STATISTICAL RESEARCH DIVISION
FY 2010 THIRD QUARTER REPORT
–April, May, June 2010–

COLLABORATION

DECENNIAL DIRECTORATE

Decennial Management Division/Decennial Statistical Studies Division/American Community Survey Office (Sponsors)

Project Number	Project Title	FTEs
5210001	Forms Development	.70
A.	*Census Questionnaire Design Features (Other than Race and Ethnicity)*	
B.	*Development of Race and Ethnicity Questions*	
5210003	Language Planning and Development	1.00
5310001	Data Collection Planning and Development	1.00
A.	*Accessible Web Surveys (Research)*	
B.	*Decennial Reinterview Internet Testing - Usability Input*	
5310008	Special Place/Group Quarters (GQ) Planning and Development	.37
5610002	Statistical Design and Estimation	6.67
A.	*Decennial Editing and Imputation*	
B.	*Decennial Record Linkage*	
C.	*Decennial Disclosure Avoidance*	
D.	*Census Unduplication Research*	
E.	*Statistical Design for Experiments and Evaluations*	
5610003	Coverage Measurement Planning and Development	1.39
A.	*Coverage Measurement Research*	
B.	*Accuracy of Coverage Measurement*	
C.	*Questionnaire Wording and Automation Team*	
5610005	Coverage Improvement Planning and Development	1.00
5610006	Evaluation Planning Coordination	3.50
A.	*Development of Questionnaires for Decennial Coverage Improvement*	
B.	*2010 CPEX Experimental Overcount Booklet*	
C.	*Evaluations, Experiments, and Assessments Operational Integration Team (EEA OIT)*	
D.	*Evaluation of CCM Interviews*	
E.	*Investigation of Study Methods for the Census Coverage Measurement (CCM) on Group Quarters (GQ) Population*	
F.	*2010 Census Language Study (CPEX)*	
G.	*2010 Census Behavior Coding Evaluation*	
H.	*Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups*	
I.	*Explaining How Census Errors Occur through Comparing Census Operations History with Census Coverage Measurement Results*	
J.	*2011 Relationship Survey*	
5385060	American Community Survey (ACS)	2.23
A.	*ACS Missing Data and Imputation*	
B.	*ACS Group Quarters Item Imputation and Micro Data Disclosure Avoidance Research*	
C.	*ACS Applications for Time Series Methods*	
D.	*ACS Small Area Estimation for Selected Characteristics*	
E.	*ACS Small Area Estimation for Group Quarters*	
F.	*ACS Data Issues*	

5385095 American Community Survey (ACS)/Methods Panel 3.32
- *A. ACS Language Research*
- *B. ACS Data Reliability Indicator Project*
- *C. ACS Messaging Project*
- *D. ACS Internet Testing – Usability Input*
- *E. ACS Internet Testing – Cognitive Input*
- *F. ACS Internet Test Experimental Design Team*
- *G. ACS Iterative Testing of the Web Site*
- *H. ACS Content Test Pilot CARI Behavior Coding*

DEMOGRAPHIC DIRECTORATE

Population Division (Sponsor)

Project Number	Project Title	FTEs
TBA	Current Population Survey (CPS)/Annual Social and Economic Supplement (ASEC) Tables.....	TBA

Demographic Surveys Division (Sponsor)

Project Number	Project Title	FTEs
0906/7374	Demographic Surveys Division Special Projects	1.16
	A. Data Integration	
	B. Using Survey Paradata to Manage Surveys in the Field and Estimate Survey Error	
	C. Usability Testing of the National Survey of College Graduates	
1465001	Quick Turnaround Pretesting of Household Surveys	1.06
	A. Rental Housing Finance Survey	
	B. National Crime Victimization Survey (NCVS)	
	C. Development of the CARI Behavior Coding System	
	D. American Housing Survey Modules	

Housing and Household Economic Statistics Division (Sponsor)

Project Number	Project Title	FTEs
1465444	Re-Engineered Survey of Income and Program Participation Research	1.75
	A. Re-Engineered Survey of Income and Program Participation Methodological Research	
	B. Model-based Imputation for the Demographic Directorate	

Data Integration Division (Sponsor)

Project Number	Project Title	FTEs
7165000	Data Integration Division Small Area Estimation Projects	2.17
	A. Research for Small Area Income and Poverty Estimates (SAIPE)	
	B. Small Area Health Insurance Estimates (SAHIE)	

ECONOMIC DIRECTORATE

Project Number	Project Title	FTEs
2370054	Editing Methods Development	.18
	Investigation of Selective Editing Procedures for Foreign Trade Programs	
2470051	Disclosure Avoidance Methods	1.14

2370052 Time Series Research 2.11
- *A. Seasonal Adjustment Support*
- *B. Seasonal Adjustment Software Development and Evaluation*
- *C. Research on Seasonal Time Series - Modeling and Adjustment Issues*
- *D. Supporting Documentation and Software for X-12-ARIMA and X-13A-S*

TBA Survey of Research and Development in Industry, Imputation and Sampling Research and Software Design ..... TBA

TBA Governments Division Project on Decision-Based Estimation ..... TBA

STRATEGIC PLANNING AND INNOVATION

Project Number	Project Title	FTEs
0359999	Remote Access - Microdata Analysis System	1.72

CENSUS BUREAU

Project Number	Project Title	FTEs
0381000	Program Division Overhead	12.25

- *A. Division Leadership and Support*
- *B. Research Computing*

GENERAL RESEARCH AND SUPPORT

Project Number	Project Title	FTEs
0351000	General Research and Support	27.26
1871000	General Research	3.26

STATISTICAL METHODOLOGY
- *A. Disclosure Avoidance*
- *B. Disclosure Avoidance for Microdata*
- *C. Seasonal Adjustment*
- *D. Household Survey Design and Estimation*
- *E. Sampling and Estimation Methodology: Economic Surveys*
- *F. Research and Development Contracts*
- *G. Small Area Estimation*

STATISTICAL COMPUTING METHODOLOGY
- *A. Record Linkage and Analytic Uses of Administrative Lists*
- *B.1. Editing*
- *B.2. Editing and Imputation*
- *C. Developed Software Support – General Variance Estimation Development and Support*
- *D. Missing Data and Imputation: Multiple Imputation Feasibility Study*
- *E. Modeling, Analysis and Quality of Data*

SOCIAL & BEHAVIORAL SCIENCES SURVEY METHODOLOGY
- *A. Usability Research and Testing*
- *B. Questionnaire Pretesting*
- *C. Questionnaire Design Experimental Research Survey 2006 (QDERS)*
- *D. Language: Interdisciplinary Research on Language and Sociolinguistic Issues Relevant to Survey Methodology*
- *E. Training for Cognitive Interviewing*
- *F. Research on Cognitive Testing of Non-English Language Survey Instruments*
- *G. Interviewer-Respondent Interactions*
- *H. Q-Bank: A Database of Pretested Questions*
- *I. Health Insurance Measurement*

J. Emerging Social Trends on Household Structure and Living Situations, Race/Ethnicity, and Linkages to Enumeration Method and Coverage
K. Using Vignettes to Explore Survey Concepts
L. Retrieval Effects on Judgments about Knowledge

RESEARCH SUPPORT AND ASSISTANCE

PUBLICATIONS
- Journal Articles, Publications
- Books/Book Chapters
- Proceedings Papers
- Statistical Research Division Research Reports
- Statistical Research Division Studies
- Other Reports

TALKS AND PRESENTATIONS

STATISTICAL RESEARCH DIVISION SEMINAR SERIES

PERSONNEL ITEMS
- Honors/Awards/Special Recognition
- Significant Service to Profession
- Personnel Notes

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.

Highlights: No significant progress this quarter.

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

B. Development of Race and Ethnicity Questions
Description: 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.

Highlights: During the third quarter of FY2010, staff continued work 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 also developed strategies for recruiting Spanish monolinguals of various nationalities, races, and ethnicities. Staff are currently recruiting respondents and conducting cognitive interviews.

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 this quarter, staff participated in a series of Lessons Learned meetings regarding the 2010 Decennial Census language program. Staff provided verbal and written comments to the Decennial Management Division.

Future Plans: Staff will continue to support the Decennial area through participation in team meetings, review of non-English forms and materials, and language research.

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: No significant progress this quarter.

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

B. Decennial Reinterview Internet Testing – Usability Input
Description: The usability team is working with the cognitive team on the development of an online instrument for the Decennial Reinterview Project. The project will consist of two rounds of usability and cognitive testing of prototypes of the internet instrument and the letters associated with the paper forms. 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: The second round of usability testing on this Web survey has been completed, and the report is currently being drafted.

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 Editing and Imputation

[See Projects 0351000 and 1871000 (B), General Research - Statistical Methodology]

B. Decennial Record Linkage

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

C. 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: 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 worked to improve the current data swapping algorithm for the ACS. They wrote an algorithm for improving pairing efficiency, developed an algorithm for cyclical swapping based on the "Traveling Salesman Problem," and developed a method of imputing geographies. All three new methods will be researched and will be compared to the current swapping method.

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. Laura participated in a webinar for this project.

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, Public Use Microdata Areas (PUMAs), and state. Staff implemented a test-run exercise using the states of 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.

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

D. 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 Census. One of the major problems is how to incorporate the effects of name frequency into the unduplication procedures. Staff also provides assistance in specifying and reviewing output from the matching and unduplication procedures for test censuses and eventually for Census 2010. We began this project in May 2004.

Highlights: Staff attended multiple sessions to review data and set cutoffs for within-response (multiple links between a pair of housing units) links from the 2010 Census Duplicate Person Identification process.

Staff: Michael Ikeda (x31756), Ned Porter

E. Statistical Design for Experiments and Evaluations

Description: The overall objective of this project is to provide statistical expertise in experimental design for decennial-related experiments and evaluations.

Highlights: No significant progress this quarter.

Staff: Thomas Mathew (x35337), Aref Dajani

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).

Nonignorable Models for Coverage Component Estimation
Highlights: Staff completed a draft of a manuscript which documents the construction of a new nonignorable model and compares this model to a well-known alternative using a combination of data analysis and simulation studies.

Developing a Model for Census Coverage Using Race-based Partitioning
Highlights: A staff member supplemented his previous work on a race-based model by incorporating Census domains and assessing the validity of the model for different racial subpopulations by using measures of racial consistency between surveys. He presented his results to Decennial Statistical Studies Division (DSSD) stakeholders in the Census Coverage Estimation Team on April 29, and to the component missing data subgroup several times. He also met with staff from DSSD to discuss conclusions and coordinate future research on the project, and wrote the first draft of a research report for this project.

Aggregate Level Small Area Estimation Modeling
Highlights: Staff are looking at new ways to estimate the probability of correct enumeration in the census using small area techniques. They would like to incorporate random effects into the model and to use Bayesian methods to estimate the parameters of interest. However, there is likely clustering in the sample, and we do not know the form of the probability density for the sample. They worked to develop a method which modifies the quasi-likelihood of a sample to more accurately reflect the sampling design. Their goal is to construct a likelihood that is "close" to the quasi-likelihood and which has a mode and curvature that match up with the design-based parameter estimate and variance estimate. Staff considered modifying the design weights so that the new weights give us a new quasi-likelihood with the desired properties. While this idea has intuitive appeal and is computationally feasible, preliminary studies show that many of the new weights can be set to zero, and there is the possibility that the new weights can "blow up." There are also technical issues which must be addressed due to the fact that the new weights are functions of the data.

Staff are considering alternative methods to modifying the design weights. One idea is to separately model the probability of clustering using the empirical distribution. A second approach, which was suggested by a participant in the *SUMMER AT CENSUS* program, is to use semiparametric methods to construct estimating functions for the parameter. Staff plans to further investigate all three methods.

Unit Level Small Area Estimation Modeling
Highlights: Staff completed the evaluation of housing unit and block-level random effects on coverage using randomization tests with the 2006 Travis county test data.

Logistic Mixed Model Estimation
Highlights: Staff are preparing a report and presentation on the project.

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. The first attempt to implement the new methodology is with data from the 2006 Census Test.

Highlights: Staff provided technical expertise and experience in the planning and implementation of coverage measurement research for the 2010 Census. This included serving on three teams formed to plan and implement census coverage measurement (CCM) research for the 2010 Census.

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 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. In addition, work began on a draft of the study plan for the CPEX study "Developing an Error Structure in Components of Census Coverage Error."

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: Staff commented on the first draft of the 2010 Census Coverage Measurement (CCM) Person Followup (PFU) interviewer training.

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: We will consult on the development of questions and questionnaires designed to improve within household coverage in the Decennial Census. We will 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.

Highlights: Staff assisted the Decennial Statistical Studies Division in planning a large-scale cognitive test and qualitative study of the Targeted CFU (TCFU) and duplicates in the 2010 Census.

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 persons' alternative addresses and where to count them. 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: Staff sent the revised final report to the sponsor for approval.

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: Staff attended meetings and responded to requests by sponsors. 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 CPEX study plans submitted by division colleagues.

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: Staff selected sites and observers, and prepared training for the observers participating in the respondent debriefing project.

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: Staff contacted two 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 Site Two in May to deliver the remaining 40 questionnaires. Staff completed data collection and implemented non-response followup procedures at both sites via email and telephone. 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 twice via telephone conference in April and June and began reviewing American Community Survey observation reports ethnography field note reports and decennial enumeration observation reports.

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: The CPEX study "Observing census enumeration of non-English-speaking households in the 2010 Census" aims 1) to identify language and cultural barriers in census interviews with respondents who speak little or no English, and 2) to examine the effectiveness of the 2010 Census Language Assistance Program. During this quarter, staff developed the theoretical framework and research methodology for the fieldwork and conducted two-day training for the contracted ethnographers in seven languages (Spanish, Chinese, Korea, Russian, Vietnamese, Arabic, and Portuguese). A total of 23 bilingual ethnographers participated in the training and translated study documents from English into the perspective target languages. Staff provided technical guidance on the translation of the observation protocol and the respondent and enumerator debriefing questions. Staff also coordinated with the Field Division to make arrangements for the ethnographers to observe NRFU interviews in the target languages to ensure successful fieldwork. Ethnographers in all seven languages observed NRFU interviews for two weeks in the designated areas of high concentration of speakers of the target languages. They also conducted debriefing 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 summarizing field notes and preliminary findings.

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 interviewer-administered instruments (e.g., NRFU, CFU, CCM) 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 all 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: This quarter, staff prepared training materials for the Nonresponse Followup (NRFU) behavior coding.

Staff: Jennifer Hunter Childs (x34927), Nathan Jurgenson

H. Comparative Ethnographic Studies of Enumeration Methods and Coverage in Race/Ethnic Groups

Description: Staff will conduct comparative ethnographic research on enumeration methods and coverage in four to 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. This field study will involve accompanying enumerators to observe, tape, and debrief respondents during three 2010 operations involving personal visit census data collection: Update/Enumerate (U/E), Nonresponse Followup (NRFU), and Census Coverage Measurement (CCM). 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. An additional component to explore factors respondents use in self-identification of race is under consideration.

Highlights: The Update/Enumerate contract ethnographer completed nine days of observation/debriefing on the Indian reservation and provided very useful comments on how to improve the project methods, taping instructions, and checklists. The ethnographer observed 35 interviews but noted that about half of his respondents declined to be taped. Reluctance to taping is not surprising, as many Indian reservations have explicit policies that prohibit taping. A staff member then went to that site herself to conduct several days of observations and debriefings to identify potential weaknesses and flaws, to get a feel for how the U/E Operation was working, and to increase our number of taped interviews from that site. She observed for three days in different areas and got first-hand experience in how difficult it is to locate widely scattered housing units in a very remote area with no GPS and with no identifying features on the census maps. She had a long and enlightening discussion with the Field Office Supervisor. She also took a long solo tour with an Indian anthropologist on Sunday to learn more about the culture and appropriate behavior for outside observers, and to ask him questions about socio-economic conditions on the reservation. On return to the office, she revised the methods, handout and checklists accordingly, then revised them to be consistent with the NRFU operation.

We worked very closely with staff from Acquisition Division (ACQ) and the Office of Security to get our seven Phase 2 NRFU contract ethnographers' contracts and security clearances approved by early May. We also worked very closely with staff from Field Division (FLD) to get Regional Census Center approvals for our seven field sites and with FLD staff to get Field Visit Request approvals. Due to complications in one Regional Census Center area, one field site had to be changed on very short notice.

We conducted two two-day training sessions with our NRFU contractor ethnographers. At the request of Field Division, we waited an additional week or so to begin observations in some sites, due to NRFU processing problems. All NRFU ethnographers were in their field sites for nine continuous days between May 11 and May 27, accompanying interviewers and observing/debriefing during live census interviews at respondents' homes. Our Phase 1 (U/E) and Phase 2 (NRFU) contractors are now sending their tapes, transcriptions, and interview summaries on a flow basis. Staff is processing and duplicating the tapes and will give one set to a division colleague for her behavior coding project. The Chinese tapes are being duplicated for possible behavior coding.

Staff also selected eight additional ethnographers to conduct the Phase 3 observations/debriefings during the CCM/PI Operation in August and September. We revised the core statement of work and sole source justification documents to be consistent with the CCM PI Operation and had them reviewed by an IT security lawyer. We consulted with staff from the Privacy Office and ACQ to learn how to complete new privacy and acquisitions documents. We prepared and submitted the eight contracting packages and worked with Administrative staff and staff from ACQ to process those. Security clearance documents for our contractors are currently being processed. We are working with staff from LAN Technology Support Office (LTSO) to get special loaner laptops configured for these new contractors before they arrive.

We are collaborating with staff from Decennial Statistical Studies Division (DSSD) and Field Division (FLD) to get approvals from eight Regional Census Centers to conduct our CCM observation studies in their areas. We collaborated with division colleague in preparing a PowerPoint overview of our similar evaluations for DSSD and FLD to present at the CCM Manager's conference in June. We prepared an overview of our CCM observation project and revised our interviewer handout, then sent these materials to be sent to the nine Regional Census Centers.

Staff: Laurie Schwede (x32611), Rodney Terry, Matthew Clifton, Greg Bulmash, Lorraine Randall, 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. We want to merge Census files from the various stages of Census operations for a subsample of CCM areas and compare them to the CCM results. This comparison is intended to help find patterns of errors in Census operations and provide insights into ways to avoid these errors.

Highlights: Staff sent out a revised draft study plan and a revised draft file listing to critical reviewers and some additional interested parties. Staff also sent the revised draft study plan to the Decennial Management Division in response to a request from the General Accounting Office for study plans for several evaluations.

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: Staff completed the contract with Westat to conduct focus groups with same sex and opposite sex couples to investigate potential changes to the relationship and marital status questions to accommodate same sex marriage. With staff from Westat, staff co-authored and presented a paper summarizing the results of focus groups. The paper, presented at the Annual Conference of the American Association for Public Opinion Research (AAPOR), was entitled "Measurement of Relationship and Marital Status: Results of Focus Groups with Same-Sex Couples."

Staff attended an all-day meeting with the Census Bureau's Relationship Study Expert Panel and presented the results of the focus groups to this audience. We also presented the results at a Survey Methodology Brownbag.

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: No significant progress this quarter.

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: Staff generated synthetic data for disclosure avoidance for the 2009 ACS group quarters sample. Synthesis was performed via a prototype version of an SRD-written R package named "pep."

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: Staff completed first draft of project for a client, to conduct a study comparing ACS Multi-Year Estimates with linear combinations of one-year Estimates. After receiving criticism on methodology, we devised a new nonparametric methodology and presented results to the client, who was satisfied. Future plans include presentation to the ACS Research and Evaluation Committee and publishing the paper.

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

D. 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: No significant progress this quarter.

Staff: Lynn Weidman (x34902)

E. 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: 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 (SEs), 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 and continue to be examined and summary reports are being drafted. One notable result is that the SEs of all the imputation methods are smaller than those of the data-only estimates. This is at least partly due to constraints on estimates of total GQ persons at the tract and county level that are used for the imputed data sets, but not for the sample only estimates.

Staff is preparing for evaluations with ACS data files that will include more variables than just demographics. Data from years 2006-09 will be used along with a derived 2010 sample data file. Meetings were held to discuss and documentation drafted to describe files, records, and variables that will be used in the imputation and weighting. Weighting programs are being modified to include a person non-interview adjustment, which was not used in the simulation.

Methods were proposed and discussed to approximately estimate the effect of imputation on variances of GQ estimates. One of these methods will be used for now until the more theoretically-correct fractional replication procedure is developed and tested.

Staff has been developing possible methods for approximating the sizes of GQs that are not in sample during each year. These are needed to determine the number of people to be imputed into a GQ and to establish the tract and county GQ population constraints. For the year 2010, a relationship between maximum size from the Census 2010 GQ Enumeration and actual size is being 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, is being developed to apply to non-sampled GQs to adjust their expected populations.

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

F. 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: No significant progress this quarter.

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 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 the third quarter of FY10, staff worked closely with the ACS language team to carry out four language projects: 1) Translation review and cognitive testing of translated ACS Language Assistance Guides in Chinese and Korean; 2) cognitive testing of the remaining half of the Spanish CATI/CAPI instrument (Project D, Part III and IV); 3) cognitive testing of the ACS Regional Office (RO) letters in five languages (Project C); and 4) Spanish translation of the ACS internet instrument. Staff worked as Census Bureau research analysts, providing technical guidance and supervision for the contractor. Staff reviewed and provided guidance on cognitive interview protocols, interview summaries, and alternative translations. Staff also translated and reviewed the Spanish version of the ACS internet instrument.

Staff developed a systematic coding scheme to code types of translation issues identified in the cognitive testing process and provided training to the Contractor's cognitive interviewers. 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 will be implemented in the ACS Spanish CATI/CAPI instrument project.

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.

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: The report for the third round of usability testing has been submitted to the ACS critical review board for review.

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: Staff 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."

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: The second round of usability testing for the Web survey has been completed and a report is being drafted. Staff regularly attend design meetings.

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

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: Staff continued working on the Computer Assisted Telephone Interview (CATI) Reinterview survey called the Attitudes and Behavior Study (ABS). Questions in this survey will ask about Internet use/nonuse, privacy concerns, and effectiveness of mailing materials. These questions will vary slightly across 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 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: Cognitive and usability testing was performed on the mailing materials for this project, and a report is being drafted.

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

G. Iterative Testing of the American Community Survey Web Site
Description: The American Community Survey (ACS) Web site is making major re-designs to their 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: Staff worked on a quick report for the iteration 3.0 usability testing. Findings highlight that the main page was the dominant source for information and participants did not efficiently use the top navigation. Most participants believed that the ACS form could be filled out online, which is not currently true. Participants said that information about penalties should be grouped together with information about whether the survey was mandatory. Staff gave the quick report to the client and is working on the final report.

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

H. ACS Content Test Pilot CARI Behavior Coding
Description: Staff are 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. This project will involve 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: Staff participated in the development and testing of the software that will be used for this project. In addition, staff participated in bi-weekly ACS Content Test team meetings to assist in project planning and had several meetings with project sponsors to discuss a detailed project schedule.

Future Plans: Staff will participate in systems testing, and when the system is completed, work on the pilot project will continue. Staff will create behavior codes and configure the system, listen to sample recordings, write a training manual, travel to one of the Census Bureau telephone centers to train coders, monitor coder performance, complete data analysis and write up a report.

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: This project is on track to be completed in FY2010 Q4. 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: Staff developed algorithms based on the matcher to identify potential targets for re-identification. More algorithms were developed to create profiled targets by relations to other people, such as families and associates. Those groups were profiled so that attacks on the ACS Public Use Microdata Samples (PUMS) file will commence.

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: Staff continued their work on Current Population Survey paradata, and began additional work on National Health Interview Survey paradata and National Crime Victimization Survey paradata.

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.

Highlights: No significant progress this quarter.

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: Staff completed an expert review of the 2011 School Crime Supplement questionnaire and recommended changes to the cyberbullying questions. The changes were accepted by the sponsor. Staff recruited respondents for cognitive interviews with the School Crime Supplement questionnaire, obtained Office of Management and Budget clearance, and made preparations to begin interviewing.

Staff continued work on a literature review documenting issues related to the role of memory and perceptual processes in proxy racial identification. This literature review is part of a project to revise the response categories for the race-of-offender question in the NCVS.

Staff: Terry DeMaio (x34894), Jennifer Beck

C. Development of the CARI Behavior Coding System
Description: This project involves consultation with staff from DSD, DSMD, DSSD, TMO, FLD, ACSO 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 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: Staff attended biweekly meetings of the CARI project and provided a behavior coding orientation to the interviewers in the TTC who tested the Behavior Coding Component (BCC) of the CARI System. They participated in a number of tests of the BCC: Systems test, Verification Test, End-to-end Systems Test, and Load Test. They provided comments on the user feedback form. They also developed a list of enhancements for the BCC system, secured division funding, and wrote up portions of a Statement of Work for additional development of the CARI System.

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 scheduled for implementation in 2011.

Highlights: Staff participated in expert reviews of the questionnaire, which was finalized for cognitive testing after consultation with staff from the Decennial Statistical Division and the Department of Housing and Urban Development, the sponsor.

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

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: The third quarter of FY2010 coincided with the retirement of Jeff Moore, our long time group leader and the chair of the Re-engineered SIPP Research Group since 2007. Staff from Housing and Household Economic Statistics Division is now leading the research group. Staff from our division continued to contribute to the research project and participate in the weekly research group meeting. As the data collection for the long-planned 2010 SIPP EHC-CAPI field test was winding down by mid-March, the Re-Engineered SIPP Research Group began planning for the 2011 EHC Wave 1 and 2012 EHC Wave 2 Field tests during the third quarter of FY2010. A new sample of approximately 4,000 households in high poverty stratum will be selected and Field Representatives in all twelve regions will participate in the 2011 Field test. The 2011 research objectives are to evaluate and improve interviewers training, incorporate as many of the identified fundamental changes for evaluation, and provide a foundation for a wave 2 instrument test in 2012. The design of the 2011 instrument will provide a foundation for a study of attrition and mover location and for evaluation of the incorporation of feedback into the 2012 instrument. The primary goals for the 2012 test would be (1) to understand and minimize attrition or mover loss and to develop and evaluate procedures for maintaining contact with sample between waves that will be one year rather than four months apart for current production SIPP; and (2) to provide data to create initial evaluations of the nature of seam bias in the re-engineered SIPP. Staff are collaborating in the design of the recontact experiment to track respondents. Staff continue to lead the evaluation effort on the EHC recordings with the goal to understand what features of EHC methods are effective at assisting respondents' recall of past events and circumstances, and how those methods work. During this quarter, staff worked closely with staff from the Tucson call center to transcribe the 196 EHC recordings. Some of the recordings had gone through a 'second pass' by another transcriber to enhance the transcripts quality and all transcription work completed by mid June. Starting in May, staff, along with three summer interns from the Joint Program in Survey Methodology, began reviewing the transcripts and developing coding schemes to conduct behavioral coding of the transcribed interviews.

Staff: Anna Chan (x38462), Joanne Pascale

B. Model-Based Imputation for the Demographic Directorate
Description: Staff has been asked to review and improve ultimately all of the imputation methodology in demographic surveys, beginning with the Survey of Income and Program Participation and the Current Population Survey.

Highlights: A flexible multiple imputation R function for filling gaps in longitudinal data has been implemented. Currently, the function performs a randomized multiple imputation hot-deck, selecting 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.

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: Staff continues to develop models for the county poverty rate and its design-based variance based on the American Community Survey. The variance models previously developed for the log-level county models had to be modified to incorporate variance features associated with rates, namely a p(1-p), or pq, structure. Another development is that the generalized variance function (GVF) has been factored into two intuitive terms. The first represents the pq structure based on the true poverty rate. The second corresponds to

effective sample size, or equivalently, to design effect. The "true poverty rate" is currently being estimated by the information contained in the design-based variance estimates. Staff is concerned that this could lead to inconsistent estimates of the true poverty rate between the variance and poverty rate models. This issue will be investigated further and may warrant coupling the poverty rate and variance models.

Staff has also uncovered empirical evidence of a relationship between the precision of the variance of the poverty rate and the true poverty rate. The variance model framework has been augmented to allow the degrees of freedom (i.e., precision) to be a function of the poverty rate.

Staff is working on a draft paper titled "Further simulation results on the distribution of some survey variance estimators," to be presented at 2010 JSM in August. In this paper, the simulation results are documented on some distribution properties of four variance estimators of mean or ratio (proportion) under different populations (Normal (0, 1), Bernoulli, Poisson, ACS 2005 data) from random samples of different sample sizes. The four variance estimators are the usual variance estimator of mean or the Taylor series method of ratio, Fay's successive difference replication variance estimator, the jackknife (delete one unit) variance estimator, and the random group (10 groups) variance estimator.

For Normal (0, 1) population, all four variance estimators of sample mean can be approximated by the chi-squared distribution for all sample sizes considered. For non-normal populations, the least sample sizes needed for all 4 variance estimators to be approximated by the chi-squared distribution are 100 and 340, for Bernoulli (0.1) and Poisson (0.02) populations, respectively.

In regard to the variance estimators of ratio (proportion), for moderate to large samples (n>100), the simulation results are supportive of using a scaled chi-squared approximation to the distributions of the 4 variance estimates of the proportion. For small samples, the chi-squared approximation may not be so good. The results are better for larger values of the population proportion (for values less than 0.5).

The degrees of freedom of the 4 variance estimates of proportions differed across the various populations and across the various different variances. Degrees of freedom of the Taylor series method and the jackknife variance increased fastest with sample size n. The degrees of freedom of Fay's variance increased slower with sample size n and remained below the number of replicates (80) used in the Fay's variance. The degrees of freedom of the random group variance estimates increased slowest with sample size n and in general did not exceed 9 (the number of groups minus 1)

Bias of the variance estimates were generally small for n >100 (except for the random group variance estimates with non-negligible bias for n>100). There are negative biases for all variance estimators of proportion for very small sample size. There are some positive biases for the jackknife variance estimator for small, but not the smallest n.

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: No significant progress this quarter.

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: Staff continued working on methods to identify suspicious records in our foreign trade data without the use of the Edit Master ratio edits parameter file. We use the selective editing software which has previously targeted manual review of rejected records on all the incoming data. Staff prepared a draft of a division report describing research for this project.

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. Recall that the Research and Methodology (R&M) team was formed to discuss the specific problems associated with the current use of the cell suppression 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 FY2010. The actual coding of the new version, likely with a modern language such as C++, will be done by a contractor, who is now on board.

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. 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. There has also been an exploration of mathematical modeling languages. Such 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: Seasonal adjustment and X-12-ARIMA support was provided to American Forest & Paper Association, The Conference Board, Haver Analytics, International Labour Organization, Intuit Inc., PC-Software Solutions (Germany), Mathworks, SAS, the *Wall Street Journal*, 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, Bureau of Labor and Statistics, Augusta State University, and the University of Valencia.

Staff also developed a version of X-12-ARIMA on a Blade server for the use of staff from Housing and Household Economic Statistics Division in their production work.

Staff 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: 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.

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 began testing this code to check for errors and inconsistencies in the revised code. Staff also continued to develop HTML output for the X-13A-S prototype.

Staff continued to update the signal extraction diagnostics in X13A-S to implement the diagnostics in McElroy and Blakely (2009).

Staff 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: 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 this quarter, staff: (a) continued empirical studies of model-based seasonal adjustment diagnostics, (b) developed algorithms and code for long memory seasonal adjustment, (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 cyclical effects are present, (e) compared performance of two competing benchmarking procedures on time series, (f) studied performance of new distribution theory for Ljung-Box statistics, (g) studied how seasonal adjustment methods facilitate short-term forecasting of cyclical and trend dynamics, and (h) studied 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: Staff finished developing web content for X-12-ARIMA and 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. 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: Staff gave a Division Seminar titled "Imputation in the Survey of Research and Development in Industry Using a Last-Value-Dependent Nonresponse Mechanism," which described joint work on this topic. The seminar first provided an overview of the new imputation procedure for longitudinal data proposed by Xu, Shao, Palta, and Wang (*Survey Methodology*, 2008). The seminar then illustrated our work in applying and customizing the procedure to the Survey of Research and Development in Industry (SRDI). Additionally, staff had a meeting with staff from the National Science Foundation in which we went over the seminar slides and discussed current and future directions for this project. Staff also has been working on a paper in which we extend from a last-value-dependent to a past-value-dependent nonresponse mechanism. The theoretical work for this paper is in place and the past-value-dependent method has been applied to analyze data from the SRDI. In addition, staff prepared some simulation results to explore the properties of this method when applied to a population similar to that of the SRDI data, and to examine the nature of the regression functions that the method uses for imputation.

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: 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 for the joint JSM paper of Yang Cheng, Eric Slud, and Carma Hogue to be presented at 2010 JSM by Yang Cheng. Staff aided in writing the JSM paper draft and presentation slides for this work.

Staff: Eric Slud (x34991)

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: Staff continues 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 qv Rule. The Drop qv Rule is a universe subsampling routine in the MAS, which is meant to protect against differencing attack disclosures. For example, a differencing attack disclosure occurs on the MAS when a data intruder creates two similar universe data sets on the MAS:

U(n): a universe with n total observations

U(n-B): a universe with the exact same n observations as U(n) less B possibly unique observations.

The difference U(n) – U(n-B) = U(B), where U(B) is a universe that contains B possibly unique observations. Let T[ ] denote an m-way contingency table. Suppose a data intruder is able to define both U(n) and U(n-B) on the MAS, then obtains the following two similar m-way contingency tables, T[U(n)] and T[U(n-B)], then performs the following matrix subtraction attack on the cells of T[U(n)] and T[U(n-B)]:

(1) T[U(n)] – T[U(n-B)] = T[U(B)]

The resulting table T[U(B)] is an m-way table of counts built upon the manipulated universe U(B) = U(n) – U(n-B). Note that the T[U(B)] only contains B << n observations. For example, if B = 1, then the differencing attack in (1) yields T[U(1)], an m-way table of counts built upon the one unique observation contained in U(1) that contains a cell count of 1 in the cell that represents the partial microdata record for the one unique observation contained in U(1), and zeros within the remaining cells.

To prevent data intruders from performing differencing attacks like the one shown above, the MAS implements a universe subsampling routine called the Drop qv Rule. In the Drop qv Rule, for any universe data set U(n) defined on the MAS, the MAS first draws a random value of Qv = qv from a Discrete Uniform (1/(k-1)) distribution, where qv = {2,…,k} and the value of k is fixed by the system administrator. Then the U(n) data set is subsampled by removing qv observations at random from U(n) to yield a new subsampled universe data set U(n-qv). On the MAS, each U(n) data set is subsampled only once. That is, if the same U(n) was selected again by the same user, or by a different user, then the exact same qv observations would be removed from U(n) to yield the exact same U(n-qv) as before. Therefore if a data intruder attempted the differencing attack in (1), they would actually perform the differencing attack in (2):

(2) T[U(n-q1)] – T[U(n-B-q2)]

where T[U(n-q1)] and T[U(n-B-q2)] are two similar m-way contingency tables based on two independently subsampled universe data sets: U(n-q1) and U(n-1-q2). Note that q1 is not necessarily equal to q2 and the resulting table from the differencing attack in (2) may or may not yield a successful joint disclosure for all m observed variable categories for all B possibly unique

observations contained in U(B). That is T[U(n-q1)] – T[U(n-B-q2)] may or may not equal T[U(B)] = T[U(n)] – T[U(n-B)]. When T[U(n-q1)] – T[U(n-B-q2)] = T[U(B)] = T[U(n)] – T[U(n-B)], then we say that the differencing attack in (2) yielded a successful joint disclosure of all m observed variable categories for all B observations contained in U(B).

Staff members analyzed the Drop qv Rule for the case of B = 1 and B = 2 and found an approximate probability distribution that models the probability of obtaining two independently subsampled m¬-way tables T[U(n-q1)] T[U(n-B-q2)] from the Drop qv Rule. Using this approximate probability distribution, staff members found a function that models the total approximate probability of obtaining T[U(n-q1)] – T[U(n-B-q2)] = T[U(B)] = T[U(n)] – T[U(n-B)]. Using the concepts of Majorization and Schur-convexity, staff members found that T[U(n-q1)] – T[U(n-B-q2)] = T[U(B)] = T[U(n)] – T[U(n-B)] with a probability no greater than 1/(k-1), for a given value of k > 3. That is, after applying the Drop qv Rule, the differencing attack shown in (2) will prevent yield a successful joint disclosure of all m observed variable categories for all B observations contained in U(B) = U(n) – U(n-B) with a probability no greater than 1/(k-1).

Staff members have started to analyze differencing attack disclosures (2) for the case when B = 2. When B = 1, the universe U(1) contains one unique observation. When B = 2, the universe U(2) contains 2 possibly unique observations. We call these two observations A and B. Staff members have started to research the approximate probability of obtaining a successful disclosure of observation A for all m observed variables, but not a full disclosure of observation B for all m observed variables, from the resulting table from the differencing attack of T[U(n-q1)] – T[U(n-2-q2)].

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 are beginning to explore other possibly ways to generate synthetic residual bivariate plots by using multivariate statistical methods.

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

Future Plans: Staff will continue to work closely with staff members of DID to develop the beta AQS/MAS prototype. Staff members will work with members of DID to determine testing algorithms for the current set of confidentiality rules within the AQS/MAS prototype. Staff will work with members of the Human Factors & Usability Research Group and members of DID to develop, test and evaluate a user interface for the beta AQS/MAS prototype. Staff will continue to explore possible model fit diagnostic plots based on synthetic data methods for logistic regressions, and explore possible ways to develop matrix scatter plots based on synthetic residual methods.

Staff: Laura Zayatz (x34955), 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, 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 the third quarter of FY2010, all of the data on research1 was migrated to the research2 cluster using the program "rsync." Significant progress was made defining the "default" user environment. When the remaining three blade servers have been built, we will allow users on the system for testing. If this goes well, it is expected that by the end of FY2010, all of the users of research1 will be using the research2 cluster for their day-to-day research computing.

All but one of the findings related to research1 that were cited during Certification and Accreditation (C&A) of CEN14 have been addressed.

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: 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 (x34955), 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: Staff refereed two papers on microdata confidentiality. Staff reviewed, commented on, and corresponded with the authors of a critique of one of the algorithms of differential privacy.

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: No significant progress this quarter.

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 area in the implementation of a newly designed simulation to investigate the properties of two methods, M-estimation and Clarke Winsorization, that treat influential values. The team is investigating treatments for influential values in the Monthly Retail Trade Survey. The new plans incorporate comments from the designer of the M-estimation method, Jean-Francois Beaumont of Statistics Canada, on the results of the first scenario. Staff began planning for a *SUMMER AT CENSUS* visit by Jean-Francois Beaumont to the Census Bureau to discuss the design of the simulation and our initial results.

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.

Highlights: During the third quarter of FY2010, five task orders were modified and two were completed. To date, there have been 96 task orders awarded under the R&D2007 contracts, with a monetary value over $129 million (over $107 million obligated). Seventy-seven task orders have been completed and one task order terminated, leaving 18 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.

Highlights: During the third quarter of FY2010, five task orders were awarded, one was modified and one was completed. To date, there have been 8 task order awarded under the R&D2014 contracts, with a monetary value of $5.8 million, ($4.5 million obligated).

Staff: Ann Dimler (x34996)

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: Staff worked to develop a Census Coverage model for race/origin, age/sex, and tenure covariates., and found variance estimates for all E-Sample and P-Sample clusters in 2006 Census Coverage, under a two-stage cluster sampling model developed by Partha Lahiri.

Staff: Don Malec (x31718), Aaron Gilary, Partha Lahiri

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: Staff wrote draft papers on record linkage for the Joint Statistical Meetings.

Staff continued investigating generalizations of the Birthday and Collision Problems.

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 three 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: 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: Staff provided editing advice, references, edit generation and ratio editing code to staff from the National Oceanic and Atmospheric Administration. There was interest in verifying the feasibility of SPEER editing software for editing their fisheries data.

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: Staff continued to develop general software (PEP) for edit and imputation. The plan for development was presented at the Planning and Information Exchange meeting of American Community Survey Office (ACSO) managers. At the meeting and in follow-ups, staff addressed the following primary concerns of ACSO staff and staff from the Demographic Directorate: 1) New methodology designed to take advantage of administrative information. Staff started implementing various kind of multiple-imputation; 2) Computational resources. Staff is preparing a test based on the group quarter data; and 3) Administrative feasibility. Staff

consulted with staff from ACSO and the Housing and Household Economic Statistics Division to anticipate procedural issues involving the review of the analysts.

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: Variance estimation software support will revert to the respective directorates as of July 1, 2010. General variance estimation development and support, including research and methodology, will remain 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: Staff generated synthetic data for disclosure avoidance for the 2009 ACS group quarters sample. Synthesis was performed via a prototype version of an SRD-written R package named "pep."

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: Work continued 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 related to ten of the proposals related to improving corporate efficiency. 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: William Winkler (x34729), Rob Creecy, William Yancey, María García

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.

Support for X-12 ARIMA Documentation and Software (Statistical Research Division)

Highlights: No significant progress this quarter.

Staff: Lawrence Malakhoff (x33688), Brian Monsell

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

Highlights: Staff worked with Administrative and Customer Services Division (ACSD) this quarter to address and correct accessibility issues in PDF and HTML data tables. 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 2 or 3 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 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: No significant progress this quarter.

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

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

Highlights: At the request of staff from Economic Statistical Methods and Programming Division, staff reviewed CIDR screen mockups for usage of color and readability. The only 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: 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 (per Census Bureau IT standard 15.0.2). 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 provided to the sponsor for their action.

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: 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: Lawrence Malakhoff (x33688), Victor Quach

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.

WebTA
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. These findings and recommendations were compiled into a report and sent to the Department of Commerce for its action.

Staff: Lawrence Malakhoff (x33688)

A.3. AFF Usability Study: Iterations 1 and 2 – Conceptual Design and Low-Fidelity Prototype Testing, and Iteration 3 – Higher Fidelity Testing
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: Staff wrote the quick report for iteration 3 usability testing. Some of the high priority problems uncovered during 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. 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 issues identified.

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

A.4. AFF Usability Study: Baseline Testing
Description: American FactFinder (AFF) is a free, public online tool that allows users to find, customize and download data on the population and economy of the United States. The AFF Web site is undergoing a thorough redesign under the sponsorship of the Data Access and Dissemination Systems Office (DADSO). In order to evaluate whether the re-designed effort is successful, the Usability Lab proposed and conducted 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 AFF Web site, expected sometime in 2011 or 2012.

Highlights: No significant progress this quarter.

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

A.5. 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: No significant progress this quarter.

Staff: Kathleen Ashenfelter (x34922)

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: Staff continued working on the final report. Staffs sent the draft to a reviewer and are currently making edits based on reviewers' comments.

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

A.7 Usability Testing of Iteration 3 of the New American FactFinder
Highlights: No significant progress this quarter.

Staff: Jennifer Romano (x33577), Jennifer Chen, Erica Olmsted-Hawala

A.8 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: Staff performed a literature review on Web 2.0 and digital agents for presentation at the International Field Directors and Technologies Conference 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, discussed potential implementations of avatars for Web surveys, and summarized future research.

Staff is drafting a paper for American Association for Public Opinion Research's on-line journal *Survey Practice* on how ECAs function and interact with a respondent in a Web survey.

Staff: Lawrence Malakhoff (x33688)

A.9 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: No significant progress this quarter.

Staff: Jennifer Romano (x33577), Jennifer Chen

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: Eleven submissions were sent to the Office of Management and Budget for approval under the generic clearance for questionnaire pretesting research: cognitive testing of CPEX experimental race and ethnicity questions in Spanish; cognitive testing of the 2010 Annual Survey of Manufactures Management and Organizational Practices (MOPS); usability testing of the 2010 Annual Survey of Manufactures and 2010 Report of Organization web application; usability testing of the Census Quality Survey web application; usability testing of the 2010 Business and Professional Classification Report; cognitive testing of the Business R&D Innovation Survey; usability testing of the American Community Survey rostering section; cognitive interviews on the 2011 NCVS School Crime Supplement; unstructured interviews on the Large Company Reporting Practices Program; unstructured interviews for the 2012 Census of Governments Finance Component and Retirement Component; focus groups and in-depth interviews on census enumeration of non-English speaking households.
Staff prepared an Office of Management and Budget package for a three-year extension of the generic clearance for questionnaire pretesting research and circulated it for approval within the Census Bureau. The package went to the Department of Commerce for approval.

Staff conducted an expert review of new questions proposed for the Survey of Fishing, Hunting, and Wildlife-Associated Recreation.

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: Staff finalized a paper comparing this Random Digit Dialing study to a field pretest.

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: 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.

During this quarter, staff continued to collaborate with researchers at the National Cancer Institute to work on examining the effectiveness of cognitive interviewing techniques for English and Chinese. Staff will conduct interviews 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 vs. cultural effect.

In addition, staff continued to be active on 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 use in organizing a panel of experts to help the Census Bureau work on language proficiency tests. Staff also worked on a series of conference papers to examine the effectiveness of translation-review process and cognitive interview probes, and analyzed ACS data to identify patterns of nonresponse among Chinese speakers. Staff presented findings of this research at a professional conference.

Staff started a new research project to develop a coding scheme 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. Staff worked on a conference paper to present findings associated with the application of the coding scheme to Census Bureau translation projects.

Additionally, staff continued to work on a book project 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: Staff conducted cognitive interview training for seven people from our division and Demographic Surveys Division.

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 this quarter, staff worked on an article for submission to academic journals.

Future Plans: The article will be completed and submitted to academic journals for consideration.

Staff: Patricia Goerman (x31819)

G. Interviewer-Respondent Interactions
Description: Survey nonresponse rates have been increasing, leading to concerns about the accuracy of (demographic) sample survey estimates. For example, from 1990 to 2004, initial contact nonresponse rates have approximately doubled for selected household sample surveys including the Current Population Survey (CPS) (from 5.7% to 10.1%). While mailout/mailback is a relatively inexpensive data collection methodology, decreases in mailback rates to censuses and sample surveys mean increased use of methodologies that bring respondents into direct contact with Census Bureau interviewers (e.g., field representatives) using CATI (computer assisted telephone interviewing) or CAPI (computer assisted personal interviewing). CAPI can include face-to-face or telephone contact. Unsuccessful interviewer-respondent interactions can lead to increased costs due to the need for additional follow-up, and can also decrease data quality.

Highlights: No significant progress this quarter.

Staff: Tommy Wright (x31702), Tom Petkunas

H. 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. And, 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: No significant progress this quarter.

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

I. 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: The SHIPP 2010 field test continued, rotating through all three field periods and all three versions of the questionnaire, and concluded on May 10. Staff returned to Hagerstown several times to train interviewers on the new version of the questionnaire, monitor the first days of production of the new field period, and conduct interviewer debriefings. During production interviewing staff carefully monitored daily progress reports and adjusted staff and sample as needed (e.g., when the Medicare sample quota was reached, it was pulled in order to focus all interviewing resources on the RDD sample cases). Upon completion of interviewing, staff calculated response rates and prepared a final methodology report.

On a separate but related project, staff carried out research examining the role of state-specific names for government health programs in the Current Population Survey and the American Community Survey. The research involved two phases. First was an analysis of CPS survey estimates and administrative records, considering differences in the nature and number of state-specific names used as fills in the questionnaire. The second phase involved a comparison of CPS and ACS estimates of Medicaid coverage, again considering differences in questionnaire stimuli for coverage under government programs.

Staff: Joanne Pascale (x34920)

J. 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: In response to the Director's call for staff to identify challenges under the new Corporate Priorities Program, a staff member 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 of those trends.

Staff: Laurie Schwede (x32611)

K. Using Vignettes to Explore Survey Concepts

Description: Vignettes are a common tool for survey pretesting. Vignettes 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 either being 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: Staff presented the findings of the exploratory research in a poster, "On the Usefulness of Vignettes in Exploratory Research," at the American Association for Public Opinion Research Annual Conference. 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

L. 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: Staff finalized data collection, analyzed the data, and presented them in a poster, "The Effects of Framing and Retrieval Fluency on Feeling-of-Knowing (FOK) Judgments," at the Association for Psychological Science Annual Convention. The results show that the framing and retrieval fluency manipulations did not appear to affect the participants' average 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

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.*

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

Herzog, T. E., Scheuren, F., and Winkler, W. E. (In Press.) "Record Linkage." *Wiley Interdisciplinary Reviews: Computational Statistics, 2.*

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

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.

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

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

3.3 PROCEEDINGS PAPERS

3.4 STATISTICAL RESEARCH DIVISION RESEARCH REPORTS

<http://www.census.gov/srd/www/byyear.html>

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

3.6 OTHER REPORTS

4. TALKS AND PRESENTATIONS

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."

5. STATISTICAL RESEARCH DIVISION SEMINAR SERIES

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

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.

.

6. PERSONNEL ITEMS

6.1 HONORS/AWARDS/SPECIAL RECOGNITION

6.2 SIGNIFICANT SERVICE TO PROFESSION

Kathleen Ashenfelter
- Associate Editor, *Frontiers in Quantitative Psychology and Measurement.*

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

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,* and the *Journal of Nonparametric Statistics.*

Brian Monsell
- Webmaster, Business and Economic Statistics Section, American Statistical Association.

Mary H. Mulry
- Refereed articles for *Public Opinion Quarterly*.

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

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.
- 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).

Asoka Ramanayake
- 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.

Jen Romano
- Reviewed a symposium 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).

Natalya Titova
- 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.*
- 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.

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.

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

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).