on mover/attrition issues and the functionality of the wave 2 instrument. This test will involve a new sample of approximately 4,000 households in high poverty stratum and Field Representatives in all twelve regions will participate. Staff are collaborating in the design of a recontact experiment to track respondents.

Staff: Anna Chan (x38462), Joanne Pascale

B. Model-Based Imputation for the Demographic Directorate

Description: Staff has been asked to review and ultimately improve all of the imputation methods in demographic surveys, beginning with the Survey of Income and Program Participation (SIPP) and the Current Population Survey (CPS).

Highlights: During FY 2010, several R functions for imputing and evaluating imputation methods for the SIPP and CPS have been implemented. The functions for imputation include a flexible time-series-appropriate multiple imputation hot deck and sequential Bayesian regression. The randomized multiple imputation hot deck selects donors at random from those who match on one or more user-specified categorical variables and/or are 'close' on one or more numeric variables. If no donor is found, the function provides the option to impute the mean or median, or drop the matching constraints one by one in a pre-specified order until a donor is found or it is determined that no donor exists. The sequential Bayesian regression function allows users to choose from Cauchy, *t*, or normal priors, and fills in missing values with random draws from the posterior predictive distribution.

Two functions for evaluating the imputation have been implemented. One function produces a plot superimposing distributions of the imputed, observed, and combined data and residual diagnostic plots. The other provides summary statistics including means, variances and the fraction of variance due to missing data.

Staff: Chandra Erdman (x31235), Ben Klemens, Yves Thibaudeau

1.15 DATA INTEGRATION DIVISION (DID) SMALL AREA ESTIMATION PROJECTS (Demographic Project 7165000)

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

Description: The purpose of this research is to develop, in collaboration with the Data Integration Division (DID) (The Small Area and Poverty Estimates Branch was previously in Housing and Household Economic Statistics Division and is now in DID), methods to produce "reliable" income and poverty estimates for small geographic areas and/or small demographic domains (e.g., poor children age 5-17 for counties). The methods should also produce realistic measures of the accuracy of the estimates (standard errors). The investigation will include assessment of the value of various auxiliary data (from administrative records or surveys) in producing the desired estimates. Also included would be an evaluation of the techniques developed, along with documentation of the methodology.

Highlights: During FY 2010, due to the observed biases in both the design based point and variance estimates of the log number of poor children for small sample sizes from the previous simulation study, the staff investigated various alternative model forms for county children poverty rates. The staff found that a normal mixture model with nonlinear mean function performs well. In modeling the county poverty rate model, we require the variance estimates of the sampling errors of the children poverty rate. Because for the county children poverty rates, which is a small sample size, the variance estimates are unstable, we need to model the county variance estimates of the estimated poverty rates to improve them. The staff developed a variance model framework to model the design-based variance for the county children poverty rate from the American Community Survey (ACS) data. Results from the county variance modeling of the estimated poverty rate will be in the *Statistical Research Division Research Report Series* under the title "Using Small Area Modeling to Improve Design-Based Estimates of Variance in County Level Poverty Rates in the American Community Survey" by Jerry Maples. The staff also examined the distribution of four survey variance estimators of the estimated proportion– linearization, Fay's successive difference replication variance estimator (used in ACS), the jackknife (delete one unit), and the random group (10 groups) via simulation for random samples from different generated populations (Bernoulli, Poisson) and 2005 ACS data. Documentation of results is underway.

Staff: Elizabeth Huang (x34923), Jerry Maples, William Bell (DIR)

B. Small Area Health Insurance Estimates (SAHIE)

Description: At the request of staff from the Data Integration Division (DID), our staff will review current methodology for making small area estimates for health insurance coverage by state and poverty level. Staff will work on selected topics of SAHIE estimation methodology, in conjunction with DID.

Highlights: <u>Bayesian Benchmarking of estimates from distinct geographic models</u>
Currently, estimates for states and counties are made independently from two distinct models. Consequently, county estimates do not add up to the corresponding state estimates. Ad hoc adjustments to the county-level estimates produce final estimates that are based on two, different, overlapping models making posterior variance determination problematic.

During FY 2010, assuming the models at higher geographic levels take precedence over the overlapping components of the lower geographic model, staff developed a new approach to estimating lower level (state level, county level) parameters by using linear transformations to modify the full conditional posterior distributions currently used in the SAHIE MCMC algorithm. We showed that this method produces estimates which are nonnegative and which have a distribution that is independent of the choice of transformation. This new method is guaranteed to produce lower level estimates which aggregate to higher level estimates. Future work involves determining the feasibility of implementing this method and conducting simulation studies to compare the current methodology to the proposed methodology.

Staff periodically reviews public releases of SAHIE methodology descriptions.

Staff: Don Malec (x31718), Ryan Janicki

1.16 EDITING METHODS DEVELOPMENT (ECONOMIC PROJECT 2370054)

Investigation of Selective Editing Procedures for Foreign Trade Programs

Description: The purpose of this project is to develop selective editing strategies for the U.S. Census Bureau foreign trade statistics program. The Foreign Trade Division (FTD) processes more than 5 million transaction records every month using a parameter file called the Edit Master. In this project, we investigate the feasibility of using selective editing for identifying the most erroneous records without the use of parameters.

Highlights: In this project, we research methods to identify errors in the Census Bureau foreign trade data without the use of the edit master parameter file. During FY 2010, we investigated the application of selective editing techniques to all incoming records (not only rejects). Research and development of score functions includes an effect term to estimate the effect of changes on Value and Quantity of shipments on the final totals. We revised the implementation code and testing to fine-tune program parameters. We also added two alternative ways to assign the suspicion term in the score function by

using the log transform rather than the Hidiroglou-Berthelot transform for symmetrizing the unit price ratios. Staff have written a draft for an *SRD Research Report* "Score functions for the US Census Bureau Trade Data."

Staff: María García (x31703), Yves Thibaudeau, Rachelle Reeder (FTD)

1.17 DISCLOSURE AVOIDANCE METHODS (Economic Project 2470051)

Description: The purpose of this research is to develop disclosure avoidance methods to be used for Census Bureau publicly available economic data products. Emphasis will be placed on techniques to implement disclosure avoidance at the stage of data processing. Disclosure avoidance research will be conducted on alternative methods to cell suppression for selected economic surveys. We will also aid in the implementation of the methods.

Highlights: Most work for the Economic Directorate was concentrated on a single major project: cell suppression modernization. During FY 2010, Research and Methodology (R&M) team was formed to discuss: (1) specific problems associated with the current use of the cell suppression program and (2) ways in which those problems could be resolved by improving and modernizing the program. At the weekly R&M team meetings, both theory and implementation issues were discussed.

Staff served as a co-leader of the R&M team, and, in that capacity, developed a list of research topics that should be addressed by the R&M team in the next two years. These were further broken down into short term (within 6 months), medium term (within a year), and long term topics. The hope is for the R&M team to analyze the short term topics thoroughly enough so that analytic improvements can be incorporated into the new version of the cell suppression program that will begin development during FY 2011. The actual coding of the new version, likely with a modern language such as C++, will be done by a contractor hired specifically for this task as well as programmers in ESMPD. Some of the short-term research topics (e.g., negative values, cost functions) have been discussed at the R&M meetings, but no final resolution has been reached. There has also been much discussion of complex aspects of the program that involve the notions of capacity and backtracking. Staff has been involved in the choice of optimization software that will be used both for research and for the production version of the modern cell suppression program. Staff stressed to the group the importance of using an optimization modeling language in the new program. Staff demonstrated to the group how this could be done by showing how such a language, AMPL, could be used

within a Fortran program. Such modeling languages are a good way to write code which accesses optimization solvers in a way that is clear and easily modifiable.

Staff: Laura Zayatz (x34955), Asoka Ramanayake, Jason Lucero, Paul Massell, Bimal Sinha, Tapan Nayak

1.18 TIME SERIES RESEARCH
(Economic Project 2370052)

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

Highlights: During FY 2010, seasonal adjustment and X-12-ARIMA support was provided to the Bundesbank, the Reserve Bank of Australia, Bank of Japan, Bank of England, Credo Consulting, Haver Analytics, Ford Corporation, Railway Association of Canada, SAS, Franklin Templeton Investimentos (Brasil) Ltda., Baker & McKenzie Consulting LLC, Fortress Investment Group, Ford Corporation, Apple Corporation, Morgan Stanley, SAS Institute, American Forest & Paper Association, The Conference Board, International Labour Organization, Intuit Inc., PC-Software Solutions (Germany), Mathworks, SAS, the Wall Street Journal, Bureau of Labor Statistics, U. S. Department of Agriculture, FHFA, Bureau of Economic Analysis, South Carolina Department of Commerce, Australian Bureau of Statistics, Office of National Statistics (UK), Statistics Canada, Central Statistical Office of Ireland, Statistics Austria, Korean National Statistical Office, Statistics Netherlands, Statistics South Africa, Statistical Research Institute (Korea), Swiss Statistical Agency, Ministry of Economy, Trade and Industry (Japan), Statistics Sweden, Government of Turkey, Office of National Statistics (UK), Australian Bureau of Statistics, Columbian Statistics Office, Ghana Statistical Service, INSEE, Norwegian Labour and Welfare Service, IMBG (Brazilian Government), Bank of Korea, Bank of Spain, Bureau of Economic Analysis, University of El Manar (Tunisia), Nankai University, Universidad Nacional de Tucumán (Argentina), Augusta State University, and the University of Valencia.

Staff prepared comments on a report by the Office of Statistical Methodology and Research for Economic Programs on results from an attempt to implement the Causey-Trager benchmarking method in SAS. Staff also developed a version of X-12-ARIMA on a Blade server for the use of staff from the Housing and Household Economic Statistics Division in their production work, and investigated how the observations at the beginning and end of a time series are replaced in the initial

iterations of the X-11 algorithm at the request of staff from the Office of the Director.

Staff: Brian Monsell (x31721)

B. Seasonal Adjustment Software Development and Evaluation
Description: 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) developing a Windows programming interface for the X-12/X-13 seasonal adjustment software in collaboration with analysts from the Bank of Belgium; (2) finishing a version of the X-13ARIMA-SEATS program with accessible output and improved performance so that, when appropriate, SEATS adjustments can be produced by the Economic Directorate; and (3) incorporating further improvements to the X-12-ARIMA/X-13A-S user interface, output and documentation. In coordination and collaboration with the Time Series Methods Staff of the Office of Statistical Methods and Research for Economic Programs (OSMREP), the staff will provide internal and/or external training in the use of X-12-ARIMA and the associated programs, such as X-12-Graph, when appropriate.

Highlights: During FY 2010, staff released an updated version of X-12-ARIMA, Build 188 of Version 0.3, to the Economic Directorate for their testing and then to the general public. Staff compared adjustments from this version of the software to the last released version of X-12-ARIMA (Build 177) and found in most cases no differences in the adjustments, and small differences between compilers on Linux-based machines. Staff also provided support for analysts and programmers in the Economic Directorate in their testing of the new release, and compared the performance of this build of X-12-ARIMA compiled with different Fortran compilers. Before release of the software, staff repaired a defect in X-12-ARIMA to ensure that the A18 table (original series adjusted for calendar effects) had leap-year preadjustment factors removed from it, and tested the program to ensure the table was produced correctly for a number of scenarios. Staff updated source files used in Eurostat's Demetra+ prototype to the most recent version of X-12-ARIMA, and updated the source code for X-12-ARIMA-FAME to latest version of X-12 for development by Statistics Norway.

Later in the year, staff repaired defects found in Build 188 of X-12-ARIMA encountered by the Manufacturing and Construction Division (MCD) in its testing of the software. A revised version of the software was provided to MCD for its use, and a version was provided to the TSAR software team to make available to MCD when it uses this version of X-12-ARIMA in production. Staff also fixed a defect with the accessibility codes for the

output generated by the *pickmdl* spec of Version 0.3 of X-12-ARIMA. Staff also changed how transformations affect the trading day regressors in the X-13A-S prototype, and the default for autocorrelation diagnostics for annual data.

Staff updated the SEATS code in the most recent X-13ARIMA-SEATS (X-13A-S) prototype to match the most recent version of SEATS and tested this code to check for errors and inconsistencies. Staff also continued to develop HTML output for the X-13A-S prototype and developed subroutines in Fortran and C++ for the new signal diagnostic routines. Staff added an option to allow users to specify whether a series is a flow or stock series into the series and composite specs of X-13ARIMA-SEATS. Staff prepared and taught a seasonal adjustment class using X-13ARIMA-SEATS at Eurostat. Staff assisted in determining how length-of-month adjustment factors could be generated for the regCMPNT program.

Staff: Brian Monsell (x31721), Christopher Blakely, Natalya Titova, David Findley (Consultant)

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

Description: The main goal of this research is to discover new ways in which time series models can be used to improve seasonal and calendar effect adjustments. An important secondary goal is the development or improvement of modeling and adjustment diagnostics. This fiscal year's projects include: (1) continuing research on seasonal adjustment diagnostics; (2) studying further the effects of model based seasonal adjustment filters; (3) examining goodness of fit diagnostics for time series modeling and signal extraction; (4) determining if information from the direct seasonally adjusted series of a composite seasonal adjustment can be used to modify the components of an indirect seasonal adjustment; (5) studying the modeling of seasonality using Bayesian methods, and determining if using such a method is feasible for short time series; (6) studying the modeling of stock holiday and trading day on Census Bureau time series; (7) examining approaches for modeling time series with heteroskedastic errors.

Highlights: During FY 2010, staff: (a) continued empirical studies of model-based seasonal adjustment diagnostics, (b) developed algorithms and code for estimation of seasonal long memory models, (c) studied Markov Chain Monte Carlo techniques for Bayesian estimation of seasonal time series models, (d) studied properties of parameter estimates for mis-specified models when sampling error or cyclical effects are present, (e) studied the performance of benchmarking methods with graphics developed by Statistics Canada and compared the performance of two benchmarking procedures on Census Bureau series, (f) studied performance of new distribution theory for Ljung-Box statistics and compared the size and power properties of an alternate Box-Pierce statistic with those of the Box-

Pierce and Ljung-Box statistics, (g) studied how seasonal adjustment methods facilitate short-term forecasting of cyclical and trend dynamics, and (h) implemented a method to fit time series models by minimizing multi-step ahead forecasting error.

Ongoing research includes: (a) examining the effects of model mis-specification on seasonal adjustment, (b) exploring seasonal adjustment for long memory models, and (c) Bayesian approaches to outlier modeling and signal extraction.

Staff: Tucker McElroy (x33227), Christopher Blakely, Brian Monsell, Ekaterina Sotiris, Natalya Titova, William Bell (DIR), David Findley (Consultant)

D. Supporting Documentation and Software for X-12-ARIMA and X-13A-S

Description: The purpose of this project is to develop supplementary documentation and supplementary programs for X-12-ARIMA and X-13A-S that enable both inexperienced seasonal adjustors and experts to use the program as effectively as their backgrounds permit. This fiscal year's goals include improving the documentation of X-12-ARIMA, improving the documentation of X-12-ARIMA, rendering the output from X-13A-S accessible, and exploring the use of component and Java software developed at the National Bank of Belgium.

Highlights: During FY 2010, staff released an updated version of the Genhol utility to the public with expanded documentation for the utility. Later, staff included the capability in the software to generate change of regime holiday regressors. Staff also updated the Seasonal Adjustment papers site and the X-12-ARIMA site, uploading new versions of the Win X-12 and X-12-Graph programs developed by the Time Series Methods Staff (OSMREP) as well as X-12-ARIMA and Genhol, and redesigned the sites to match the Census Bureau's new Look and Feel Guidelines. New content was added by staff related to new users of X-12-ARIMA, and provided listings of the Seasonal Adjustment Papers website by Author and Topics. A similar redesign was performed by staff for the division's *Research Report Series*.

Staff developed a utility for using X-13A-S in simulation studies by creating a C++ interface for the Fortran routines used in X-13A-S. A user interface and graphics in Java are also being developed.

Staff: Brian Monsell (x31721), Christopher Blakely

1.19 SURVEY OF RESEARCH AND DEVELOPMENT IN INDUSTRY, IMPUTATION AND SAMPLING RESEARCH AND SOFTWARE DESIGN
(Economic Project TBA)

Description: This project undertakes research on the imputation of unreported mandatory items in the Survey of Research and Development in Industry, sponsored by the National Science Foundation. It also examines what estimators are more appropriate under alternative sampling plans; in particular, it evaluates using calibration estimators to compensate for missing data. The possibility of extending calibration to new sampling plans, such as balanced sampling, is investigated. Both traditional linear regression techniques and nonparametric regression techniques are examined.

Highlights: In the context of the SIRD, staff implemented new imputation methods that are designed for nonmonotone nonresponse, and thus do not require any observed data to be discarded.

Staff: Yves Thibaudeau (x31706), Martin Klein, Jun Shao

1.20 GOVERNMENTS DIVISION PROJECT ON DECISION-BASED ESTIMATION (Economic Project TBA)

Description: This project involves joint work with Governments Division on variance estimation for total government employment and payrolls in the Survey of Public Employment and Payroll, within a new method of stratumwise GREG estimation based on PPS sampling. In this method, substrata based on small and large units are either treated as separate strata or combined according to the results of hypothesis tests on equality of regression slopes.

Highlights: During FY 2010, staff coded the decision-based estimation method along with Monte Carlo and Bootstrap simulations in R. Many runs have been made and preliminary results collected.

An ongoing research project employed Monte Carlo simulation to investigate on Bootstrap variance estimation for these decision-based estimators and their comparison with existing variance formulas that ignore the randomness inherent in the decision-based stratum definition. Preliminary indications are that both the bootstrap and naive variance estimators can be far away from the true variances due to small substratum samples and highly influential points in the GREG regression fits.

Staff: Eric Slud (x34991), Yang Cheng (GOVS), Carma Hogue (GOVS)

1.21 REMOTE ACCESS - MICRODATA ANALYSIS SYSTEM (Strategic Planning and Innovation Project 0359999)

Description: Researchers and sophisticated data users' demand for Census Bureau microdata, both for general research and programmatic needs, continues to grow. Microdata allows virtually any type of analysis, and it is the desired form of data that allows modeling. Internal Census Bureau microdata files contain levels of detail and variables, which are not available in public use files. Methods are applied to reduce detail, both by suppressing and coarsening variables in public use files, in order to protect the identity of respondents and to ensure confidentiality of responses under Title 13 of the U.S. Code. As data on individuals accumulate, and identifiable public and commercial data becomes more and more accessible, the ability to publish quality microdata while maintaining a sufficient level of ambiguity is becoming an issue.

Highlights: During FY 2010, staff members continued to work with members of the Data Integration Division (DID) on the development of a new Advanced Query System (AQS) / Microdata Analysis System (MAS). Staff members performed some further testing on a MAS confidentiality rule known as the *Drop q_v Rule*. The *Drop q_v Rule* is a universe subsampling routine in the MAS, which is meant to protect against differencing attack disclosures. A differencing attack disclosure occurs on the MAS when a data intruder creates two similar universe data sets on the MAS. Staff members analyzed the *Drop q_v Rule* and found an approximate probability distribution that models the probability of obtaining two independently subsampled m-way tables. Using this approximate probability distribution, staff members found a function that models the total approximate probability of a subtraction problem. Using the concepts of Majorization and Schur-convexity, staff members were able to recommend the values of q to be used for a given universe.

Staff members continued to work on the synthetic residuals program for the MAS. Staff members optimized the current R program that generates synthetic residuals and synthetic fitted values to run faster. Results show that the new program runs many times faster than the previous program. Staff members explored other possible ways to generate synthetic residual bivariate plots by using multivariate statistical methods. These new methods show a great improvement to those previously explored. Documentation on the new methods has been drafted.

Staff members are currently looking into possible ways to test the confidentiality rules within the MAS.

Staff members continue to work on a cutpoint program for the MAS. The cutpoint program is used to reduce disclosure risk by bucketing observations into bins. We require that each bin contain at least 80 observations. While programs like this have been developed before, there has not been a systematic study of the strengths and

weaknesses of each binning strategy. To date, we have programmed 3 different strategies. During the summer we generated a set of cutpoints using a frequency based strategy for generating the bins. This is currently what is used in the MAS. We are replacing that with a strategy based on recursive partitioning. The code is complete and the new cutpoints will be added to the MAS in January. At this stage, it is unclear that this strategy is optimal. Therefore, we will continue to program other strategies and evaluate their strengths and weaknesses on variables with different frequency distributions.

Staff: Laura Zayatz, Asoka Ramanayake, Jason Lucero, Paul Massell, Lisa Singh

1.22 PROGRAM DIVISION OVERHEAD
(Census Bureau Project 0381000)

A. Division Leadership and Support
This staff provides leadership and support for the overall collaborative consulting, research, and operation of the division.

Staff: Tommy Wright (x31702), Robert Creecy, Matt Gore (HRD), Michael Hawkins, Gloria Prout, Stephanie Sheffield, Sarah Wilson, Kelly Taylor

B. Research Computing
Description: This ongoing project is devoted to ensuring that Census Bureau researchers have the computers and software tools they need to develop new statistical methods and analyze Census Bureau data.

Highlights: During FY 2010, division staff, in cooperation with the IT Directorate, planned and provisioned a cluster of blade servers (research2) that will replace the current SGI Altix 3700 (research1). After detailed planning in Q1, the Computer Services Division (CSVD) built and delivered four identical IBM LS42 blade servers in Q2, each with four quad-core processors and 128 GB of RAM. Computing throughput did not meet our expectations, and three more nodes were added, for a total of seven. As of the fourth quarter, a few problems remain, in particular disappointing aggregate I/O performance for writes, and various file locking issues with the shared filesystems (GFS2). We are working with CSVD and RedHat engineering to resolve or minimize these problems before moving the bulk users to the new environment. Because of imminent Decennial related projects, research1 will likely remain in production for at least a portion of FY 2011.

LTSO completed the life cycle replacements of staff desktops. All staff desktop computers now meet or exceed the current desktop standard.

The Certification and Accreditation of CEN14 was completed and interim authority to operate has been granted for the system. All but one of the findings requiring action by our staff have been addressed, and it is expected that all of them will be addressed on schedule (October 23, 2010).

In their role as members of the Data Management Committee (DMC), division staff proposed that the Census Bureau work toward a shared file system/computing space that would provide greater access to data via the network and support the goals of "utility computing." These ideas as well as others developed by the committee will be developed further in an upcoming pilot project in FY2011.

Face Reader, a software package that automatically analyzes facial expressions, was acquired for use by the Usability and Human Factors Research group. The ability to automatically detect facial cues related to a subject's emotional state is a valuable tool for testing survey questions and potentially other applications as well.

Staff: Chad Russell (x33215)

2. RESEARCH

2.1 – 2.2 GENERAL RESEARCH AND SUPPORT TOPICS
(Census Bureau Projects 0351000, 1871000)

Statistical Methodology

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

Highlights: During FY 2010, the Disclosure Review Board worked with staff at the Social Security Administration (SSA) to develop an agreement on the review process for data products developed at SSA that include Title 13 data.

Staff members presented an issue to the Data Stewardship Executive Policy Committee on whether to treat some variables or subpopulations as more sensitive than others when undergoing disclosure review. The DSEP ruled that all variables and subpopulations should be treated in the same fashion.

Staff members wrote numerous documents on the correction of public use microdata files from Census 2000, ACS03-06, and CPS04-09. Meetings were held with researchers and with staff from other agencies with Disclosure Review processes in an attempt to improve the Census Bureau's disclosure avoidance procedures and processes.

Staff developed, evaluated, and documented data swapping software that may be used to perturb ages for a subset of respondents in the Current Population Survey.

Staff documented research on estimating quantiles from data to which multiplicative noise was added.

Staff: Laura Zayatz, Asoka Ramanayake, Jason Lucero, Paul Massell, Bimal Sinha, Lisa Singh, Tapan Nayak

B. Disclosure Avoidance for Microdata
Description: Our staff investigates methods of microdata masking that preserves analytic properties of public-use microdata and avoid disclosure.

Highlights: During FY 2010, staff reviewed literature, refereed four papers related to microdata confidentiality, and reviewed a book proposal for J. Wiley. Staff reviewed, commented on, and corresponded with the authors of a critique of one of the algorithms of differential privacy. Staff wrote a new version of very fast software for generating synthetic data with valid analytic properties and reduced/eliminated re-identification risk that was demonstrated as part of a short course at the University of London. What is surprising is how easy it is to re-identify in synthetic data that has been generated from very high quality edit/imputation models. The use of convex constraints in the general EMH algorithm (Winkler 1993, 1990 *Ann. Prob.)* allows significant reduction or elimination of re-identification risk while maintaining very high quality in synthetic data.

Staff: William Winkler (x34729), William Yancey

C. Seasonal Adjustment (See Economic Project 2370052)

D. Household Survey Design and Estimation
Description: The household surveys of the Census Bureau cover a wide range of topics but use similar statistical methods to calculate estimation weights. It is desirable to carry out a continuing program of research to improve the accuracy and efficiency of the estimates of characteristics of persons and households. Among the methods of interest are sample designs, adjustments for nonresponse, proper use of population estimates as weighting controls, and the effects of imputation on variances.

Highlights: See projects "ACS Small Area Estimation for Selected Characteristics" and "ACS Small Area Estimation for Group Quarters (GQ)."

Staff: Lynn Weidman (x34902)

E. Sampling and Estimation Methodology: Economic Surveys
Description: The Economic Directorate of the Census Bureau encounters a number of issues in sampling and estimation in which changes might increase the accuracy or efficiency of the survey estimates. These include estimates of low-valued exports not currently reported, alternative estimation for the Quarterly Financial Report, and procedures to address nonresponse and reduce respondent burden in the surveys. Further, general simulation software might be created and structured to eliminate various individual research efforts. An observation is considered influential if the estimate of total monthly revenue is dominated by its weighted contribution. The goal of the research is to find methodology that uses the observation but in a manner that assures its contribution does not dominate the estimated total or the estimates of period-to-period change.

Highlights: Staff collaborated with staff from the Economic Directorate in the implementation of a newly designed simulation to investigate the properties of two methods, M-estimation and Clarke Winsorization, which treat influential values. The team is investigating treatments for influential values in the Monthly Retail

Trade Survey. The team completed initial work that incorporated comments from the designer of the M-estimation method, Jean-Francois Beaumont of Statistics Canada, on the results of the first scenario. Staff hosted Jean-Francois Beaumont in a visit to the Census Bureau to discuss the design of the simulation and our initial results as part of SRD's *SUMMER AT CENSUS*. The discussions led to a refinement in the research plan. Initial work on the revised plan has begun.

Staff: Mary Mulry (x31759)

F. Research and Development Contracts
Description: 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. 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 problems through research. Many of the prime contractors are teamed with one or more organizations and/or have arrangement with outside experts/consultants to broaden their ability to meet all of the potential needs of the Census Bureau. These 5-year contracts allow Census Bureau divisions and offices to obtain outside advisory and assistance services to support their research and development efforts quickly and easily.

R&D 2007 Contracts
Twenty-five contracts were awarded during Fiscal Year 2002 in six technical areas: 1) assessment, planning, and analysis; 2) data analysis and dissemination; 3) statistical analysis, 4) methodological research, 5) sub-population research, and 6) survey engineering. The contracts ended September 30, 2009; however, the task orders awarded prior to the end date will be allowed to continue until completion.

During FY 2010, twenty-seven modifications were awarded and five were completed. To date, there have been ninety-six (96) task orders awarded under the R&D 2007 contracts, with a monetary value of over $118 million. Seventy-nine task orders have been completed and one task order terminated, leaving sixteen active tasks.

R&D 2014 Contracts
Thirty-seven contracts were awarded during Fiscal Year 2009 to thirty-one firms in five technical areas: 1) assessment, planning, and analysis; 2) data analysis and dissemination; 3) statistical analysis and evaluation, 4) methodological research, and 5) survey engineering.

During FY 2010, thirty (30) new task orders were awarded, two were modified and three were completed. To date, there have been thirty task orders awarded under the R&D2014 contracts with a monetary value of $16 million. Three task orders have been completed, leaving twenty-seven active task orders.

Staff: Ann Dimler (x34996), Christina Cooper

G. Small Area Estimation
Description: Methods will be investigated to provide estimates for geographic areas or subpopulations when sample sizes from these domains are inadequate.

Highlights:
Aggregate Level Small Area Estimation Modeling
Staff are exploring new ways to estimate the probability of correct enumeration in the census using small area techniques. This work involves incorporating random effects into the model and accounting for clustering in the sample. Staff showed how to modify the design weights to construct a new quasi-likelihood with certain desirable properties. However, further work is needed to show how to overcome potential undesirable results, such as many modified weights being set to zero or being made too large. We have also considered semiparametric methods. As there is no available likelihood in this problem it is natural to construct (estimated) estimating functions. The idea is to use splines, or other nonparametric smoothers, to estimate the mean and the covariance function for clustered data. This approach involves far fewer model assumptions and should produce estimates which are more robust to model misspecification.

Unit Level Small Area Estimation Modeling
Staff completed the evaluation of housing unit and block-level random effects on coverage using randomization tests with the 2006 Travis county test data. Results were presented at the 2010 JSM.

Staff: Don Malec (x31718), Ryan Janicki

Statistical Computing Methodology

A. Record Linkage and Analytic Uses of Administrative Lists
Description: Under this project, our staff will provide advice, develop computer matching systems, and develop and perform analytic methods for adjusting statistical analyses for computer matching error.

Highlights: During FY 2010, staff reviewed a considerable amount of literature on parallel and threaded programming methods. Staff also reviewed parallel programming associated with graphics processors (GPUs), particularly those related to an invited JSM

session on fast computation for MCMC. Staff reviewed literature on record linkage in the health sciences. Staff reviewed seven papers on record linkage from Statistics Canada. Staff sent papers, software, and advice on record linkage to Statistics Canada.

Staff wrote draft papers on record linkage for the Joint Statistical Meetings. Staff continued investigating generalizations of the Birthday and Collision Problems. Staff provided the latest list of references related to record linkage to a professor at Harvard who maintains a web site of methods/papers useful in the survey literature.

Staff refereed two papers on record linkage. Staff reviewed and commented on one paper on privacy-preserving record linkage. Staff made comments for a professor at the University of Western Australia who was performing a comparison of various record linkage software packages. Staff sent her additional software for standardization. Staff sent a researcher working on the openempi record linkage software (www.openempi.org) software for one of the NIH institutes detailed comments related to an implementation of the EM algorithm for parameter estimation. Staff provided detailed comments to the authors of two papers on improved string comparator metrics, the best of which appears to improve over the basic Jaro-Winkler string comparator that is widely used by computer scientists. Staff provided detailed advice on two proposed record linkage projects to a professor at Duke University. Staff wrote an overview document of key record linkage issues affecting the use of administration lists for the Committee on Administrative Records and Record Linkage.

Staff taught the short course "Record Linkage" at the Institute of Education at the University of London. During the course, staff demonstrated the latest version of *BigMatch* software (Yancey and Winkler 2004-2009), two new results using the EM algorithm for unsupervised learning, and some new applications of record linkage.

BigMatch is used for two very large Decennial Census production matching projects and may be used for evaluation projects. *BigMatch* continues as the fastest, high quality record linkage software (sustained rates of 900,000+ pairs per second with the fastest version on a Core I7 chip). *BigMatch* runs on all known computer architectures.

Staff: William Winkler (x34729), William Yancey, Ned Porter

B.1 Editing
Description: This project covers development of methods for statistical data editing. Good methods allow us to produce efficient and accurate estimates and higher quality microdata for analyses.

Highlights: During FY 2010, we presented research at the work session on Statistical Data Editing. This research covered new ideas on creating edit rules for a given set of data, applying the edit rules efficiently to maximize quality of the resultant database, providing strategies for monitoring the edits, and providing partial measures for assessing the quality of the edit set (M. Garcia and W. Winkler, "Determining a set of edits and quality of a database"). Staff wrote a new version of the *genbnds* program for generating a complete set of ratio edits for a given set of explicit ratio edits that could be embedded within the general package for edit/imputation. The FORTRAN version of the code implements a variation of the shortest path algorithm (with no cost transformation) similar to the one we previously implemented in the SAS/IML and SAS/OR code.

Staff: María García (x31703)

B.2 Editing and Imputation
Description: Under this project, our staff provides advice, develops computer edit/imputation systems in support of demographic and economic projects, implements prototype production systems, and investigates edit/imputation methods.

Highlights: During FY 2010, staff integrated several new methodologies, in particular new model-based multiple imputation methods and hot-deck multiple imputation methods to PEP. Staff submitted results to professional publications.

Staff: Yves Thibaudeau (x31706), Chandra Erdman, María García, Martin Klein, Ben Klemens, Rolando Rodriguez

C. Developed Software Support – General Variance Estimation Development and Support
Description: This project will develop new methods and interfaces for general variance estimation software including VPLX, WesVar, and SUDAAN. Our staff will provide training for variance estimation software applications, and will provide support for complex applications such as the Survey of Income and Program Participation and the Survey of Construction.

Highlights: During FY 2010, staff provided ongoing VPLX support to the Economic Directorate. Staff were contacted by the Manufacturing and Construction Division on the topic of calculating variance estimates for medians. Staff were contacted by members of the Economic Statistical Methods and Programming Division on the topic of creating a new VPLX executable for a new server in their division. Variance estimation software support reverted to the respective directorates as of July 1, 2010. General variance estimation development and support, including research and methodology, remained in our division.

Staff: Aref Dajani (x31797), Ned Porter

D. Missing Data and Imputation: Multiple Imputation Feasibility Study

Description: Methods for imputing missing data are closely related to methods used for synthesizing sensitive items for disclosure limitation. One method currently applied to both issues is multiple imputation. Although the two issues may be addressed separately, techniques have been developed that allow data users to analyze data in which both missing data imputation and disclosure limitation synthesis have been accomplished via multiple imputation techniques (e.g., synthetic data). This project ascertains the effectiveness of applying multiple imputation to both missing data and disclosure limitation in the American Community Survey (ACS) group quarters data. Statistical models are used to generate several synthetic data sets for use within the multiple-imputation framework.

Highlights: During FY 2010, we generated a new package in R, tentatively called PEP, to generate edit-compliant synthetic data for American Community Survey. PEP integrates previously-used modeling strategies from past ACS years with additional software to improve workflow and insure edit consistency (notably we adapted Bill Winkler's DISCRETE program into C for this purpose). Having tested the software on both SRD and ACS production systems, we successfully used PEP for statistical disclosure control (via synthetic data methods) for the 2009 ACS group quarters sample. Future research will focus on integrating multiple imputation methods into PEP and testing these methods on the ACS.

Staff: Rolando Rodríguez (x31816), Ben Klemens, Yves Thibaudeau

E. Modeling, Analysis, and Quality of Data

Description: Our staff investigates methods of the quality of microdata primarily via modeling methods and new software techniques that accurately describe one or two of the analytic properties of the microdata.

Highlights: During FY 2010, staff reviewed a Ph.D. dissertation and several journal articles on imputation. One staff person provided software, documentation, and advice related to generalized edit/imputation software to other staff within our division. Staff reviewed literature on Statistical Matching. Staff continued work on generalized nonresponse-variance estimation routines for discrete data.

Staff put together a short document related to the need for greatly increased algorithm speed. In many production (or even analysis) situations, a 10- or 100-fold increase in speed may decide whether certain advanced methods can be used. The document included four examples of generalized systems successfully developed at the Census Bureau. Staff provided advice and comments related to three proposed 2020 Decennial-Census research projects. Staff provided detailed comments and links to research

papers on edit/imputation, microdata confidentiality, and record linkage to staff at the Australian Bureau of Statistics.

Staff taught the short course "Cleaning Administrative Data: Improving Quality with Edit and Imputation" at the Institute of Education at the University of London. Staff provided new versions of the modeling/edit/imputation software and the general synthetic-data-generation software.

Staff: William Winkler (x34729), Rob Creecy, William Yancey, María García

F. Top Coding Investigation

Description: During the second quarter, staff was asked to carry out the statistical analysis and evaluation of an income sample following a lognormal distribution, when the sample includes top coded values; i.e., incomes above a specified level are not reported in the sample for the purpose of disclosure control. Thus, top coding results in a singly right censored sample. Even when the lognormality assumption holds, data analysis can be complicated under top coding. For analyzing income data in this scenario, a methodology that has been recommended in the literature consists of the imputation approach. The goal of the project is to develop data analysis techniques based on other novel methodologies available in the statistics literature, avoiding the use of imputations.

Highlights: During FY 2010, we investigated the application of the novel concept of a generalized confidence interval for analyzing top coded income data, under the assumption of lognormality. We derived confidence intervals for a variety of income parameters that are of practical interest: the mean, median, percentiles, Gini coefficient, fraction of incomes that are top coded in the population etc. We are currently investigating the performance of the confidence intervals so obtained. We anticipate that the generalized confidence interval approach will result in a unified methodology for analyzing top coded data.

Staff: Thomas Mathew (x35337), Aref Dajani

Social & Behavioral Sciences Survey Methodology

A. Usability Research and Testing

A.1. Web Applications Accessibility
Description: This project focuses on the accessibility of Internet and Intranet applications by blind and low vision users in accordance with the *Section 508* regulations.

Census 2010 Web Site (Systems Support Division)

Description: This Web site provides information via text, audio, and video to people interested in the 2010 Census.

Highlights: During FY 2010, the following accessibility issues of the site were identified and corrected:

The main feature of the site, the marquee, was found to be inaccessible because there was no keyboard access to the video histories. The remedy was to create a work-around to access the video histories.

An interactive 2010 Census form was made accessible by providing a plain text version.

On the main page, the "Navigate to" combo box was accessible, but automatically goes to the first selection when using the down arrow. The remedy is to add the instruction "Use the alt and down arrow key to open the list" as title text to select an option.

Although accessible, a combo box was too narrow to display all text, so a recommendation was given to widen the field so text could be read by all users.

The links to access video stories in English and Spanish were not accessible. Staff recommended labeling and placing these links in the tab order so they would be made accessible to screen-reader users.

The media player provided to play video stories was not accessible. Users could tab to the image of the person, but could not play the video. The visual focus did not follow the visual order of moving from upper left to lower right and no labels for player controls were spoken by the JAWS screen-reader.

Proposed slider windows for a horizontal menu view were determined to be inaccessible and drop-down windows were recommended as an accessible alternative.

Staff: Larry Malakhoff (x33688), Lisa Wolfisch (SSD), Carollyn Hammersmith (SSD)

2010 Title 13 Awareness Training E-Learning Application (Systems Support Division)
Description: This application permits Census Employees to satisfy the annual requirement for Title 13.

Highlights: During FY 2010, testing revealed what appeared to be text and graphical links were all button labels. Issues with button labels were reported because users expect items appearing as links to function as links. All text and graphical buttons were unlabeled, making pop-up windows inaccessible. In the course of accessibility testing, two main usability issues were detected. 1) There was no visual indication a link had been visited (clicked) by turning from blue to purple. 2) Response option selection was overly complex. Conventional radio button groups as response options were replaced by a layout where users were directed to

select a letter for an item instead of selecting the item directly. A report detailing these findings was written and delivered to the sponsor for their action.

Staff: Larry Malakhoff (x33688), Lisa Lawler (SSD)

CPSI-PAL High-Level Expert Reviews (Systems Support Division)
Description: This project is a follow-up of the review of the high-level concept document of the CPSI-PAL web site performed in 2009. This Web site will allow users to research roles and responsibilities for usage in process improvement.

Highlights: During FY 2010, staff performed and delivered two expert reviews in preparation for the planned release in January 2010. Staff found in the first review Census jargon was used on the welcome screen which was potentially confusing to users, acronyms were not clearly defined, and the alignment of text varied between pages. The second review confirmed recommended changes were made. Staff found the usefulness of the search widget to be problematic. First, search results were not meaningful and the position of the search box made users believe pressing the "Go" button next to the search box would also work for the "Process," "Role," or "Topic" searches.

Staff: Larry Malakhoff (x33688), Temika Holland, Jennifer Romano, Jennifer Chen

NotifyMe (Economic Planning and Coordination Division)
Description: NotifyMe allows persons to select manufacturing reports and be notified when they become available.

Highlights: During FY 2010, the following aspects of the site were evaluated and corrected:

1) Neither of the extended list boxes to select states or sectors is labeled. Further, text describing step 1, step 2, and the instructions to use keys to make selections precedes the states extended list box. The step 2 text should not precede the states list box because it will cause confusion. As a result of the current layout, there is no text preceding the sectors extended list box. A report was provided to the sponsor detailing the correct navigation sequence.
2) JAWS users cannot use control-left click to select single options. There is a conflict between the way JAWS uses the control key and how this key is defined for everyone else. We recommend these JAWS instructions: "Use shift-down-arrow to select multiple options next to each other. Use shift-F8 and the arrow keys and press the space bar to select different items." The JAWS instructions should not be visible (e.g., programmed either as alt text in a transparent graphical link or as the title text for the extended list box).

Staff: Larry Malakhoff (x33688)

Support for X-12 ARIMA Documentation and Software
(Statistical Research Division)
Description: Accessibility review support for X-12 ARIMA products.

Highlights: During FY 2010, two versions of a table with annual data were evaluated to determine which version provided the best accessibility. The first preserves the ordering of the data as it appeared in the original output format, and the other table presented the data in a more intuitive way. The first table was recommended because it had better accessibility due to row and column headers referencing data cells properly. For JAWS users, coding of the second table had the effect of nesting column headers and row stubs together, making them difficult to understand.

Staff: Lawrence Malakhoff (x33688), Brian Monsell

Data Tables (Administrative Customer Services Division and Systems Support Division)
Description: Staff reviews various data tables for accessibility and provides recommendations if the table is not coded properly.

Highlights: During FY 2010, staff from ACSD requested a review of a PDF report that included data tables. Tables with one level of column headers did conform to *Section 508*. However, other tables with nested column and row headers are not coded correctly. Screen-reader users can only hear headers closest to the data vocalized instead of the two or three levels present on the screen. The recommendation for the non-compliant tables is to reconfigure them by moving the column sub-headers to the left side of the table, making them row headers.

An SSD HTML data table was evaluated for compliance and it was determined the second level column headers were not coded correctly, i.e., they were not vocalized by the screen-reader when accessed by the table reading key commands. Some row headers included a footnote after a four-digit year, but the footnote is just announced as a number, which impacts data quality (e.g., if the year is 1990 and the footnote is a 2, the string sounds like "1992"). A reference to a recommendation for accessible footnotes was provided to the sponsor.

Staff: Lawrence Malakhoff (x33688), Laura Yax (SSD)

Censtats & USA Counties (Administrative Customer Services Division)
Description: These Web sites permit users to get national, state, and county statistics.

Highlights: During FY 2010, staff provided feedback to a programmer about lack of labels for combo boxes used to select parameters to generate tables. Tables generated by these applications used numeric codes as a description of the remaining items in the row instead of an actual text description. Further, all tables had buttons in them to show different views of the data that were all labeled with the same text. This project is on hold until the ACSD programmer makes corrections.

Staff: Lawrence Malakhoff (x33688), Tina Egan (ACSD)

Central Indicator Data Repository (CIDR) (Economic Statistical Methods and Programming Division)

Highlights: During FY 2010, this application required a low to moderate effort to correct issues identified in the accessibility review. Buttons, links, and combo boxes do not show focus when accessed by the tab key throughout the application. Instructions to use a combo box directed users to select more than one option which is not possible. Staff recommended using extended list-boxes which do support multiple option selection. A link to Adobe Reader is needed for users to read PDF content. Staff reviewed CIDR screen mockups for usage of color and readability. The recommendation provided was to change the direction of the arrowheads to point into the figure to indicate where steps were to be followed.

Staff: Lawrence Malakhoff (x33688)

Census Quality Survey (CQS)
Description: The purpose of this project was to evaluate the CQS for conformance to *Section 508* standards.

Highlights: During FY 2010, portions of the CQS did not conform to Census IT Standard 15.0.2 and *Section 508*. Screen reader users would not know how to proceed because buttons were announced using variable names instead of the visible label text. Some links were not shown in underlined blue and did not change to magenta (purple) when visited. Labeling for data in three parts (Date of Birth, Name) is not programmed correctly because the third label is associated with the first data entry field (e.g., Last name label with first name field and YYYY label with MM field). These findings were compiled into a report, provided to the sponsor, and corrected.

Staff: Lawrence Malakhoff (x33688)

Regional Offices Web Site
Description: The goal of this project is to identify usability and accessibility issues of the Regional Offices Web Site through an expert review and usability testing with test participants.

Highlights: During FY 2010, staff collaborated to draft an expert review of accessibility and usability issues of the Census Bureau's Regional Offices Web site. The review revealed the map of the Los Angeles Region does not show Guam, American Samoa, and the Northern Mariana Islands. The table and map of the United States also do not list Guam, America Samoa, and Northern

Mariana Islands.   Staff also collaborated to create a test plan for the upcoming usability test. Goals include seeing how well test participants do at finding employment, becoming a Census partner, requesting a workshop, getting directions, and getting population counts. The test also includes screen reader users who will perform the same tasks as those test participants with normal vision. Staff worked with Columbia Lighthouse for the Blind to recruit JAWS users who will use the screen reader software to test usability goals.

Staff: Lawrence Malakhoff (x33688), Victor Quach

Business Help Site (Economic Planning and Coordination Division)
Description: The goal of this project is to identify usability and accessibility issues of the Business Help Site through an expert review.

Highlights: During FY 2010, staff collaborated to draft an expert review of accessibility and usability issues of the Business Help site. The following issues were revealed during the evaluation and are being remedied by the sponsor:

- The Web site does not reflect the Census Bureau's new "look and feel" template for public Web sites, which includes standardized banners across the top and bottom, and menus for top navigation, left-pane navigation, a center-panel for main content, and a right panel for additional information.

- The search box on the left panel can be used to search only for forms with numbers. This search box does not help users searching for other information. The web site lacks a general search box option, which could search by form name or key word.

- A blind person must have assistance from a sighted person to get their username and password from the paper form to report on-line, but have no idea they can accomplish this task because there is no ALT text describing this step.

- All the information sheets for all sectors, both single and multi units, are inaccessible. The information sheets were all scanned images instead of converted forms.

Staff: Lawrence Malakhoff (x33688), Sabin Lahke

A.2. Desktop Applications Accessibility
Description: This project focuses on accessibility of desktop applications by blind and low vision users in accordance with the *Section 508* regulations. Desktop applications are either downloaded or sent to the respondent on disk.

Census In the Schools (CIS) Documents (CLMSO)
Description: Teachers instructing children in all grades would use these lesson plans and exercises to learn about what information the Census Bureau collects in the 2010 Census and how it is used.

Highlights: During FY 2010, staff performed accessibility evaluations for 65 documents for the CIS program. These included two documents for grades K-8 English Language Learners; three documents for Puerto Rico grades 9-12; thirteen documents for Diversity training; four documents for English as a Second Language Adult learners; and four documents for the Census Road tour.

Staff: Larry Malakhoff (x33688)

WebTA (DOC)
Description: The WebTA application permits employees to enter their time and attendance on-line. The purpose of this project was an evaluation for compliance to *Section 508* due to a complaint before being released for general use.

Highlights: The WebTA application had several violations of *Section 508* standards. The login, logout, and help buttons did not show a dotted rectangle (focus) surrounding them. The help link was not accessible by keyboard commands. Labels were not properly associated with fields on top of the Account Selector form screen. Dollar Transactions and Mile Transactions tables on the T&A screen were not in tab order. During FY 2010, these findings and recommendations were compiled into a report and sent to the Department of Commerce for its action.

Staff: Lawrence Malakhoff (x33688)

American FactFinder (AFF) User Acceptance Testing (DACMO)
Description: As part of AFF user acceptance testing, staff was requested to contribute an accessibility evaluation of the AFF interface.

Highlights: During FY 2010, a report was provided to the sponsor for their action. The following accessibility issues of the AFF application were identified and are being remedied:

- Overlays generated by users performing searches for geographies, population groups, and industry codes were inaccessible by keyboard commands to JAWS users.

- After submission of a query on an overlay, a table is generated but cannot be accessed because keyboard focus reverts to the main screen and the overlay cannot be accessed by keyboard commands.

- The table of output on the overlay is not marked up correctly because JAWS users hear tool-tips; "click to sort descending" and "click to sort ascending," for the column headers instead of the actual displayed text.

Staff: Lawrence Malakhoff (x33688), George Huffman (DACMO)

ACSD Section 508 Implementation Team (ACSD)
Description: This goal of the team is to make all electronic content conform to Section 508 regulations.

Highlights: At the request of the Administrative Customer Services Division, tutoring was given to 24 ACSD staff members in eight sessions in August 2010 on how to create and test accessible tables for the on-line version of the 2010 Statistical Abstract and within PDF publications.

A.3. American FactFinder (AFF) Website Redesign
Description: The U. S. Census Bureau releases much of the nation's economic and demographic data on the American FactFinder (AFF) Web site. In conjunction with the massive data release anticipated at the conclusion of the 2010 Census, AFF is currently undergoing a major redesign, which is scheduled to launch sometime in 2011 or 2012. The Data Access and Dissemination Systems Office (DADSO) has asked the Census Bureau's Usability Lab to participate in the redesign effort. This effort encompasses the full spectrum of user-centered design activities, from iterative low-fidelity paper prototype testing to high-fidelity testing with a working prototype. The purpose of the testing is to identify usability issues. Recommendations made to resolve the issues are intended to improve the usability of the Web site for all users.

Highlights: During FY 2010, staff conducted usability testing on Iterations 1, 2, and 3 of the New AFF Web site. Staff wrote up the results of usability testing for Iterations 1 and 2, presented the results to the AFF development team and published the report in the *SRD Research Report Series Survey Methodology #2010-07*. Staff then worked on usability testing the revised AFF design: Iteration 3. Some of the high priority problems uncovered during the latest round of testing include issues with the geography overlay – users were confused as to how they were supposed to add geographies. In addition, users did not see the search results that load beneath the geography overlay. We recommended the development team make the geography overlay smaller either by making it narrower or pushing it lower on the page. We recommended creating an "intermediate" state for the geography overlay between open and closed or automatically minimizing the geography overlay once a filter has been selected. In addition, we recommended changing the way the other filters work, such that they *all* open to the right of the tabs on the main page. Users expect consistency within a Web site, and the geography

filter does not work the same way that the other filters do. We highlighted how this can be confusing and frustrating to users. After giving the team time to read the report, staff met with the development team and discussed high priority findings as well as documented team solutions to the identified issues.

Staff: Erica Olmsted-Hawala (x34893), Jennifer Romano

A.4. Spatial Ability Research with Iowa State University
Description: The purpose of this research is to continue the Census Bureau's investigations of the role of spatial ability in mediating the success of field personnel in performing computer-based tasks.

Highlights: Although there was no significant progress during FY 2010, staff have met over the phone with the Iowa State collaborators and plan to contribute to analysis and interpretation of data in FY 2011.

Staff: Kathleen Ashenfelter (x34922)

A.5 Usability Evaluation of the Business & Industry Web Site
Description: Our division was asked to provide usability guidance and testing support to the Economic Current Surveys Web Site Redesign Team. This effort entailed attending regular team meetings and commenting on proposed design elements; providing input to occasional planning meetings; and implementing a wide range of usability evaluation methods, from expert review to high-fidelity prototype testing.

Highlights: During FY 2010, staff met with the team and provided expert advice on screen shots of some of the emerging designs of the Business and Industry pages. Staff submitted the final report of the Business and Industry round 1 usability study to the division's *Research Report Series.* This project is complete.

Staff: Erica Olmsted-Hawala (x34893)

A.6. Baseline Usability Testing of the American Community Survey Web site
Description: Our division was asked to provide a baseline measure of the current American Community Survey (ACS) Web site. Within a year, the development team of the ACS Web site plans to make some major re-design changes to the interface of the Web site. In order to evaluate whether the re-designed effort is successful, the Usability Lab proposed a baseline usability study to measure user performance and satisfaction with the current site. Ultimately the results will be used to compare user performance and satisfaction with the same measures taken on the final release of the new American Community Survey Web site.

Highlights: During FY 2010, staff ran expert users through the baseline study, identified usability problems,

31

wrote a quick report which detailed the usability problems, and presented the results to the client. Results, as in the first study with novice users, show that the main tabs across the top of the Web site convey little about what is found under each section. For example, the word "data" is used frequently in the top navigation tabs and gives the impression that data is available within the ACS web site. However, this is misleading as the data is housed on the American FactFinder Web site. Thus we showed how for users, the current tab labels were not distinct enough to allow easy discrimination between sections. We recommended that the tabs need to be more descriptive to accurately reflect the contents of each tab. The average task success rate across all participants was 59%. The average amount of time it took for participants to complete their tasks was five minutes and seven seconds. These are below the goals set for the site. Staff wrote the results up in a final report that is published in the *SRD Report Series: Survey Methodology #2010/09*. Staff also presented the results of the baseline usability testing to the ACSO team.

Staff: Erica Olmsted-Hawala (x34893), Temika Holland

A.7. Usage of Avatars and Web 2.0 to Collect Survey Information

Description: This research effort will use Web 2.0 technology to perform interviewing, recruiting, and distance learning using digitally animated agents (avatars) and virtual worlds such as Second Life. The initial focus will be on creation of a type of digital agent known as an Embodied Conversational Agent (ECA). The ECA can process reactions of the respondent through a webcam with a microphone and can react to modify a question and use facial expressions to obtain a response.

Highlights: During FY 2010, staff performed a literature review on Web 2.0 and digital agents for presentation at the International Field Directors and Technologies Conference (IFD&TC) in Chicago, Il in May. The presentation described differences between Web 1.0 and Web 2.0, provided examples of Web 2.0 data collection widgets, and discussed potential implementations of avatars for Web surveys. An ECA can conduct an effective interview if it can have a conversation with the respondent, including all the verbal and non-verbal cues present in a face-to-face interview. Conversation includes spoken language, facial expressions, body postures, and hand gestures. Research has begun on the facial expression component of conversation using the Noldus Face Reader 3.0.5 software package. The software will classify what expressions human respondents display when asked sensitive questions. Staff has drafted a paper on how ECAs function and interact with a respondent in a Web survey.

Staff: Lawrence Malakhoff (x33688)

A.8. Usability Testing of the National Survey of College Graduates

Description: The National Survey for College Graduates (NSCG) is an online data collection Web site that collects education and job information from diverse users who have received bachelor's degrees from American schools or abroad.

Highlights: During FY 2010, the National Survey of College Graduates (NSCG) is an online data collection Web site that collects education and job information from diverse users who have received Bachelor's degrees from American schools or abroad. The NSCG was developed by the Special Surveys Branch (SSB) of the Demographic Surveys Division (DSD), and the Usability Lab tested it to identify usability issues. Recommendations made to resolve the issues are intended to improve the usability of the Web-based survey for all users.

Staff conducted two rounds of usability testing on the web-based National Survey of College Graduates. The results from the first round of testing were presented to the design and development team and were published in the *SRD report series: Survey Methodology #2010-05*. Staff then conducted another round of testing, in which we tested four versions of the instrument on 30 participants. The different versions examined the placement of the "Previous" and "Next" navigation buttons and the presentation of a long list of items (one column or two columns). Eye tracking was conducted. Results demonstrated that overall, participants were satisfied with the survey. Some issues uncovered in testing include (1) people did not read dense text on the log-in screen; (2) participants preferred the "Previous" button on the left side of the screen and the "Next" button on the right side of the screen, (3) participants preferred the long job-code list to be in two shorter columns rather than one long, scrolling column. Recommendations were made and presented to the sponsor, and the final report is in progress.

Staff: Jennifer Romano (x33577)

B. Questionnaire Pretesting

Description: This project involves coordinating the Census Bureaus 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.

Highlights: During FY 2010, requests for 32 pretesting activities requests involving 3388.5 respondent burden hours were approved by the Office of Management and Budget (OMB). A three-year extension of the generic clearance for questionnaire pretesting research until August 30, 2013 was approved. Staff also provided to OMB an annual report documenting all the pretesting activities conducted under the clearance.

Staff participated in the Demographic Surveys Division training program by giving a presentation on pretesting.

Staff: Terry DeMaio (x34894)

C. Questionnaire Design Experimental Research Survey 2006 (QDERS)

Description: QDERS 2006 is an omnibus survey designed to facilitate independent research related to questionnaire design issues and other survey methodology issues. The QDERS 2006 was conducted from the Hagerstown Telephone Center. The focus of the 2006 QDERS is a questionnaire design experiment examining different ways to determine a person's place of residency on Census day.

Highlights: During FY 2010, staff finalized a paper comparing this Random Digit Dialing study to a field pretest. In addition, staff worked on completing a final, comprehensive report for this study.

Staff : Jennifer Hunter Childs (x34927), Elizabeth Nichols, Rolando Rodríguez, Aref Dajani

D. Language: Interdisciplinary Research on Language and Sociolinguistic Issues Relevant to Survey Methodology

Description: There is a need for both qualitative and quantitative interdisciplinary research on how to best develop and successfully use non-English language collection instruments and other survey materials. Interdisciplinary research is also needed to determine the quality of the data that respondents with little or no knowledge of English provide the Census Bureau using both English and non-English language data collection instruments.

Highlights: During FY 2010, staff worked collaboratively with researchers from academia and survey research organizations on cross-cultural issues in survey interviews and translation methods. Specifically, we studied the following problems: 1) cross-cultural communication norms and survey interviews, 2) the use of interpreters in survey interviews, 3) language and cultural effects on conducting cognitive interviews in non-English languages, 4) methods to encourage survey participation from speakers of languages other than English, 5) creation of best practices for the management of non-English language cognitive testing research, and 6) methods to categorize translation issues identified in the translation-review or cognitive testing process.

In collaboration with researchers at the National Cancer Institute, we started a research project comparing English and Chinese cognitive interviews. This research effort aims at identifying methodological issues, including protocol development, probing techniques, and optimal number of interviews for cognitive testing in multiple languages. During FY 2010, staff worked to fine tune the linguistic coding scheme to code key linguistic features associated with the probing questions and responses. We continued interviewing with four groups of respondents: monolingual English interviewed in English, bilingual Chinese interviewed in English, bilingual Chinese interviewed in Chinese, monolingual Chinese interviewed in Chinese. With this design we will be able to dissect language effect versus cultural effect.

Staff members continued to be actively involved with the Federal Government Interagency Roundtable on Languages (IRL) to work on the development of the Census Bureau language assessment tool. Staff conducted a literature review on language proficiency testing and compiled a list of experts for organizing a panel of experts to help the Census Bureau work on language proficiency tests.

To understand how respondents who speak a language other than English answer survey questions, we analyzed ACS data to identify patterns of nonresponse among Chinese speakers and we found that item nonresponse rate is much higher for Chinese speakers than for English speakers. We also designed a coding scheme using discourse analysis to detect Chinese respondents' linguistic behavior in an interview setting. We discovered that as a group, Chinese respondents are more likely to provide indirect or irrelevant answers to survey questions than English respondents. This indirect and irrelevant response pattern poses a threat to data quality and validity of the constructs measured by survey questions. More research is needed on how to design survey questions and train interviewers to elicit information from respondents of diverse linguistic and cultural backgrounds.

Along this line of research, we investigated what messages were motivating to speakers of non-English languages for their census and survey participation. In addition, we started a new research project to develop a translation quality assessment tool guided by sociolinguistic theories to categorize translation issues identified through the translation-review process or the cognitive testing projects. The coding scheme will help us better understand the nature of the problems and identify solutions to address translation issues.

We presented findings from the aforementioned research at conferences and professional meetings. We were also actively writing up research results for peer-reviewed publications. In collaboration with university researchers, we completed one book on politeness communication, and are continuing to work on another edited volume on discourse and interaction.

Staff: Yuling Pan (x34950), Patricia Goerman, Anna Chan, George Higbie, Marisa Fond

E. Training for Cognitive Interviewing

Description: Our staff will train members of other divisions in the Census Bureau to carry out cognitive

interviewing and provide consultation and support for projects.

Highlights: During FY 2010, staff conducted two cognitive interview training sessions for people from our division and the Demographic Directorate.

Staff: Jennifer Hunter Childs (x34927), Yuling Pan, Patricia Goerman, Terry DeMaio

F. Research on Cognitive Testing of Non-English Language Survey Instruments
Description: The staff is currently engaged in a study designed to test and identify best practices for conducting cognitive interviews with Spanish-speaking respondents. We have tested both widely accepted and new techniques and probes (e.g., "What does the term foster child mean to you in this question?") with Spanish-speaking respondents of high and low educational levels. The research was based on a segment of the CAPI version of the American Community Survey. Future applications of this research should extend to cognitive interview techniques for use with respondents who speak additional non-English languages.

Highlights: During FY 2010, staff worked on an article for submission to academic journals.

Staff: Patricia Goerman (x31819)

G. Q-Bank: A Database of Pretested Questions
Description: Q-Bank was developed through an interagency committee, led by the National Center for Health Statistics (NCHS), of which the Census Bureau is a member. The objective of Q-Bank is to have an online interagency database of pretested survey questions and research results. The database is maintained at NCHS and is guided and used by other participating Federal statistical agencies, including the Census Bureau. Q-Bank serves many purposes. When survey questions and questionnaires are being developed, Q-Bank can be used by survey methodologists and subject matter experts to search through previously tested questions. Q-Bank provides a forum to catalog pretesting reports in a manner that is easy to search by content or subject matter. Q-Bank also will allow us to produce meta-data about our pretesting findings. Finally, Q-Bank will be an additional resource for analysts to interpret survey data. Q-Bank has just reached the production phase and is currently being populated with cognitive test reports, which is a necessary step before it becomes available to a broader audience.

Highlights: During FY 2010, staff participated in the Questionnaire Evaluation Methods workshop at NCHS, with leaders in the field of questionnaire evaluation. Staff continued development of the Q-Bank database and began investigating eTraining to use the Q-Bank system.

Staff: Jennifer Hunter Childs (x34927), Jennifer Beck, Yuling Pan, Patricia Goerman

H. Health Insurance Measurement
Description: The U.S. health care system is a patchwork of public and private programs and plans, thus there are no definitive centralized records on the number of individuals without insurance. Researchers must rely on surveys for this estimate, and the Current Population Survey (CPS) is the most widely-cited source for this statistic. It is not without its critics, however, and recent official reports have included caveats regarding the data quality. The purpose of this research is to identify particular features of the CPS questionnaire that are associated with measurement error, and to explore alternative designs to reduce that error.

Highlights: As part of a broad effort to assess health insurance measurement error in Census Bureau surveys, a large-scale field experiment was conducted in March 2010 which included three panels: one with questions modeled on the American Community Survey, one with questions modeled on the Current Population Survey, and one with an experimental set of questions on health insurance. Staff led efforts on writing the Office of Management and Budget clearance package, instrument design, development and testing (including users, systems, and verifications tests), specifications for sampling (from both an RDD frame and a list of Medicare enrollees), developing interviewer training materials and conducting training, and specifying and monitoring daily interviewing production reports. With regard to evaluation methods, staff developed respondent debriefing questions included in the automated instrument, organized the tape recording of a sample of interviews for later behavior coding, conducted interviewer debriefings, and developed and implemented a hybrid open/closed-ended question-by-question log for interviewers to record their feedback as the interviewing progressed. Staff wrote specifications for data processing and output and reviewed the final dataset. Staff then calculated response rates and prepared a final methodology report. Staff presented a paper on findings from a pretest of the experimental panel at the Survey Methodology Brown Bag series. Finally, on a related topic, staff wrote and presented a paper examining the role of state-specific program names for Medicaid and other government-assistance health programs in the ACS and CPS.

Staff: Joanne Pascale (x34920)

I. Emerging Social Trends on Household Structure and Living Situations, Race/Ethnicity, and Linkages to Enumeration Methods and Coverage
Description: In 2006, the National Academies of Science (NAS) Panel on Residence Rules recommended that the Census Bureau establish a trends office with an ongoing

research program on social trends, enumeration methods, and coverage. This program would include monitoring emerging social trends and their impact on the accuracy of basic residence information and census coverage. It would also include developing, conducting, and synthesizing new research to suggest changes in enumeration methods and improve census coverage. Specifically recommended ongoing research topics include: "research on changing factors influencing people's attachments to locations where they are counted," "living situations," "large households," "sources of omissions in the census, as well as duplications," and "questionnaire strategies" (NRC 2006: 175-178).

Highlights: During FY 2010, in response to Director Groves' request for staff to identify challenges under the new Corporate Priorities Program, we submitted a brief proposal to set up an inter-disciplinary Social Trends Office. This office would identify emerging social, demographic and economic trends affecting households and living situations and their linkages to enumeration methods and coverage. The aim would be to analyze our own Census Bureau data to identify emerging trends, evaluate how well our questions and measures are reflecting those trends, and develop and conduct research to revise our questions and measures accordingly to improve our documentation. This is included as our division's Challenge 7, posted for the last six months on the Director's site at <http://cww.census.gov/ioe/challenges.asp>.

Staff: Laurie Schwede (x32611)

J. Using Vignettes to Explore Survey Concepts
Description: Vignettes are a common tool for survey pretesting. They depict hypothetical situations and allow us to evaluate concepts without actually having to recruit people in those situations. Vignettes are also useful when evaluating survey topics that may be highly sensitive. This research will identify and explore how teens classify their contacts with online strangers and the degree to which they are aware of the danger in such interactions. The study will be a mixed-design qualitative and quantitative study. Participants will classify vignettes depicting online contacts with strangers and online contacts with non-strangers as being either appropriate and harmless or inappropriate and dangerous. Participants will also answer open-ended questions about why they feel these contacts are or are not dangerous. The results of the vignette classification task and the open-ended questions will help to identify how teens conceptualize their online relationships and reveal potential online vulnerabilities.

Highlights: During FY 2010, staff completed a preliminary content analysis on previously collected data. The paper discussed the psychological framework for why vignettes are a useful tool for uncovering people's judgments and decision-making processes. The paper

interpreted findings from pretesting vignettes on online activities and how they reveal information about how kids and teens think about activities that are relevant to their vulnerability to Internet predators. These findings suggest several key variables for researchers to explore in an effort to improve education about internet crimes against children.

Staff: Jennifer Beck (x31736), Terry DeMaio

K. Retrieval Effects on Judgments about Knowledge
Description: Surveys are a common way to collect information on a variety of topics. It is easy to assume that if people understand the intended meaning of and know the answer to a survey question, they should have relatively little problem providing an accurate answer. However, research on human memory and knowledge assessment casts significant doubt on this assumption. Context, in the form of both situational variables and individual differences, can have a significant effect on how accurately people answer questions.

In an attempt to investigate the effects of these variables on evaluations of knowledge, we have developed a set of experiments that will investigate the effects of retrieval context on how people evaluate their knowledge of general, factual information. This research will be jointly conducted with researchers at SUNY Stony Brook.

Highlights: During FY 2010, staff collected and analyzed data. The results show that the framing and retrieval fluency manipulations did not appear to affect the participants' average Feeling-of-Knowing (FOK) ratings. Participants were also still able to predict their knowledge accurately. However, in the condition where people thought about a few times when their sense of their knowledge was not accurate, there were fewer, more accurate high-confidence judgments. These results may suggest that a framing and retrieval fluency interaction have differential effects on metacognition.

Staff: Jennifer Beck (x31736)

Research Support and Assistance
This staff provides substantive support in the conduct of research, research assistance, technical assistance, and secretarial support for the various research efforts.

Staff: Matthew Clifton, Matt Gore (HRD), George Higbie, Temika Holland, Gloria Prout, Lorraine Randall, Kelly Taylor, Sarah Wilson

3. PUBLICATIONS

3.1 JOURNAL ARTICLES, PUBLICATIONS

Chan, A. and Pan, Y. (2010). "The Use of Cognitive Interviewing to Explore the Effectiveness of Advance Supplemental Materials among Five Language Groups." *Journal of Field Methods.*

Childs, J. and Goerman, P. (2010). "Bilingual Questionnaire Evaluation and Development through Mixed Pretesting Methods: The Case of the U.S. Census Nonresponse Follow-up Instrument." *Journal of Official Statistics, Vol 26, No 3.*

Fitzsimmons, P. and McElroy, T. (2010). "On Joint Fourier-Laplace Transforms." *Communications in Statistics, 39*, 1883-1885.

Goerman, P. and Casper, R. (2010). "A Preferred Approach for the Cognitive Testing of Translated Materials: Testing the Source Version as a Basis for Comparison." *International Journal of Social Research Methodology, 13.4,* 303-316.

Goerman, P. and Clifton, M. (In press). "The Use of Vignettes in Cross-Cultural Cognitive Testing of Survey Instruments." *Field Methods.*

Herzog, T. N., Scheuren, F., and Winkler, W. E. (2010). "Record Linkage." *Wiley Interdisciplinary Reviews: Computational Statistics, 2, Issue 5.*

Herzog, T. N., Scheuren, F., and Winkler, W. E. (2010). "Data Quality." *Wiley Interdisciplinary Reviews: Computational Statistics*, *2, Issue 5.*

Holan, S. and McElroy, T. (In press). "Tail Exponent Estimation via Broadband Log Density-Quantile Regression." *Journal of Statistical Planning and Inference*.

Klein, M., Sinha, B., and Subramaniam, R. (In press). "Statistical Inferences from Formaldehyde DNA-protein Cross-link Data: Improving Methods for Characterization of Uncertainty." *Journal of Biopharmaceutical Statistics.*

Ma, Y., Hart, J. D., Janicki, R., and Carroll, R. J., (In press). "Local and Omnibus Tests in Classical Measurement Error Models." *Journal of the Royal Statistics Society*, *Series B, 73.*

Maples, J., et al. (In press). "Bayesian Benchmarking with Applications to Small Area Estimation." *TEST.*

McElroy, T. (In press). "A Nonlinear Algorithm for Seasonal Adjustment in Multiplicative Component Decompositions." *Studies in Nonlinear Dynamics and Econometrics*.

McElroy, T. (In press). A Nonparametric Method for Asymmetrically Extending Signal Extraction Filters." *Journal of Forecasting*.

McElroy, T. and Findley, D. (In press). "Discerning Between Models Through Multi-Step Ahead Forecasting Errors." *Journal of Statistical Planning and Inference*.

McElroy, T. and Trimbur, T. (In press). "On the discretization of continuous-time filters for nonstationary stock and flow time series." *Econometric Reviews*.

McElroy, T. and Wildi, M. (2010). "Signal Extraction Revision Variances as Goodness-of-Fit Measure." *Journal of Time Series Econometrics, 2.1.*

McElroy, T. and Wildi, M. (In press). "Signal Extraction Revision Variances as a Goodness-of-Fit Measure." *Journal of Time Series Econometrics*.

Mulry, M. and Cantwell, P. (2010). "Overview of the 2010 Census Coverage Measurement Program and Its Evaluations." *Chance, 46, No 3,* 46-51.

Nagaraja, C., et al. (In press). "An Autoregressive Approach to House Price Modeling." *The Annals of Applied Statistics.*

Nayak, T., Sinha, B., and Zayatz, L. (In press). "Statistical Properties of Multiplicative Noise Masking for Confidentiality Protection." *Journal of Applied Statistics.*

Pan, Y. (In press). "Epilogue for Special Issue on Institutional Politeness in (South) East Asia." *Journal of Asian Pacific Communication, 20.2.*

Roberts, C. G., Holan, S. H., and Monsell, B. C. (2010). "Comparison of X-12-ARIMA Trading Day and Holiday Regressors with Country Specific Regressors." *Journal of Official Statistics, 26.2,* 371–394.

Romano, J. C., Howard, J. H. Jr., and Howard, D. V. (2010). "One-year Retention of General and Sequence-Specific Skills in a Probabilistic, Serial Reaction Time Task." *Memory, 18,* 427-441.

Schwede, L. (2010). "Who is Counted? Subpopulation Coverage in U.S. Censuses." *Anthropology News.* 5-6.

Shao, J. and Thompson, K. (2009). "Variance Estimation in the Presence of Nonrespondents and Certainty Strata." *Survey Methodology, 35 No. 2,* 215-225.

Terry, R. L. and Winston, C. E. (2010). "Personality Characteristic Adaptations: Multiracial Adolescents' Patterns of Racial Self-Identification Change." *Journal of Research on Adolescence, 20.2,* 432-455.

Willimack, D. and Nichols, E. (2010). "A Hybrid Response Process Model for Business Surveys." *Journal of Official Statistics, 26.1,* 3-24.

3.2 BOOKS/BOOK CHAPTERS

Chan, Anna. (In press). "Discourse Analysis of the Indirectness of Chinese Speakers' Responses to Survey Interview Questions." *Chinese Discourse and Interaction: Theory and Practice,* edited by Y. Pan and D. Z. Kádár. London: Equinox Publishing, Ltd.

Lucero, J. and Zayatz, L. (2010). "The Microdata Analysis System at the U.S. Census Bureau." *Privacy in Statistical Databases 2010 LNCS 6344*, Springer, 234-248.

Pan, Yuling and Dániel Z. Kádár. (2010). *Politeness in Historical and Contemporary Chinese Communication.* London/New York: Continuum International Publishing.

3.3 PROCEEDINGS PAPERS

2009 American Association for Public Opinion Research Meeting, Hollywood, FL, May 14-18, 2009.
2009 Proceedings of the American Statistical Association.

- Chan, A. and Pan, Y. (2009). "The Use of Cognitive Interviewing to Explore the Effectiveness of Advance Supplemental Materials among Five Language Groups," 5822-5836.
- Goerman, P. and Clifton, M. (2009). "Vignettes in Cross-Cultural Cognitive Testing: Adaptation for Spanish-Speaking Respondents of Lower Educational Levels," 5794-5806.
- Pan, Y. and Landreth, A. (2009). "Conveying Translated Informed Consent Concepts: Effects of Language and Culture on Interpretation of Legally Required Messages," 5902-5916.
- Park, H., Sha, M., and Pan, Y. (2009). "Cognitive Testing as a Method of Pre-Testing Questionnaires in High and Low Context Cultures," 5676-5690.
- Schwede, L. and Sorokin, A. (2009). "To Link or Not to Link?: Exploring Approaches to Maintaining

American Community Survey Response Rates During Census 2010," 5844-5858.
- Sha, M. and Pan, Y. (2009). "The Use of Vignettes in Evaluating Multilingual Questionnaires," 6122-6133.

2009 Joint Statistical Meetings (American Statistical Association), Washington, D.C., August 1-6, 2009.
2009 Proceedings of the American Statistical Association.
- Childs, J., Norris, D., and Jurgenson, N. (2009). "Pretesting 2010 Census Questionnaires for People with Atypical Living Situations," 1880-1891.
- Findley, D. (2009). "Stock Series Holiday Regressors Generated by Flow Series Holiday Regressors," 477-486.
- Huang, E. and Bell, W. (2009). "A Simulation Study of the Distribution of Fay's Successive Difference Replication Variance Estimator," 5294-5309.
- Ikeda, M. (2009). "Developing Guidelines Based on CVs for when One-Year Estimates Can Be Used Instead of Three-Year Estimates in the American Community Survey," 1670-1677.
- Lucero, J., Singh, L., and Zayatz, L. (2009). "The Current State of the Microdata Analysis System at the Census Bureau," 3762-3772.
- Maples, J., Bell, W., and Huang, E. (2009). "Small-Area Variance Modeling with Application to County Poverty Estimates from the American Community Survey," 5056-5067.
- Massell, P. (2009). "An Overview of Uncertainty Creation to Protect Statistical Data," 1550-1558.
- McDonald-Johnson, K., Findley, D., and Cepietz, E. (2009). "Investigating Quarterly Trading Day Effects," 3305-3317.
- McElroy, T. and Holan, S. (2009). "The Detection of Cycles in Raw and Seasonally Adjusted Data," 702-707.
- Monsell, B. (2009). "Update on the Development of X-13ARIMA-SEATS," 1907-1920.
- Mule, V., Malec, D., Imel, L., Nguyen, N., and Moldoff, M. (2009). "Missing Data Methods for the CCM Component Error Estimation," 3265-3278.
- Mulry, M. and Adams, T. (2009). "Overview of Evaluations of the 2010 Census Coverage Measurement Program," 3117-3128.
- Mulry, M. and Oliver, B. (2009). "A Simulation Study of Treatments of Influential Values in the Monthly Retail Trade Survey," 2979-2993.
- Mulry, M. and Olson, T. (2009). "Analyses for Partnerships Based on the Census Barriers, Attitudes, and Motivator Survey," 5727-5741.
- Schwede, L., Sorokin, A., and Wake, V. (2009). "'You Really Have to Puzzle This Out': Checking Residence and Coverage Duplications on a Census 2010 Questionnaire," 4504-4518.
- Slud, E. and Thibaudeau, Y. (2009). "Simultaneous Calibration and Nonresponse Adjustment," 2263-2272.
- Titova, N. and Monsell, B. (2009). "Detecting Stock Calendar Effects in U.S. Census Bureau Inventory Series," 2108-2122.
- Weidman, L. and Ashenfelter, K. (2009). "Results of a Survey on Choice of Sampling Error Display in American Community Survey Data Products," 1137-1150.
- Winkler, W. (2009). "General Discrete-data Modeling Methods for Producing Synthetic Data with Reduced Re-identification Risk that Preserve Analytic Properties," 2568-2576.

2009 Proceedings of the Federal Conference on Statistical Methodology, Washington, D.C., November 2-4, 2009.
- Bates, N. and Pan, Y. (2009). "Motivating Non-English-Speaking Populations for Census and Survey Participation."
- McElroy, T. and Holan, S. (2009). "Using Spectral Peaks to Detect Seasonality."
- Pascale, J. (2009). "Findings from a Pretest of a New Approach to Measuring Heath Insurance in the Current Population Survey."

Proceedings of the Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality, Bilbao, Spain, December 2-4, 2009.
- Zayatz, L., Lucero, J., Massell, P., and Ramanayake, A. (2009). "Disclosure Avoidance for Census 2010 and American Community Survey Five-year Tabular Data Products."

3.4 STATISTICAL RESEARCH DIVISION RESEARCH REPORTS
<http://www.census.gov/srd/www/byyear.html>

RR (Statistics #2009-07): Christopher G. Roberts, Scott H. Holan, and Brian Monsell, "Comparison of X-12-ARIMA Trading Day and Holiday Regressors With Country Specific Regressors," October 8, 2009.

RR (Statistics #2009-08): William E. Winkler, **"**Sample Allocation and Stratification," November 16, 2009.

RR (Statistics #2009-09): Jason Lucero, Lisa Singh, and Laura Zayatz, "Recent Work on the Microdata Analysis System at the Census Bureau," November 16, 2009.

RR (Statistics #2009-10): Laura Zayatz, Jason Lucero, Paul Massell, and Asoka Ramanayake, **"**Disclosure Avoidance for Census 2010 and American Community Survey Five-year Tabular Data Products," November 23, 2009.

RR (Statistics #2010-01): Tucker S. McElroy and David F. Findley, "Selection Between Models Through Multi-Step-Ahead Forecasting," January 28, 2010.

RR (Statistics #2010-02): William E. Winkler, "General Discrete-data Modeling Methods for Producing Synthetic Data with Reduced Re-identification Risk that Preserve Analytic Properties," January 28, 2010.

RR (Statistics #2010-03): Eric V. Slud and Yves Thibaudeau, "Simultaneous Calibration and Nonresponse Adjustment," February 22, 2010.

RR (Statistics #2010-04): Asoka Ramanayake and Laura Voshell Zayatz, "Balancing Disclosure Risk with Data Quality," February 22, 2010.

RR (Statistics #2010-05): Tapan K. Nayak, Bimal Sinha, and Laura Voshell Zayatz, "Statistical Properties of Multiplicative Noise Masking for Confidentiality Protection," March 5, 2010.

RR (Statistics #2010-06): Tucker S. McElroy and Marc Wildi, "Signal Extraction Revision Variances as a Goodness-of-Fit Measure," March 5, 2010.

RR (Statistics #2010-07): Mary H. Mulry and Timothy P. Olson, "Analyses for Partnerships Based on the Census Barriers, Attitudes, and Motivators Survey," March 11, 2010.

RR (Survey Methodology #2009-07): Joanne Pascale, **"**Findings from a Pretest of a New Approach to Measuring Health Insurance in the Current Population Survey," November 16, 2009.

RR (Survey Methodology #2009-08): Leticia Fernández, Eleanor Gerber, Matt Clifton, George Higbie, and Mikelyn Meyers, "Cognitive Pretesting of 2010 Alternative Questionnaire Experiment (AQE) Race and Hispanic Origin Treatment Panels," November 24, 2009.

RR (Survey Methodology #2009-09): Jeffrey Moore, Jason Fields, Gary Benedetto, Martha Stinson, Anna Chan, and Jerry Maples, **"**The 2008 Survey of Income and Program Participation Event History Calendar Field Test: Study Design and Initial Results," December 30, 2009.

RR (Survey Methodology #2010-01): Kent H. Marquis and Jeffrey C. Moore, "Measurement Errors in SIPP Program Reports," January 14, 2010.

RR (Survey Methodology #2010-02): Jennifer Beck, "On the Usefulness of Pretesting Vignettes in Exploratory Research," January 21, 2010.

RR (Survey Methodology #2010-03): Jeffrey C. Moore and Daniel Kasprzyk, "Month-to-Month Recipiency Turnover in the ISDP," April 11, 2010.

RR (Survey Methodology #2010-04): Jeffrey C. Moore, Kent H. Marquis, and Karen Bogen, "The SIPP Cognitive Research Evaluation Experiment: Basic Results and Documentation," April 12, 2010.

RR (Survey Methodology #2010-05): Dawn R. Von Thurn, Jeffrey C. Moore, and Elizabeth A. Martin, "National Health Interview Survey Redesign: An Anthropological Investigation of Mental Health Concepts," April 12, 2010.

RR (Survey Methodology #2010-06): Jeffrey C. Moore, Karen Bogen, and Julie Klein Griffiths, "Wave 8 Adult Well-Being Topical Module: Cognitive Interview Results and Recommendations," April 12, 2010.

RR (Survey Methodology #2010-07): Jeffrey C. Moore, Karen Bogen, and Kent H. Marquis, "A 'Cognitive' Interviewing Approach for the Survey of Income and Program Participation: Development of Procedures and Initial Test Results," June 9, 2010.

RR (Survey Methodology #2010-08): Jeffrey C. Moore, "Evaluating the Public Information Campaign for the 1980 Census-Results of the KAP Survey," June 9, 2010.

RR (Survey Methodology #2010-09): Jeffrey C. Moore, "Proxy Reports: Results from a Record Check Study," 2010.

RR (Survey Methodology #2010-10): Jeffrey C. Moore and Kent H. Marquis, "Using Administrative Record Data to Evaluate the Quality of Survey Estimates," June 9, 2010.

RR (Survey Methodology #2010-11): Jeffrey C. Moore and Kent H. Marquis, "Using Administrative Record Data to Describe SIPP Response Errors," June 9, 2010.

RR (Survey Methodology #2010-12): Kent H. Marquis and Jeffrey C. Moore, "SIPP Record Check Results: Implications for Measurement Principles and Practice," June 9, 2010.

RR (Survey Methodology #2010-13): Robert E. Fay, Nancy Bates, and Jeffrey C. Moore, "Lower Mail Response in the 1990 Census: A Preliminary Interpretation," June 15, 2010.

RR (Survey Methodology #2010-14): Jeffrey C. Moore, "Report of the Workshop on Applying Cognitive Psychology to Recall Problems of the National Crime Survey," June 15, 2010.

RR (Survey Methodology #2010-15): Eleanor R. Gerber, "Summary Report of Cognitive Testing on Residence Rules Instructions," June 29, 2010.

RR (Survey Methodology #2010-16): Eleanor R. Gerber, "Residence: A Cognitive Approach to Household Membership Judgments Among Low Income Blacks," June 29, 2010.

RR (Survey Methodology #2010-17): Eleanor R. Gerber, "The Language of Residence: Respondent Understandings and Census Rules," June 29, 2010.

RR (Survey Methodology #2010-18): Jeffrey C. Moore, "Effects of the 1986 Motivational Insert on Census Form Mail Response ," June 29, 2010.

3.5 STATISTICAL RESEARCH DIVISION STUDIES
<http://www.census.gov/srd/www/byyear.html>

SS (Survey Methodology #2009-16): Yuling Pan and Ashley Landreth, "Conveying Translated Informed Consent Concepts: Effects of Language and Culture on Interpretation of Legally Required Messages," October 15, 2009.

SS (Survey Methodology #2009-17): Jennifer Hunter Childs, Leticia Fernández, Matt Clifton, and Mikelyn Meyers, "CFU CPEX Experimental Question Cognitive Testing: Undercount, Overcount and Duplicate Experimental Question Sequences," November 16, 2009.

SS (Survey Methodology #2010-01): Kathleen T. Ashenfelter, "Eye-tracking Study Report: Examining User Patterns for Demographic Items on the 2007 and 2008 ACS Mail Forms," January 5, 2010.

SS (Survey Methodology #2010-02): Jennifer Chen, Jenna Beck, Jennifer Romano, and Elizabeth D. Murphy, "Usability Evaluation of the Governments Division Public Web Site," January 21, 2010.

SS (Survey Methodology #2010-03): Jennifer Romano, Jennifer Chen, Erica L. Olmsted-Hawala, and Elizabeth D. Murphy, "A Card-Sorting Study for the History of the Census Bureau: 'Sights and Sounds: Photos' Web Page," January 21, 2010.

SS (Survey Methodology #2010-04): Kathleen T. Ashenfelter, "Data Reliability Indicator Based on the Coefficient of Variation: Results from the Second Round of Testing," March 31, 2010.

SS (Survey Methodology #2010-05): Jennifer C. Romano and Jennifer M. Chen, "A Usability Evaluation of the Online National Survey for College Graduates (NSCG)," April 19, 2010.

SS (Survey Methodology #2010-06): Elizabeth Nichols, Jennifer Hunter Childs, and Kyra Linse, "RDD versus Site Test: Mode Effects on Gathering a Household Roster and Alternate Addresses," April 19, 2010.

SS (Survey Methodology #2010-07): Jennifer C. Romano, Jennifer M. Chen, Erica L. Olmsted-Hawala, and Elizabeth D. Murphy, "A Medium-Fidelity Usability and Eye-Tracking Evaluation of Iteration 2.0 and Iteration 2.5 of the New American FactFinder Web Site: Capabilities and Functions," July 28, 2010.

SS (Survey Methodology #2010-08): Nancy Bates and Yuling Pan, "Motivating Non-English-Speaking Populations for Census and Survey Participation," August 20, 2010.

SS (Survey Methodology #2010-09): Jennifer M. Chen, Temika Holland, Erica L. Olmsted-Hawala, and Andy Su. "Baseline Usability Evaluation of the American Community Survet Web Site with Novices and Expert Users." September 14, 2010.

SS (Survey Methodology #2010-10): Kathleen T. Ashenfelter and Victor Quach. "Data Reliability Indicator Based on the Coefcient of Variation: Report for the Third Round of Usability Testing." September 16, 2010.

SS (Survey Methodology #2010-11): Theresa DeMaio, Kathleen Denny, Nathan Jurgenson, Andrea Schwanz. "Pretesting of the 2011 School Crime Supplement: Final Results and Recommendations." September 22, 2010.

SS (Survey Methodology #2010-12): Laurie Schwede and Anissa Sorokin. "Cognitive Testing Results for the ACS Field Flyer for Use in 2010: ACS Messaging Project Phase 3 Final Report." September 22, 2010.

3.6 OTHER REPORTS

Lucero, J. (2009). "Confidentiality Rules for Universe Formation and Geographies for the Microdata Analysis System," *Statistical Research Division Census Confidential Report No. CCRR-2009/01.*

Lucero, J. (2010). "Evaluation of the Effectiveness of the Drop Q Rule Against Differencing Attack Disclosures," *Statistical Research Division Census Confidential Report No. CCRR-2010/03.*

Nichols, B. (2010). "Observation of the 2010 Census Coverage Measurement Initial Housing Unit Followup (IHUFU) Operation in Maryland on March 9, 2010," *DSSD 2010 Census Coverage Measurement Memorandum Series #2010-F-18.*

Ramanayake A. and Appelbaum S. (2010). "An Evaluation of the ACS Swapping Procedure - 2005-2008," *Statistical Research Division Census Confidential Report No. CCRR-2010/01.*

Ramanayake A. and Appelbaum S. (2010). "Swapping Evaluation at County Level for the State of California Using 2007 ACS Data," *Statistical Research Division Census Confidential Report No. CCRR-2010/02.*

4. TALKS AND PRESENTATIONS

UNECE Work Session on Statistical Data Editing, Neuchâtel, Switzerland, October 5-7, 2009.
- Maria Garcia and William Winkler, "Determining a Set of Edits and Quality of a Database."

The Washington Statistical Society, Washington, D.C., October 8, 2009.
- Yuling Pan, "The Sociolinguistics of Survey Translation." (Invited lecture.)

Eurostat Training Session, Luxembourg, Luxembourg, October 20-22, 2009.
- Tucker McElroy, "X-13 ARIMA-SEATS."

Statistics Canada Symposium 2009 on Longitudinal Surveys: From Design to Analysis, Gatineau, Canada, Oct. 27-29, 2009.
- Eric Slud and Leroy Bailey, "Evaluation and Selection of Models for Attrition Nonresponse Adjustment." (Invited talk.)

2009 Conference for the Federal Committee on Statistical Methodology, Washington, D.C., November 2-4, 2009.
- Nancy Bates and Yuling Pan, "Motivating Non-English-Speaking Populations for Census and Survey Participation."
- Tucker McElroy, "Using Spectral Peaks to Detect Seasonality."
- Joanne Pascale, "Findings from a Pretest of a New Approach to Measuring Health Insurance in the Current Population Survey."

Joint UNECE/Eurostat Work Session on Statistical Data Confidentiality, Bilbao, Spain, December 2-4, 2009.
- Laura Zayatz, "Disclosure Avoidance for Census 2010 and American Community Survey Five-year Tabular Data Products."

Annual Conference of the Linguistic Society of America, Baltimore, Maryland, January 9-11, 2010.
- Yuling Pan, "Language and Measurement Research at the U.S. Census Bureau." (Poster presentation).

National Institute of Statistical Sciences/Education Statistics Service Institute Workshop, Washington, D.C., February 16, 2010.
- Laura Zayatz, "U.S. Census Bureau Disclosure Avoidance Practices and Research: An Update for NISS/ESSI."

FedCASIC 2010, Bureau of Labor Statistics, Washington, D.C., March 16-18, 2010.
- Kathleen T. Ashenfelter, "Establishing Cutting-edge Technology in a Federal Government Usability Lab: Eye Tracking at the Census Bureau in 2010."
- Matt Jans, Kathy Creighton, and Chris Laskey, "Using Paradata at the US Census Bureau: Demographic Current Surveys' History and Future."
- Larry Malakhoff, "Accessible Web Survey Tools."

Invited Seminar, Naval Postgraduate School, Monterey, California, April 2, 2010.
- Tucker McElroy, "Goodness of Fit Testing for Time Series in the Frequency Domain."

Cognitive Aging Conference 2010, Atlanta, Georgia, April 15-18, 2010.
- Jennifer Romano, Darlene V. Howard, and James H. Howard, Jr., "Early Bilingualism is Associated with Better Implicit Sequence Learning and Executive Control in Old Age." (Poster).

65th Annual Conference of the American Association for Public Opinion Research (AAPOR), Chicago, Illinois, May 13-16, 2010.
- Anna Chan, Aref N. Dajani, and Yuling Pan, "Survey Item Nonresponse Among American Community Survey's Chinese Respondents."
- Jennifer Hunter Childs and Rodney Terry, "Measuring Race and Hispanic Origin: Cognitive Test Findings Searching for 'Truth'."

- Patricia Goerman, "Number and Nationality of Spanish- Speaking Immigrant Respondents for U.S. Cognitive Testing Studies."
- Joanne Pascale, "Health Insurance Estimates from the CPS versus the ACS: the Role of State-Specific Program Names."
- Laurie Schwede and Anissa Sorokin, "Do Respondents Read Those Key Messages in Our Questionnaire Package Cover Letters?: What Comes Out of the Envelope First Matters."
- Mandy Sha, Hyunjoo Park, and Yuling Pan, "Developing a Systematic Process for Translation Expert Review: The Translation Appraisal System."
- Rodney Terry and Jennifer Hunter Childs, "Exploring Race and Hispanic Origin Questions with Multiracial Individuals: Diverse Responses from a Diverse Population." (Poster).

International Field Directors and Technologies Conference, Chicago, Illinois, May 16-19, 2010.
- Matt Jans, Chris Laskey, and Kathy Creighton, "Using Paradata at the US Census Bureau: Demographic Current Surveys' History and Future."

Conference on Nonparametric Statistics and Statistical Learning, Columbus, Ohio, May 17, 2010.
- William E. Winkler, "Machine Learning for Record Linkage, Text Categorization, and Edit/Imputation."

Monthly Meeting of the Usability Professionals Association, D.C. Chapter, Washington, D.C., May 17, 2010.
- Kathleen Ashenfelter, "Eye Tracking and Web Survey Research."

Quantitative Methods in Defense and National Security, George Mason University, Fairfax, Virginia, May 25-26, 2010.
- Paul Massell, "Releasing Statistical Data while Keeping the Underlying Individual Data Confidential."

2010 International Conference on Cognitive Processes in Language, Łódź, Poland, May 25-27, 2010.
- Kathleen Ashenfelter, "Differential Mental Strategies and Reading Patterns for Demographic Surveys: Question Organization and Its Impact on Response Strategy."

Association for Psychological Science Annual Convention, Boston, Massachusetts, May 27-30, 2010.
- Jennifer Beck, "The Effects of Framing and Retrieval Fluency on Feeling-of-Knowing (FOK) Judgments." (Poster).

2010 International Total Survey Error Workshop (ITSEW), Stowe, Vermont, June 13- 16, 2010.
- Mary H. Mulry, "The Structure of Error Components in 2010 Census Coverage Error Estimation: P-sample Estimates."

International Symposium on Forecasting, San Diego, California, June 20-23, 2010.
- Tucker McElroy, "Discerning Between Models through Multi-Step Ahead Forecasting Error."
- Brian Monsell, "Update on the Development of X-13ARIMA-SEATS."

Joint Statistical Meetings, American Statistical Association, Washington D.C., July 31-August 5, 2010.
- William Bell and Ekaterina Sotiris, "Seasonal Adjustment to Facilitate Forecasting: Empirical Results."
- Anna Chan and Yuling Pan, "Analysis of Chinese Speakers' Responses to Survey Intention Questions."
- Scott Holan and Tucker McElroy, "On the Seasonal Adjustment of Long Memory Time Series."
- Elizabeth Huang and William Bell, "Further Simulation Results on the Distribution of Some Survey Variance Estimators."
- Ryan Janicki, "Estimation and Hypothesis Testing in Submodels Using Fisher Estimating Functions."
- Matt Jans, "Verbal Paradata and Survey Question Content: How Question Sensitivity and Cognitive Complexity Influence the Way Answers (and Nonanswers) to Survey Questions are Delivered."
- Martin Klein and Robert Creecy, "Steps Toward Creating a Fully Synthetic Decennial Census Microdata File."
- Donald Malec, "An Evaluation of Housing Unit and Block Cluster Effects on Small-Area Census Coverage Variability in the 2006 Census Test."
- Tucker McElroy, "Fixed-Bandwidth Asymptotics for the Studentized Mean for Long and Negative Memory Time Series."

- Brian Monsell, "An Evaluation of Kan and Wang's Adjusted Box-Pierce Test Using Seasonal Time Series."
- Mary Mulry and Bruce Spencer, "Estimating Nonsampling Errors in Estimates of Omissions and Erroneous Enumerations in the 2010 Census."
- Yuling Pan, "Developing Multilingual Questionnaires: A Sociolinguistic Perspective."
- Joanne Pascale, "Health Insurance Estimates from the CPS vs. ACS: An Analysis of State-Specific Program Names."
- Yves Thibaudeau, Eric Slud, and Alfred Gottschalck, "Modeling Log-Linear Restricted Conditional Probabilities for Prediction in Surveys."
- Natalya Titova, David Findley, and Brian Monsell, "Comparing the Causey-Trager Method to the Dagum-Cholette Regression Method of Benchmarking Subannual Data to Annual Benchmarks."
- Laura Zayatz, Paul Massell, Jason Lucero, and Asoka Ramanayake, "Disclosure Avoidance for Census 2010 and American Community Survey Five-Year Tabular Data Products."

5. STATISTICAL RESEARCH DIVISION SEMINAR SERIES

Seminar Series Team: Aref Dajani, Richard Griffin (DSSD), Paul Massell,
Laurie Schwede, Katherine Thompson (ADEP)

Partha Lahiri, (ASA/NSF/Census Research Fellow) University of Maryland, College Park, "Hierarchical Modeling and Related Inferential Issues In Small Area Estimation," October 21, 2009.

Jaejik Kim, University of Georgia, "Clustering Techniques for Histogram-valued Data," October 27, 2009.

Erica Olmsted-Hawala, SRD, U.S. Census Bureau, "Think-Aloud Protocols: A Comparison of Three Think-Aloud Protocols for Use in Testing Data Dissemination Web Sites for Usability," November 9, 2009.

Jennifer Romano, SRD, U.S. Census Bureau, "Benefits of Iterative Usability Testing on a Web Site," November 9, 2009.

Larry R. May and Rahul Tikekar, Office of Research, Internal Revenue Service, "Graph Query: A Pattern Matching Tool to Detect Structures on Linked Data," November 12, 2009.

Anton Korinek, University of Maryland, and Johan Mistiaen, World Bank, "Survey Nonresponse and the Distribution of Income," December 7, 2009.

Victoria Leaver, Australian Bureau of Statistics, "Statistical Disclosure Control: Methods and Directions in the ABS," December 14, 2009.

Joseph Sakshaug, University of Michigan, (U.S. Census Bureau Dissertation Fellow), "Synthetic Data for Small Area Estimation in the American Community Survey: A Preliminary Evaluation," February 23, 2010.

Andrew Fiore, University of California, Berkeley, "Assessing Attraction and Interpersonal Judgments in Online Dating Through a Web-Based Survey," March 2, 2010.

Virginia Wake, SRD, U.S. Census Bureau, "A Sociolinguistic Approach to Institutional Interaction Among Three People," March 15, 2010.

Courtney Kennedy, University of Michigan, (U.S. Census Bureau Dissertation Fellow), "Measurement Error in Cell Phone Surveys," March 18, 2010.

William Winkler, SRD, U.S. Census Bureau, "Machine Learning for Record Linkage, Text Categorization, and Edit/Imputation," May 5, 2010.

Barbara Anderson, University of Michigan, *SUMMER AT CENSUS,* "Adult Mortality in South Africa, 1997-2005: AIDS and Other Causes," May 13, 2010.

Gauri Datta, University of Georgia, *SUMMER AT CENSUS,* "Model Selection by Testing for the Presence of Small-Area Effects in Area-Level Data," May 18, 2010.

Jennifer Van Hook, The Pennsylvania State University, *SUMMER AT CENSUS,* "Estimates of the Mexican Foreign-born Population: Comparisons of ACS with Mexican Census and Administrative Data," May 19, 2010.

Kathleen Ashenfelter, SRD, U.S. Census Bureau, "Moving from Paper to the Internet: Preliminary Results from Usability Testing of an Online Version of the ACS," May 20, 2010.

Byran Smucker, The Pennsylvania State University (U.S. Census Bureau Dissertation Fellow), "Model-Robust Experimental Design," May 24, 2010.

Tzee-Kiu Edwin Chow, Texas State University – San Marcos, *SUMMER AT CENSUS,* "What Does the Web Have to Do With Counting?," May 27, 2010.

Robert V. Remini, Historian of the U.S. House of Representatives, *SUMMER AT CENSUS,* "The Compromise of 1850," May 27, 2010.

Bikas K. Sinha, Indian Statistical Institute, Kolkata, *SUMMER AT CENSUS,* "Estimation of the Size of a Finite Population," June 1, 2010.

William Seltzer, Fordham University, *SUMMER AT CENSUS,* "The Role of Ethics in a Federal Statistical Agency with Special Reference to the U.S. Census Bureau," June 9, 2010.

Margo Anderson, University of Wisconsin-Milwaukee, *SUMMER AT CENSUS,* "The Census in American History: A Story of Success . . . Mostly, and How Census Takers in the Past Dealt with the Inevitable Bumps in the Road," June 10, 2010.

Martin Klein, SRD, U.S. Census Bureau, "Imputation in the Survey of Research and Development in Industry Using a Last-Value-Dependent Nonresponse Mechanism," June 16, 2010.

Robert H. Shumway, University of California, Davis, *SUMMER AT CENSUS*, "Analysis of Longitudinal Mixed Dynamic Models, July 8, 2010.

Irma Elo, University of Pennsylvania, *SUMMER AT CENSUS*, "Early Life Family and Socioeconomic Conditions and Cause-Specific Mortality in Finland," July 13, 2010.

Lily Wang, University of Georgia, *SUMMER AT CENSUS*, "Semiparametric Marginal Mean Models for Longitudinal Survey Data," July 13, 2010.

Zhen-Qing Chen, University of Washington, *SUMMER AT CENSUS*, "Metropolis Algorithm and Distorted Brownian Motion," July 14, 2010.

Scott Holan, University of Missouri, *SUMMER AT CENSUS*, "A Bayesian Approach to Long Memory," July 15, 2010.

John W. Emerson, Yale University, "The Bigmemory Project," July 20, 2010.

Jerry Reiter, Duke University, *SUMMER AT CENSUS*, "Sampling with Synthesis: A New Approach to Releasing Public Use Files of Census Data," July 20, 2010.

Amanda Holmes, Baylor University, "Interview and Test Procedures Affect Confidence and Accuracy in Eyewitness Memory for Product Brand Identification," July 26, 2010.

Joel Andress, Florida State University, "Shared Mental and Physical Health Risks among Spouses," July 27, 2010.

Cleo Redline (U.S. Census Bureau Dissertation Fellow) University of Maryland, College Park, "Clarifying Survey Questions," July 28, 2010.

Rachel Freidus, University of Maryland, College Park, "The Goal of Low Self-Monitors: To Thine Own Self Be True?" July 29, 2010.

Michael Kubovy, University of Virginia, *SUMMER AT CENSUS*, "On Narratives and the Structure of Human Lives," August 9, 2010.

Jean-Francois Beaumont, Statistics Canada, *SUMMER AT CENSUS*, "Variance Estimation Under Composite Imputation: The Methodology Behind SEVANI," August 10, 2010.

Michael Kubovy, University of Virginia, *SUMMER AT CENSUS*, "Audio-visual Objects," August 11, 2010.

Wendy Rogers, Georgia Institute of Technology, *SUMMER AT CENSUS*, "Supporting Successful Aging Through Scientific Research on Design and Training," August 16, 2010.

Wendy Rogers, Georgia Institute of Technology, *SUMMER AT CENSUS*, "Involving Older Adults in Research and Usability Studies," August 17, 2010.

William Winkler, SRD, U.S. Census Bureau, "Record Linkage: Introductory Overview," August 17, 2010.

Anna Marie Trester, Georgetown University, "The Language of Silence: Ethnographic Research with a Quaker Vigil," August 25, 2010.

Jennifer Leeman, George Mason University, "The Role of Language in the Classification of Latinos: From Mother Tongue to Hispanic Origin," August 26, 2010.

Antonio Simoes, University of Kansas, "Towards an Interface of Sociolinguistics and Population Analysis," August 27, 2010.

Rachel Levenstein (U.S. Census Bureau Dissertation Fellow) University of Michigan, "Nonresponse and Measurement Error in Mixed-Mode Designs," September 1, 2010.

Yves Thibaudeau, SRD, U.S. Census Bureau, "Modeling Log-Linear Conditional Probabilities for Prediction in Surveys," September 2, 2010.

Tapan Nayak, The George Washington University and SRD, U.S. Census Bureau, "Statistical Properties of Multiplicative Noise Masking for Confidentiality Protection," September 7, 2010.

Judith M. Tanur, Stony Brook University, *SUMMER AT CENSUS*, "An Overview of the Respondent-Generated Intervals (RGI) Approach to Sample Surveys," September 20, 2010.

MINI-CONFERENCE ON INTERNET AND MULTI-MODE DATA COLLECTION
(Co-Sponsors: Statistical Research Division & American Community Survey Office)
September 9, 2010 / Census Bureau
Planning Committee: Kathleen Ashenfelter, Todd Hughes, Tommy Wright (Chair)

- David Cantor, Westat, "Incorporating a Web Option in a Two-Phase Mail Survey."
- Morgan Millar, Washington State University, "Improving Response to Web and Mixed-Mode Surveys: The Effects of Sequential vs. Simultaneous Mode Choice, Incentives, and Contact Strategies Using Mail and E-Mail."
- Virginia Lesser, Oregon State University, "Does Providing a Choice of Survey Modes Influence Response."
- Stacia Jorgensen, University of Nebraska, "The Effectiveness of Multiple Modes Among Young Adults."
- Courtney Reiser and Mike Bentley, "The 2010 Census Quality Survey (CQS)."
- Mary Frances Zelenak, "The 2011 ACS Internet Test."
- Kathleen Ashenfelter, "Mail Materials and Usability Testing for the 2010 CQS and 2011 ACS Internet Test."
- Tim Gilbert, "The National Survey of College Graduates (NSCG)."
- Matt Jans, "Web Surveys and Total Survey Error."
- Deborah Griffin, "Closing Remarks."

6. PERSONNEL ITEMS

6.1 HONORS/AWARDS/SPECIAL RECOGNITION

Fellow, American Statistical Association

- **Laura Zayatz** – "For effectively moving theoretical techniques in the field of statistical disclosure control into practical application in official statistics; for leadership nationally and internationally in educating statisticians and policy officials on relevant disclosure issues, indentifying new research opportunities, and developing innovative technical solutions; and for service to the statistical community."

Silver Medal Award, U.S. Department of Commerce

- **Jennifer Hunter Childs (Team Award)** – While on detail at the Office of the Inspector General as a part of the Executive Leadership Development Program, Mrs. Childs worked on a team that completed a comprehensive review of the service used by the National Marine Fisheries Service to set catch limits for the heavily-fished New England commercial fisheries. For this work, she was awarded a Department of Commerce Silver Medal Award for Meritorious Service.

Bronze Medal Award, U.S. Bureau of the Census

- **Elizabeth Ann Dimler** – "…administers the Census Bureau R&D Contracts Program. The contracts provide a pool of contractors to assist the Census Bureau in conducting research on all sample survey and census methods and processes to improve products and services. Through her leadership, these 5-year contracts allow Census Bureau divisions and offices to obtain outside advisory and assistance services to support their research and development efforts quickly and easily."

- **Theresa DeMaio (Team Award)** – "This Census Bureau team planned and developed a prototype CARI Interactive Data Access System to retrieve and assess sound files to evaluate interviewer performance, respondent reactions to questions, and the interview environment during CAPI and CATI interviewing for the American Community Survey and Demographic reimbursable programs."

- **Yuling Pan and Patricia Goerman (Team Award)** – "In 2009, several topics were identified by Federal agencies for ACS testing and implementation. This team is recognized for managing the process, negotiating versions for testing, analyzing results, integrating all information into a final set of questions, and for the extraordinary measures taken to foster collaboration with other Federal agencies so that their data user needs and agency priorities were heard and taken into account."

- **Michal Ikeda and Edward Porter (Team Award)** – "This group is recognized for their significant contributions to the coverage improvement of the 2010 Census. They developed methodology and operations to identify and remove potential duplicates in the decennial responses."

6.2 SIGNIFICANT SERVICE TO PROFESSION

Kathleen Ashenfelter
- Associate Editor, *Frontiers in Quantitative Psychology and Measurement.*
- Representative, Capital Area Social Psychological Association (CASPA) to the American Association for the Advancement of Science's (AAAS) Science and Human Rights Coalition.
- Reviewed papers for the Computer Human Interaction (CHI) conference.

Jen Beck
- Refereed a paper for *Journal of Official Statistics.*

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

Anna Chan
- Refereed a paper for *Public Opinion Quarterly.*

Jenny Childs
- Refereed papers for *Field Methods* and *Journal of Official Statistics.*

Terry DeMaio
- Refereed papers for the *Journal of Official Statistics* and *Public Opinion Quarterly*.

Ryan Janicki
- Refereed papers for *Journal of the Royal Statistical Society, Series B.*

Matt Jans
- Refereed papers for *Journal of Official Statistics* and *Public Opinion Quarterly.*

Martin Klein
- Refereed a paper for *Environmental Modeling and Assessment.*

Jason Lucero
- Member, Confidentiality and Data Access Committee (CDAC).

Donald Malec
- Refereed papers for *Computational Statistics and Data Analysis*, *Journal of the American Statistical Association*, and *Journal of the Royal Statistical Society, Series B.*
- Associate Editor, *Survey Methodology*.

Jerry Maples
- Refereed a paper for *The American Statistician.*
- Reviewed a proposal for the National Science Foundation.

Paul Massell
- Member, Confidentiality and Data Access Committee (CDAC).
- Member, Bureau of Transportation Statistics Disclosure Review Board.

Tucker McElroy
- Refereed papers for the *Journal of Official Statistics, Studies in Nonlinear Dynamics and Econometrics,* the *Journal of Economic Surveys, Computational Statistics and Data Analysis,* the *Journal of Nonparametric Statistics,* the *Journal of the American Statistical Association, Statistics and Probability Letters,* the *Journal of Multivariate Analysis*, *Statistica Sinica,* and *Econometric Theory.*
- Publications Officer, Business and Economic Statistics Section, American Statistical Association.
- Sessions Organizer, 2010 Joint Statistical Meetings.

Brian Monsell
- Webmaster, Business and Economic Statistics Section, American Statistical Association.
- Refereed a paper for *Pakistan Journal of Statistics*
- Sessions Organizer, 2010 Joint Statistical Meetings.

Mary H. Mulry
- Refereed articles for *Public Opinion Quarterly*.
- Past-Chair, Survey Research Methods Section, American Statistical Association.
- Associate Editor, *Journal of Official Statistics*.

Chaitra Nagaraja
- External Examiner, PhD Dissertation in Statistics, George Washington University.

Tapan Nayak
- Member, Committee on Privacy and Confidentiality, American Statistical Association.
- Member, Editorial Board, *Communications in Statistics*.

Erica Olmsted-Hawala
- Refereed a paper for *Interacting with Computers.*
- Reviewed papers, 2010 Usability Professionals Association Meeting.

Yuling Pan
- Refereed manuscripts for the *Encyclopedia of Applied Linguistics.*
- Served as an outside reviewer for a promotion-to-full-professor case for Indiana University-Purdue University.
- Served as an expert in the Ad Hoc Translation Committee for the National Assessment of Educational Progress, housed in the National Center for Education Statistics in the U.S. Department of Education.
- Reviewed a research proposal for the Hungarian Scientific Research Fund.
- Co-organizer, Session on Cross-Cultural Studies, 2010 Joint Statistical Meetings.
- Member, Editorial Board, *Journal of Chinese Language and Discourse.*
- Member, Advisory Board, *Journal of Politeness Research.*
- Member, Editorial Board, *Advances in Pragmatics and Discourse Analysis* Book Series*,* Cambridge Scholar Publishing.
- Member, Multilingual Interest Group, American Association for Public Opinion Research (AAPOR).

Joanne Pascale
- Refereed a papers for *Field Methods Journal*, *Journal of Official Statistics*, and *Inquiry Journal.*

Asoka Ramanayake
- Refereed papers for the *Sri Lankan Journal of Applied Statistics*.
- Member, Confidentiality and Data Access Committee (CDAC).
- Member, NAS CTPP Project Panel for Producing Transportation Data Products from the American Community Survey that Comply with Disclosure Rules.

Rolando Rodriguez
- Refereed a paper for *Survey Methodology.*

Jen Romano
- Reviewed proposals for the American Association for the Advancement of Science (AAAS) 2011 Annual Meeting.
- Refereed a paper for *Journal of Experimental Psychology: Applied*.
- Vice President, Usability Professionals Association-DC Chapter (UPA-DC).
- Program Chair, American Association of Public Opinion Research- DC Chapter (DC-AAPOR).

Laurie Schwede
- Reviewed abstracts for the 2010 AAPOR meeting.

Yves Thibaudeau
- Refereed a paper for *Journal of the American Statistical Association.*

Natalya Titova
- Refereed a paper for *Journal of Official Statistics.*

Lynn Weidmann
- Refereed a paper for *Journal of Official Statistics.*

Bill Winkler
- Refereed papers for *Transactions on Data Privacy*, *Annals of Applied Statistics*, *Statistical Data Protection 2010*, and *Statistical Papers, Survey Methodology,* and the *Journal of Official Statistics.*
- Associate Editor, *Journal of Privacy Technology*.
- Associate Editor, *Journal of Privacy and Confidentiality.*
- Associate Editor, *Transactions on Data Privacy.*
- Member, Program Committee, Statistical Data Protection 2010 in Corfu, Greece.
- Member, Program Committee, QDB 2010 at the 2010 Very Large Database Conference in Singapore.
- Member, Committee on Voter Registration Databases, National Academies of Science.

Tommy Wright
- Refereed papers for *The American Statistician, Public Opinion Quarterly,* and *Journal of Statistics Education.*
- Associate Editor, *The American Statistician.*
- Member, Fellows Committee, American Statistical Association.
- Member, Morris Hansen Lecture Committee, Washington Statistical Society.
- Member, Advisory Board, Mathematics and Statistics Department, Georgetown University.
- Member, Cochran-Hansen Prize Committee, International Association of Survey Statisticians.

Laura Zayatz
- Refereed a book on disclosure avoidance for Springer-Verlag.
- Organizer, Privacy in Statistical Databases 2010.
- Member, Confidentiality and Data Access Committee (CDAC).
- Member, NAS CTPP Project Panel for Producing Transportation Data Products from the American Community Survey that Comply with Disclosure Rules.
- Advisor, Disclosure Review Board, Social Security Administration.
- Member, Advisory Board, *Journal of Privacy Technology.*
- Member, Committee on Privacy and Confidentiality, American Statistical Association.
- Member, UK Census Design and Methodology Advisory Committee.
- Member, Advisory Board, *Journal of Empirical Research on Human Research Ethics*.

6.3 PERSONNEL NOTES

Pam Ferrari retired from the Census Bureau after 31 years of Federal service.

Thomas Mathew (member of the statistics faculty and Presidential Research Professor at the University of Maryland-Baltimore County) accepted a Schedule A Appointment in our Computing Applications Group.

Chandra Erdman joined our Missing Data Methods Research Group.

Victor Quach (graduate student in human factors psychology at The Catholic University of America) joined our Human Factors and Usability Research Group as an intern.

Jon Krosnick (Frederic O. Glover Professor in Humanities and Social Sciences as well as Professor of Communication and of Political Science at Stanford University) accepted a Schedule A Appointment in our Language and Measurement Research Group.

Matt Gore joined the Human Resources Division and is on a temporary assignment to our division through the Mixed-Tour Program.

Marisa Fond (Ph.D. student in sociolinguistics at Georgetown University) joined our Language and Measurement Research Group as an intern.

Amelia Tseng (Ph.D. student in sociolinguistics at Georgetown University) joined our Language and Measurement Research Group as an intern.

Jeff Moore retired from the Census Bureau after 39 years of Federal service.

Tina Arbogast retired from the Census Bureau after approximately 23 years of Federal service.

Michael DePersio (graduate student in mathematics and statistics at Georgetown University) joined our division as an intern.

Jiashen You (Ph.D. candidate in statistics at UCLA) joined our division as an intern.

Summer Visitors:
- Kathleen Denny (Ph.D. student in sociology at University of Maryland, College Park).
- Jenna Fulton (Ph.D. student in survey methodology at University of Maryland, College Park).
- Rebecca Medway (Ph.D. student in survey methodology at University of Maryland, College Park).
- Andrea Schwanz (senior in psychology at The College of William and Mary).
- David Vannette (graduate student in survey methodology at the University of Michigan).
- Joint Program in Survey Methodology Junior Fellows
  - Julia Goldstein (junior in mathematics and government at Hamilton College).
  - Joseph Mauro (junior in mathematics and economics at Fairfield University).
  - Kaitlin Woo (senior in mathematics at Georgetown University).

Joo Chung (Ph.D. student in sociolinguistics at Georgetown University) joined our Language and Measurement Research Group as an intern.

Christina Cooper (sophomore student in mathematics at Christopher Newport University) joined our division as an intern.

Sabin Lakhe (senior student in information systems at University of Maryland Baltimore County) joined our division as an intern.

Jason Lucero accepted a position in private industry.

Matt Clifton started graduate study at the University of Texas, Austin.

Aref Dajani accepted a position in the Demographic Statistical Methods Division.

Jennifer Beck accepted a position in the Economic Directorate.

Stephanie Sheffield accepted a university position.

Sarah Wilson (graduate student in linguistics at Georgetown University) joined our division as an intern.

Katherine Drom (senior student in sociology at George Washington University) joined our division as an intern.

APPENDIX A	Statistical Research Division's FY 2010 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		
5210001	Forms Development		
	1. Census Questionnaire Design Features (Other Than Race & Ethnicity)	Jenny Childs.............................Keith Woodling	
	2. Development of Race and Ethnicity Questions.............................	Rodney Terry .... Nicholas Jones	
5210003	*3. Language Planning and Development*	Patti Goerman ....Keith Woodling	
5310001	Data Collection Planning and Development		
	4. Accessible Web Surveys Research....................................	Larry Malakhoff........................ Jean Fox (BLS)	
	5. Decennial Reinterview Internet Testing – Usability Input	Kathy Ashenfelter ....Joan Hill	
5610002	Statistical Design and Estimation		
	6. Decennial Record Linkage	William Winkler .... Maureen Lynch	
	7. Decennial Disclosure Avoidance....................................	Laura Zayatz ....Marie Pees	
	8. Census Unduplication Research.....................................	Michael Ikeda.......................... Maureen Lynch	
5610003	Coverage Measurement Planning and Development		
	9. Coverage Measurement Research....................................	Don Malec................................... Tom Mule	
	10. Accuracy of Coverage Measurement....................................	Mary Mulry................................Pat Cantwell	
	11. Questionnaire Wording and Automation Team...........................	Beth Nichols.............................. Magda Ramos	
5610005/ 5610006	Coverage Improvement Planning and Development/ Evaluation Planning Coordination		
	12. Development of Questionnaires for Decennial Coverage Improvement..	Jenny Childs........................... Elizabeth Poehler	
	13. 2010 CPEX Experimental Overcount Booklet..............................	Laurie Schwede...........................Sarah Heimel	
	14. Evaluations, Experiments, and Assessments Operational Integration Team (EEA OIT).......................................	Laurie Schwede........................ Karen Medina	
	15. Evaluation of CCM Interviews	Beth Nichols.............................. Magda Ramos	
	16. Investigation of Study Methods for the Census Coverage Measurement (CCM) on Group Quarters (GQ) Population	Anna Chan .... Magda Ramos	
	17. 2010 Census Language Study (CPEX)..	Yuling Pan ....Keith Woodling	
	18. 2010 Census Behavior Coding Evaluation..................................	Jenny Childs.................................. Nancy Bates	
	19. Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups..............................	Laurie Schwede......................... Karen Medina	
	20. Explaining How Census Errors Occur through Comparing Census Operations History with Census Coverage Measurement (CCM) Results	Michael Ikeda...........................Tamara Adams	
	21. 2011 Relationship Survey..	Terry DeMaio.............................Bob Kominski	
5385060	American Community Survey (ACS)		
	22. ACS Group Quarters Item Imputation and Micro Data Disclosure Avoidance Research	Laura Zayatz ....Mark Asiala	
	23. ACS Small Area Estimation for Group Quarters (GQ)................	Lynn Weidman........................Alfredo Navarro	
	24. ACS Data Issues..	Lynn Weidman........................Alfredo Navarro	
5385095	American Community Survey (ACS) / Methods Panel		
	25. ACS Language Research..................................	Yuling Pan ....Todd Hughes	
	26. ACS Data Reliability Indicator Project..	Kathy Ashenfelter.........................Tony Tersine	
	27. ACS Messaging Project....................................	Laurie Schwede....................Herman Alvarado	
	28. ACS Internet Testing –Usability Input...............................	Kathy Ashenfelter ....Debbie Klein	
	29. ACS Internet Testing –Cognitive Input...........................	Jenny Childs......................... Jennifer Tancreto	
	30. ACS Internet Test Experimental Design Team............................	Kathy Ashenfelter ....Mary Frances Zelenak	
	31. Iterative Testing of the American Community Survey Web Site	Erica Olmsted-Hawala .. Margaret Gill, Kai Wu	
	32. ACS Content Test Pilot CARI Behavior Coding............................	Patti Goerman ....Todd Hughes	

Project #	Project/Subproject Sponsor(s)	SRD Contact	Sponsor Contact
	DEMOGRAPHIC		
1443000	*33. Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) Tables*..	Aref Dajani .... Marcella Jones-Puthoff	
0906/7374	Demographic Surveys Division (DSD) Special Projects		
	34. Data Integration...	Ned Porter ....Marie Pees	
	35. Using Survey Paradata to Manage Surveys in the Field and Estimate Survey Error	Matt Jans ....Cheryl Landman	
	36. Usability Testing of the National Survey of College Graduates	Jennifer Romano.............................Tim Gilbert	
1465001	Quick Turnaround Pretesting of Household Surveys		
	37. National Crime Victimization Survey (NCVS)......................	Terry DeMaio ....Jeremy Shimer	
	38. Development of the CARI Behavior Coding System..................	Terry DeMaio ....Candice Barnes	
	39. American Housing Survey Modules	Terry DeMaio .... Joe Huesman	
7048001	*40. Survey of International Air Travelers*..	Terry DeMaio ....Chris Laskey	
1465444	Re-Engineered Survey of Income and Program Participation (RE-SIPP) Research		
	41. RE-SIPP Methodological Research...	Anna Chan ....David S. Johnson	
	42. Model-Based Imputation for the Demographic Directorate	Chandra Erdman......................... Martha Stinson	
7165000	Data Integration Division (DID) Small Area Estimation Projects		
	43. Research for Small Area Income and Poverty Estimates (SAIPE)	Elizabeth Huang .... Lucinda Dalzell	
	44. Small Area Health Insurance Estimates (SAHIE)	Don Malec .... Donald Luery	
	ECONOMIC		
2370054	*45. Editing Methods Development (Investigation of Selective Editing Procedures for Foreign Trade Programs)*	María García...................................Ryan Fescina	
2470051	*46. Disclosure Avoidance Methods*	Laura Zayatz....................................Rita Petroni	
2370052	Time Series Research		
	47. Seasonal Adjustment Support	Brian Monsell .... Kathleen McDonald-Johnson	
	48. Seasonal Adjustment Software Development and Evaluation ...	Brian Monsell .... Kathleen McDonald-Johnson	
	49. Research on Seasonal Time Series - Modeling and Adjustment Issues	Tucker McElroy ... Kathleen McDonald-Johnson	
	50. Supporting Documentation and Software for X-12-ARIMA and X-13A-S..	Brian Monsell .... Kathleen McDonald-Johnson	
TBA	*51. Survey of Research and Development in Industry, Imputation and Sampling Research and Software Design*.............................	Yves Thibaudeau ....Jeri Mulrow	
TBA	*52. Governments Division Project on Decision-Based Estimation* .	Eric Slud ....Yang Cheng	
	STRATEGIC PLANNING AND INNOVATION		
0359999	*53. Remote Access – Microdata Analysis System*...........................	Laura Zayatz............................. Nancy Gordon	
	STATISTICAL RESEARCH DIVISION		
Other	*54. Current Population Survey (CPS) Health Insurance Measurement Research*..	Joanne Pascale......................Charles T. Nelson	
	55. American FactFinder (AFF) Website Redesign	Erica Olmsted-Hawala.................Marian Brady	

**FY 2010 PROJECT PERFORMANCE
MEASUREMENT QUESTIONNAIRE
STATISTICAL RESEARCH DIVISION**

Dear

In a continuing effort to obtain and document feedback from program area sponsors of our projects or subprojects, the Statistical Research Division will attempt for the eleventh year to provide *seven measures of performance,* this time for the fiscal year 2010.   For FY 2010, the *measures of performance* for our division are:

Measure 1. Overall, Work Met Expectations: Percent of FY 2010 Program Sponsored Projects/Subprojects where sponsors reported that work met their expectations.

Measure 2. Established Major Deadlines Met: Percent of FY 2010 Program Sponsored Projects/Subprojects where sponsors reported that all established major deadlines were met.

Measure 3a. At Least One Improved Method, Developed Technique, Solution, or New Insight: Percent of FY 2010 Program Sponsored Projects/Subprojects reporting at least one improved method, developed technique, solution, or new insight.

Measure 3b. Plans for Implementation: Of the FY 2010 Program Sponsored Projects/Subprojects reporting at least one improved method, developed technique, solution, or new insight, the percent with plans for implementation.

Measure 4. Predict Cost Efficiencies: Number of FY 2010 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 2010.

Measure 6. Proceedings Publications: Number of proceedings publications documenting research that appeared in FY 2010.

These measures will be based on response to the five questions on this form from our sponsors as well as from members of our division and will be used to help improve our efforts.

To construct these seven measures for our division, we will combine the information for all of our program area sponsored projects or subprojects obtained during November 1 thru November 15, 2010 using this questionnaire. Your feedback is requested for:

Project Number and Name: _____
Sponsoring Division(s): _____

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

We very much appreciate your assistance in this undertaking.

Tommy Wright   Date
Chief, Statistical Research Division

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

*Brief Description of Results/Products from FY 2010 (**SRD Contact will provide**):*

(over)

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)**

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

  2. Listed below are at most 2 of the top improved methods, developed techniques, solutions, or new insights offered or applied on this project or subproject in FY 2010 where an SRD staff member was a significant contributor. Review "a" and "b" below **(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.

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

  □ Yes as listed below. (See a and b.)

                                          Plans for
                                          Implementation?
  a. _____   Yes □   No □
     _____
     _____
     _____
     _____

  b. _____   Yes □   No □
     _____
     _____
     _____
     _____

  Comments (Sponsor Contact):

COST:
  Predict Cost Efficiencies

  3. Listed **(provided by SRD Contact)** below are at most two research results or products produced for this project or subproject in FY 2010 that predict cost efficiencies. Review the list, and make any additions or deletions as necessary. Add any comments.

      □ No cost efficiencies predicted.
      □ Yes as listed below. (See a and b.)

   a.

   b.

  Comments (Sponsor Contact):

OVERALL:
  Expectations Met/Improving Future Communications

  4. Overall, work on this project or subproject by SRD staff during FY 2010 met expectations. **(Sponsor Contact)**

      □ Strongly Agree
      □ Agree
      □ Disagree
      □ Strongly Disagree

  5. Please provide suggestions for future improved communications or any area needing attention on this project or subproject. **(Sponsor Contact)**

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

First_____
      Sponsor Contact Signature   Date

Second_____
      SRD Contact Signature   Date

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

Third_____
      Sponsor Division Chief Signature   Date

Fourth_____
      SRD Division Chief Signature   Date

Statistical Research Division

Assistant Division Chief for
Computing & Technology
Robert Creecy
  Chad Russell
  Matthew Gore (HRD)
  VACANT

Machine Learning & Computational Statistics Research
Bill Winkler
  William Yancey

Computing Applications
VACANT
  Thomas Mathew (UMBC)
  Tom Petkunas
  Ned Porter

Missing Data Methods Research
Yves Thibaudeau
  Chandra Erdman
  Maria Garcia
  Martin Klein
  Ben Klemens
  Rolando Rodriguez
  Jun Shao (U. of WI)
  VACANT

Assistant Division Chief for
Mathematical Statistics
VACANT
  Gloria Prout

Sampling & Estimation Research
Lynn Weidman
  Mike Ikeda
  Patrick Joyce
  Mary Mulry
  Chaitra Nagaraja
  Eric Slud (U. of MD)
  Julie Tsay
  VACANT

Small Area Estimation Research
Don Malec
  Aaron Gilary
  Elizabeth Huang
  Ryan Janicki
  Partha Lahiri (F)
  Jerry Maples
  Joe Sedransk

Disclosure Avoidance Research
Laura Zayatz
  Michael DePersio (S)
  Michael Freiman
  Marlow Lemons (S)
  Paul Massell
  Tapan Nayak (GWU)
  Asoka Ramanayake
  Lisa Singh (Georgetown U)
  Bimal Sinha (UMBC)

Time Series Research
Brian Monsell
  Chris Blakely (Postdoc)
  David Findley
  Tucker McElroy
  Ekaterina Sotiris (S)
  Natalya Titova
  VACANT

Assistant Division Chief for
Survey Methodology
VACANT
  VACANT

Language & Measurement Research
Yuling Pan
  Anna Chan
  Jenny Hunter Childs
  Matt Clifton
  Patti Goerman
  George Higbie
  Nathan Jurgenson (S)
  Jon Krosnick (Stanford U)
  Stephen Lubkemann (GWU)
  Laurie Schwede
  Anissa Sorokin (S)
  Rodney Terry (Postdoc)
  Virginia Wake (Postdoc)
  VACANT

Questionnaire Pretesting for Household Surveys
Terry DeMaio
  Katherine Drom (S)
  Rachel Freidus
  Joanne Pascale
  Lorraine Randall
  VACANT

Human Factors & Usability Research
Kathleen Ashenfelter
  Temika Holland
  Matt Jans
  Sabin Lakhe (S)
  Larry Malakhoff
  Beth Nichols
  Erica Olmsted-Hawala
  Victor Quach (S)
  Jennifer Romano (Postdoc)
  VACANT

Tommy Wright, Chief
  Kelly Taylor
  Ann Dimler
  Michael Hawkins
  Christina Cooper (S)
  Sarah Wilson (S)

(F) ASA/NSF/Census Research Fellow
(S) Student