STATISTICAL RESEARCH DIVISION
FY 2010 SECOND QUARTER REPORT
–January, February, March 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 Variances*	
E.	*ACS Small Area Estimation for Selected Characteristics*	
F.	*ACS Small Area Estimation for Group Quarters*	
G.	*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*

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	

Housing and Household Economic Statistics Division (Sponsor)

Project Number	Project Title	FTEs
1465444	Re-Engineered Survey of Income and Program Participation Methodological Research	1.75

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

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: Staff continued to provide expert guidance on the development of and testing plan for the experimental Internet census form for the 2010 Census Quality Survey.

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

B. Development of Race and Ethnicity Questions
Description: 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 second quarter of FY2010, staff began working with Decennial Management Division and Population Division to pretest 11 different Alternative Question (AQE) experimental 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. Staff 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: Rodney Terry (x35475), Jennifer Hunter Childs, Patricia Goerman, Terry DeMaio, Yuling Pan, Matthew Clifton, George Higbie, Nathan Jurgenson, Amelia Tseng

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: No significant progress this quarter.

Staff: Patricia Goerman (x31819), Yuling Pan, Virginia Wake Yelei, Matthew Clifton, Anissa Sorokin

1.3 DATA COLLECTION PLANNING AND DEVELOPMENT (Decennial Project 5310001)

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

Highlights: Providers of Web survey tools claim to conform to *Section 508*, but are the applications actually accessible? Staff corresponded with software developers at the University of Michigan, ISR, Indiana University, Research Triangle Institute, Washington State University, and the Bureau of Labor Statistics to obtain sample surveys with common questionnaire elements and performed a comparison of the accessibility of question grids, radio button groups, check-boxes, logical keyboard navigation with the presence of visual focus, and data-entry field labels. The products evaluated include Microsoft .NET 3.5, Illume 4.7.0.46, Blaise IS 4.8.1.1460, Hatteras 3.0, SurveyMonkey (2010), and Jdeveloper & HomeSite (J&H). These packages were evaluated with Internet Explorer 7 and the Job Access With Speech (JAWS) 11 screen-reader software.

The evaluation revealed Illume and Blaise IS developers need to inform screen-reader users about table commands to make their question grids accessible. The only web survey to provide visual focus when tabbing with JAWS 11 running was developed with J&H. Microsoft .NET developers need to provide an alternative to color to inform respondents about missing data. JAWS users can read J&H and SurveyMonkey screens from top to bottom without back-tracking to locate the question. Vocalized labels did match text visible on the screen, except for

Blaise IS and Hatteras. SurveyMonkey developers should correct display issues before posting a survey. The evaluation criteria ranked the Web survey development tools in this order (most to least accessible) J&H, SurveyMonkey, Hatteras, Illume, Blaise IS, and Microsoft .NET.

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

B. Decennial Reinterview Internet Testing – Usability Input

Description: The usability team 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: Staff performed a first round of usability testing on a wireframe instrument and reported the overall results to the sponsors. Staff also began cognitive and usability testing of the mailing materials for this project.

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

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 software for the correction of the previously released public use microdata sample (PUMS) files from Census 2000. Staff created the input files for the program based on the text files and explored 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.

Staff developed household swapping software for the ACS 5-year data product. The goal was to redesign the swapping procedure so that the data product will provide disclosure protection at the block group level. An analysis was performed after the new procedure was applied to the data, and a *Census Confidential Report* was written on the findings. Staff evaluated and documented the current swapping procedure for the states of New Jersey and California at the county level.

Staff developed and tested new software for 1- and 5-year ACS group quarters data. The software includes several new features including consistency verification (via Bill Winkler's DISCRETE system as recoded by Ben Klemens in C), automatic model selection via AIC, and improved fingerprinting of at-risk records. Staff also worked on the ACS03-06 PUMS age correction.

Staff members applied measures of data utility and disclosure risk to compare a newly developed method of imputing geographies for comparison with the current swapping method. A report is forthcoming.

Staff developed Decennial 2010 data swapping software for the Island Areas.

Staff members worked extensively with WESTAT staff on the Census Transportation Planning Package ACS 5-year special tabulation research project sponsored by the National Academies.

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

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: Two related specifications have been revised: "2010 Decennial Census Coverage Followup and Census Coverage Measurement Match Modeling Software Requirements Specification" (procedures for evaluating person links) and "2010 Decennial Census Census Coverage Followup and Census Coverage Measurement Person Matching Parameter Software Requirements Specification" (parameter settings for matching). Staff are co-authors for both specifications. The revisions include a modification of the review limits for the within-response modeling. Based on the results of staff's examination of a national matching of Census 2000 data, the new review limits are expected to slightly reduce the number of links that need to be reviewed.

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: One project during the Second Quarter focused on the statistical analysis and evaluation of an income sample following a lognormal distribution, when the sample includes top coded values—i.e., incomes above a specified level are not reported in the sample, for the purpose of disclosure control. Thus top coding results in a singly right censored sample. Even when the lognormality assumption holds, data analysis can be complicated under top coding. For analyzing income data in this scenario, a methodology that has been recommended in the literature consists of the imputation approach. The goal of the project is to develop data analysis techniques based on other novel methodologies available in the statistics literature, avoiding the use of imputations. Application of the novel concepts of generalized p-values and generalized confidence intervals will be investigated for analyzing top coded income data, under the assumption of lognormality. Confidence intervals will be derived for a variety of income parameters that are of practical interest: the mean, median, percentiles, Gini coefficient, fraction of incomes that are top coded in the population, etc. It is anticipated that the above concepts will result in a unified methodology for analyzing top coded data.

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).

Estimation of synthetic estimators in the CCM program

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

Empirical Best Prediction (EBP) method to estimate the proportion of erroneous enumeration for the 2006 Travis County, Texas test data
Highlights: The project is very much related to the Census Bureau's effort to understand how future censuses can be improved. The main goal of this project is to develop a robust statistical method to estimate the proportion of erroneous census enumeration, an important component of possible census error that the CCM group is studying. A model and the associated estimation methodology have been developed based on certain moment conditions without fully specifying parametric distributions. The method will be tested on the 2006 Travis County, Texas test data.

Developing a model for using race-based partitioning
Highlights: Staff analyzed data from Census 2000 and Census Coverage follow-up work to develop a race-based, data-driven model for match rates and correct enumeration rates. The modeling is based on recursive partitioning techniques staff have worked on previously. Staff presented results to Decennial Statistical Studies Division stakeholders in the Census Coverage Estimation Team on March 4, and to the component missing data subgroup several times.

Nonignorable component estimation coverage models
Highlights: Staff is working on a nonignorable model based on the work of Nandram (2002) to be used for assessment of the nonignorable models for census coverage estimates. The model has been modified to avoid truncating the multiplicative effects in order to keep the estimates in a valid range. Staff is also

investigating a generalization of the Nandram model to non-tabular data, where each individual record has different values of predictor variables. The tabular data will be a special case of the generalization. Comparisons of these models to their nonignorable counterparts are planned.

Nonignorable component estimation coverage models
Highlights: Work continues on developing a nonignorable/ignorable model hybrid model that consists of a mixture of saturated models. This new model is being compared to some of the more popular nonignorable models that have appeared in the literature using a combination of simulation studies and data analysis. This study will be a useful tool for analyzing the sensitivity of the estimates to different model assumptions.

Logistic Mixed Model Estimation
Highlights: The evaluation of the use of random effects to account for terms left out of the fixed-effect logistic coverage model is nearly complete. Treating all "ROAST" interactions as random effects from the same distribution produces a very small REML variance component estimate and, hence, allows little influence to the model. Staff are looking at some simple ways allow more influence, such as sub-setting the "ROAST" interactions by their significance level or clustering their estimates via a fixed effects model. Once this evaluation is concluded, the results will be disseminated and a written report will be provided.

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

In addition to their work on the projects above, staff also attended weekly meetings of each of the three Decennial Statistical Studies Division coverage measurement subgroups: the CCM estimation group, the component missing data group, and the small area estimation group.

Staff: Don Malec (x31718), Aaron Gilary, Ryan Janicki, Jerry Maples, Julie Tsay, Partha Lahiri

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 continued previous work on the error structure for estimates of components of census coverage error by refining the structure and identifying sources of the errors. The focus was on preparing a comprehensive document that describes the work to date.

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 observed the 2010 Census Coverage Measurement Initial Housing Unit Followup operation.

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 in the field test plans for the Targeted CFU (TCFU), which will be fielded in 2011 using cases that have been identified as duplicates in the 2010 Census. Staff also assisted the Decennial Statistical Studies Division in planning a large-scale cognitive test and qualitative study of the TCFU and duplicates in the 2010 Census.

Staff: Jennifer Hunter Childs (x34927), Matthew Clifton, Anissa Sorokin, Nathan Jurgenson, George Higbie, 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: At the invitation of the sponsor, one staff member served as a critical reviewer of the Decennial Statistical Studies Division's 2010 CPEX Experiment Study Plan: "The Avoid Followup Experiment Study Plan" and submitted an extended list of comments and suggestions on additional research questions that would strengthen this study plan.

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

C. Evaluations, Experiments, and Assessments Operational Integration Team (EEA OIT)

Description: The purpose of the EEA OIT is to facilitate planning and timely implementation of 2008 Census Dress Rehearsal and 2010 Census evaluations, experiments, and assessments. The group guides and monitors the development, implementation, and reporting of the 2010 evaluations, experiments and assessments. It ensures that program integration and implementation of the 2010 Census Program of Evaluations and Experiments (CPEX) meets the guidance provided by the Census Integration Group and prepares and monitors the 2010 Census Program for Evaluations and Experiments Master Plan.

Highlights: Staff attended team meetings and reported on progress of our division's CPEX evaluations.

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 received and incorporated Census Program for Evaluation and Experiments (CPEX) study plan comments from critical reviewers for the respondent debriefing project.

Staff supplied the 2006 Questionnaire Design Experimental Research Survey (QDERS) telephone center training material to the Decennial Statistical Studies Division (DSSD). DSSD will use these materials to conduct the recall bias study for CCM. DSSD modified the materials for their project and staff then commented on DSSD's draft versions of the materials. Staff drafted a Request for Information to independently

study memory of landmark events. This will be a separate project not connected with the CCM recall bias study.

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 prepared and finalized a statement of work, budget, and all other necessary documents to secure six contracts with seven ethnographers for the GQ study. Staff handled all the initiation and followup procedures for background checks, Title 13 training and security badges for the contractors. Staff organized a two-day orientation and training conference for the seven ethnographers who will be conducting the ethnography study of group quarters in 2010. Staff reviewed and provided comments to the final proposals submitted by each of the ethnographers. Staff monitored and provided continual guidance to the ethnographers regarding their field activities throughout the quarter. In February, the lead researcher contacted and met with administrative staff of one of the study sites to prepare for the CCM-like study. Staff designed and pre-tested a new questionnaire to conduct a CCM-like study for the university housing population. Staff submitted the final instrument to Administrative & Customer Services Division for imaging. Staff submitted an Office of Management and Budget clearance package to obtain clearance for all post-enumeration interviewing and focus group activities.

Staff: Anna Chan (x38462), George Higbie, Temika Holland

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 of "Observing census enumeration of non-English-speaking households in the 2010 Census" aims to identify language and cultural barriers in census interviews with respondents who speak little or no English and to examine the effectiveness of the 2010 Census Language Assistance Program. During this quarter, staff completed the study plan and sent it out for internal review. The study plan was revised based on comments from six critical reviewers. Staff briefed the Census Integral Group (CIG) of the study plan and received CIG approval of the plan. We worked to identify bilingual ethnographers and select observation sites. We have identified seven research sites and have assembled a multilingual research team with seven sub-teams consisting of 23 ethnographers in seven languages (Spanish, Chinese, Korea, Russian, Vietnamese, Arabic, and Portuguese). Contracts were issued for the seven teams. In addition, we developed a theoretical framework to guide the study, and drafted an observation protocol guide, as well as respondent and enumerator debriefing questions.

Staff: Yuling Pan (x34950), Anissa Sorokin, Steven Lubkemann, Marissa Fond, Amelia Tseng

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 presented study plans and schedules for the 2010 Behavior Coding of NRFU, CFU and CCM PI to senior staff. Staff also continued to coordinate data collection and analysis with other divisions.

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, Nonresponse Followup, and Census Coverage Measurement. 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: Staff developed a short project overview flyer to identify ethnographers interested in conducting one of up to 18 observation/debriefing studies during three Census field data collection operations based on personal visit interviews: Update/Enumerate, Nonresponse Followup, and Census Coverage Measurement. We shared information with ethnographers at the Washington Association of Professional Anthropologists and sent out the flyer on listservs for specific professional interest groups within the American Anthropological Association, including Indigenous Anthropologists, Latina/o Anthropologists, Black Anthropologists, and Middle Eastern Anthropologists. We also sent the flyer to the American Sociological Association's Sections on Population Group, Race and Ethnic Minorities, and Sociology of the Family. We received statements of interest from more than 120 ethnographers around the country.

Staff initiated a meeting with staff from LTSO to explore the possibility of our outside contractors being issued long-term loaner laptops and flashdrives for the duration of this project to enable them to transcribe Title 13 Census Personally Identifiable Information (PII) from tapes while working from their homes. Our reason for requesting this significant change in IT policy (regular Census Bureau employees are not permitted to work with Title 13 data on loaner Bureau-issued laptops while out of Census Bureau Headquarters unless certain changes are made to the laptops) was that we were doing a coverage study where it was critical to the success of the project that we capture detailed information such as names, addresses, birthdates and ages accurately in order to conduct later matching successfully. We conducted negotiations over several weeks and ultimately reached an agreement to go forward with this new policy. This agreement has now been written into the overall Cen08 IT Security Policy for the Census Bureau, allowing long-term loaner laptops to be used at remote locations to process and store Title 13 data and secure encrypted Safeboot flashdrives to be used to transfer data from the loaner laptop to Census secure systems via secure FEDEX packaging and mailing. This is a new IT policy program that can be used for other projects, with approval from our division chief. A memo documenting the contractors' routine reporting requirements is currently being developed jointly by division staff and staff from LTSO.

Staff selected the outside contractor to observe/debrief in the Update/Enumerate Operation on the Indian reservation from candidates with previous experience working on Indian reservations. Staff wrote a contract with a detailed description of the observer/debriefing tasks, a statement of work with a justification for selection of that contractor, an IT Security Checklist, and a cover report document, and got successive approval signatures from ITSO, LTSO, the Policy Office and the Privacy Office, as well as from our division's admin office and then worked with Acquisitions to get the contract through. Staff also worked with Security officials and the contractor to get his security clearance completed. Staff worked with staff from Field Division the Regional Census Center to select the Indian reservation which fulfilled the project's eligibility criteria, and we requested and received tribal council approval to conduct this study. Staff prepared a brief written overview of the project to distribute to Local Census Office staff. This includes instructions telling the enumerators what they should be doing during the interviews and cautioning them not to change their standard interviewing procedures as a result of anything the observer asks during the debriefing. This handout also provides guidelines and verbatim text on how the enumerators should introduce the observer to the respondents and how the observers should request the respondents' permission to tape. We also developed the Update/Enumerate checklist for the observer to complete for each interview. Lorraine Randall assembled all equipment and interviewing materials and sent them to the contractor in the field site via FedEx. The contractor began his observation debriefing study at the end of this quarter.

Staff also began selecting the seven additional contractors for the 2nd observation/debriefing phase in the NRFU Operation to take place in May.

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

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 draft study plan and draft file listing to critical reviewers and some additional interested

parties. Both documents incorporated the recent dropping of field followup from the evaluation.

Staff: Michael Ikeda (x31756), Mary Mulry

J. 2011 Relationship Survey
Description: Recent changes in the American legal and social landscape with respect to family composition, relationships, and same sex marriages have potential impacts on the content of the relationship and marital status questions in Census Bureau censuses and surveys. This project involves a program of research and testing that will guide the development of revised questions. Exploratory focus groups will be conducted across the United States with members of cohabiting couples to collect qualitative information about alternative terms, definitions, categories, and/or questions that most accurately measure relationship status and partnership situations. Cognitive interviews will be conducted to evaluate questions developed from the results of the focus groups. The revised questions will be subjected to further testing and evaluation before being implemented by the Census Bureau.

Highlights: Staff attended most of the 18 focus groups that were held in Washington, DC; San Francisco, CA; Houston, TX; Boston, MA; Ft. Lauderdale, FL; Augusta, GA; and Topeka, KS with members of cohabiting same-sex and opposite-sex couples. These groups were conducted under contract and observed by project staff from our division, as well as from the Office of the Director and Population Division. Summaries of all the focus groups have been completed.

Future plans: Staff will present a paper on the focus groups at AAPOR. Revised relationship and marital status questions will be developed based on focus group results, and cognitive interviews will be conducted with these revised versions.

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

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 produced, installed, and executed new production code for ACS group-quarters disclosure avoidance, which generated new synthetic data for three collection years (2006-2008). Execution of code for 2009 data is forthcoming.

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.

Highlights: Staff interacted with client Freddy Navarro to conduct a study comparing ACS Multi-Year Estimates with linear combinations of one-year Estimates. This project was approved by the ACS Research and Evaluation Committee. Staff formulated a testing framework, implemented corresponding statistical procedures in SAS, and presented preliminary results to client. Upon receiving feedback, staff undertook additional computations.

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

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

*Highlights: S*taff completed revisions for a co-authored journal paper to be submitted under the title "Simultaneous calibration and nonresponse adjustment." Revisions were made to a second paper, "Modeling log-linear conditional probabilities for prediction in surveys" by Thibaudeau, Slud, and Gottschalk, which is to be submitted to a journal and is also the draft paper for Thibaudeau's JSM 2010 paper. Both of these papers are

the culmination of 1-2 years work on alternative methods of estimating variances for complex survey estimates: the first paper does this in the context of nonresponse-adjusted and population-controlled estimates, and the second in so-called `hybrid' estimates containing some design-based cell size estimates along with model-based conditional-probability estimates within the cells.

Staff: Eric Slud (x34991), Yves Thibaudeau

E. ACS Small Area Estimation for Selected Characteristics

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

Highlights: A staff member reviewed file information for StARS data sources in order to determine what data is available for income and employment. There is income information from two Department of Housing and Urban Development files—Public and Indian Housing Information and the Tenant Rental Assistance Certification System—as well as from IRS tax returns. There are no direct sources of employment information.

Staff: Lynn Weidman (x34902)

F. ACS Small Area Estimation for Group Quarters (GQ)

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

Highlights: Major changes were made to the purpose and schedule for this project. Senior ACS staff imposed the requirement that the 1- and 3-year data files must have GQ data present in every county with a GQ, in addition to the 5-year data files having GQ data present in every tract with a GQ, which was the sole focus of previous work. Due to the additional work needed to fulfill this requirement and the short time available before this year's production weighting, research will continue this year with the goal of applying it in 2011 production.

Evaluation of the imputation procedure and accompanying weighting was extended to include a third data set: a preliminary evaluation on a 4-year ACS file constructed from 2006-2008 data files. This allows staff to adapt the imputation software to work with ACS files and test how well geography-related variables are imputed before the final evaluation with full ACS data.

The approach for the Census 2000-based simulation was changed to use a single GQ population for each of 25 5-year samples and their imputation, so that the combined sampling and imputation variance can be estimated without being confounded with variability among different GQ populations. Software for simulating the population was redesigned, rewritten, and tested, and the 25 samples were generated. As part of the population simulation, the changing size of the GQ frame by major type across the five years was also simulated. ACS sampling files were investigated, matched as necessary, and the results analyzed to determine the effective decrease in frame size, by identifying GQs deleted and added in each year. Then a procedure to select GQs for deletion to meet this changing size was developed and software written to apply it.

The details of methodology for selecting donors for imputation were continually discussed and refined. Software to analyze the number of GQs by county and tract for each major type in order to identify where imputation of small GQs is required was designed and is being written. The imputation software has been redesigned to accommodate recent developments and is being written. In addition, staff has reviewed an updated proposed weighting procedure and is determining the details of exactly how it will be applied to produce the 3- and 5-year estimates.

On March 17, an ACS GQ Small Area Estimation Workshop was held to present the proposed methods to staff from other divisions in order to get their reaction and input. Staff working on this project prepared written documents describing their particular tasks and presented it at the Workshop.

As an alternative approach, a staff member proposed a two-step synthetic estimation method to estimate totals for small areas (e.g., census tract) that do not have any ACS sample. The first step involves assigning each person in a given GQ to one of several groups formed by certain GQ-groups and person characteristic combinations. The second step involves estimation of the population size and the true mean of the characteristics of interest for each group using a synthetic method that assumes certain similarities between the small area and a large area that covers the small area. The method has not been applied to the ACS GQ data since the goal of this project is to obtain an imputation-based method. Using the similarity between the imputation and synthetic methods, development of a corresponding imputation-based method is planned for the future.

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

G. ACS Data Issues

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

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

Staff: Lynn Weidman (x34902), Chaitra Nagaraja

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

A. ACS Language Research
Description: This project provides technical and research support for addressing language issues in 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 second quarter of FY2010, staff worked closely with the ACS language team to carry out three 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 II); and 3) Cognitive testing of the ACS CAPI letters in five languages (Project C). Staff worked as Census Bureau research analysts, providing technical guidance and supervision for the contractor (RTI). Staff developed analytical frameworks encompassing theories in sociolinguistics and psychology to provide guidance for the contractor in summarizing interview data and reporting findings. Staff conducted training for the contractor's interviewers on the ACS language assistance guides project and Project C. We also developed two types of coding schemes for these two projects. One is a coding scheme using a sociolinguistic framework to code translation issues and probe technique problems, and the other is a coding scheme that uses principles of influence from a psychological perspective to code cognitive testing recruitment strategies and the interpretation of messages conveyed in ACS CAPI letters. Additionally, we have introduced a method (also a type of coding) of collecting respondents' survey experiences and reporting respondents' behaviors in the consent procedure as indicators and supporting evidence for analyzing respondents' verbal responses in the multilingual cognitive interviewing event. Staff also reviewed and provided guidance for interview summaries in the aforementioned three projects, and conducted cognitive interviews in English and Chinese for Project C.
Staff: Yuling Pan (x34950), Patricia Goerman, Virginia Wake Yelei, Rodney Terry, Matthew Clifton, George Higbie

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: Staff conducted a third round of usability testing and delivered a quick report of the result to the sponsors during Q2. Further analysis is planned for next quarter.

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 are reviewing sponsor comments on the draft report and will finalize the report in the next quarter.

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: Staff conducted a first round of usability testing during this quarter and delivered a quick report to the sponsors. The second round of testing was also begun in conjunction with cognitive and usability testing of the mailing materials for this project.

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

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 drafted a CATI Reinterview survey called the Attitudes and Behavior Study (ABS). Questions in this survey will ask about Internet use/non-use, privacy concerns, and effectiveness of mailing materials. These questions will vary slightly across the three respondent groups: Internet respondents, mail respondents, and nonrespondents. The Decennial Statistical Studies Division is currently reviewing the draft survey and creating specifications for the Technologies Management Office.

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: Staff wrote a test plan and prepared to conduct cognitive and usability testing of the mailing materials during this quarter. Testing began at the end of this quarter on these mailing materials along with a wireframe version of the Web instrument.

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

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 revised tasks for the Iterative 3.0 study of the ACS Website and then recruited participants, ran a dry run session, and ran 10 participants through the usability study. Staff met and discussed design issues. Staff met with the sponsor and discussed usability sessions and user performance.

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

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: No significant progress this quarter.

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: No significant progress this quarter.

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 presented initial progress and plans for proof-of-concept study using Current Population Survey production data as part of a session on paradata at the FedCASIC 2010 conference. The title of the presentation was "Using Paradata at the US Census Bureau: Demographic Current Surveys' History and Future."

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

C. 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: A second round of usability testing was conducted on the web-based National Survey of College Graduates. We tested four versions of the instrument on 30 participants between February 17 and March 19, 2010. The different versions examined the placement of the Previous and Next navigation buttons and the presentation of a long list of items (one column or two columns). Eye tracking was conducted. Results demonstrated that overall, participants were satisfied with the survey. Some issues uncovered in testing include (1) people do not read dense text on the log-in screen; (2) participants preferred the Previous button on the left side of the screen and the Next button on the right side of the screen; and (3) participants preferred the long job-code list to be in two shorter columns rather than one long, scrolling column. Recommendations were made and presented to the sponsor, and the final report is in progress.

Staff: Jennifer Romano (x33577), Jennifer Chen

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 began work on a project that seeks to provide guidance on revising the response category on the NCVS race of offender question at the request of the Office of Management and Budget. Staff submitted an outline of a report containing a literature review and proposed experiments that directly measure participants' ability to accurately identify the race of a face using current and proposed modified questions.

Future Plans: Staff will also conduct research on the NCVS School Crime Survey.

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 meetings of the CARI Project Team and the Quality Assurance Component JAD, attended Alpha and Beta demonstrations of the Behavior Coding Component (BCC) participated in the Performance and Acceptance Testing of the BCC, provided comments on the User Feedback tool, and provided comments on the E-training for the BCC.

Staff: Terry DeMaio (x34894), Jennifer Beck, Jennifer Hunter Childs, Matt Jans, Joanne Pascale,

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

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 second quarter of FY2010 coincided with the field period for the long-planned 2010 SIPP EHC-CAPI field test, for which interviewing began in early January and ended in mid-March. Field representatives (FRs) in six regions, only about one third of whom had prior experience working with SIPP (about one-third had no prior interviewing experience at all), completed 4,165 household interviews, for a response rate of 82%. Despite adverse circumstances (a brand new instrument, new interviewing procedures, and an interview felt by many to be excessively long), the field test actually achieved a substantially higher response rate than that of the wave 1 interview for the 2008 SIPP panel in comparable sample areas. Over many months, the Re-engineered SIPP Research Group, chaired by staff from our division, put in place a wide variety of evaluation efforts to address several basic research questions, including the following: Can the re-engineered survey capture data of comparable quality for a 12-month reference period as the current SIPP design does for a 4-month reference period? Was FR training on implementing the new EHC-based instrument effective, and how can it be improved? What features of EHC methods are effective at assisting respondents' recall of past events and circumstances, and how do they work? Research activities incorporated into the field test design include data analysis to compare estimates from the production survey with the re-engineered survey, and, where possible, to assess and compare data quality for both surveys against administrative records; collection of detailed cost data, to assess cost savings attributable to the new design; intensive observation of both FR training and field test interviews; FR debriefings, concerning both the instrument and training; training evaluation by trainers; respondent debriefings concerning EHC "landmark event" procedures; evaluation of recorded interviews; and item-level FR notes concerning instrument issues and problems.

Staff: Jeff Moore (x34975), Anna Chan, Joanne Pascale

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: As an alternative for the log number of poor children in the county model, staff has examined various model forms for county children poverty rates. After some exploratory analysis with nonlinear normal, beta-binomial, and logistic-normal models, staff selected the model using the normal distribution with a nonlinear mean link function for further investigation. The nonlinear normal model allows for a link function to restrict the mean to the interval of [0,1] for the poverty rate but still allows for the observed values of 0 percent poverty rate from the sample. With the nonlinear normal model, a big concern is how to parameterize the variance. The staff is considering three main alternatives: constant variance, binomial variance, and logistic variance. The logistic variance is proportional to the square of [p(1-p)], where p is the mean. Initial results seem to favor the logistic variance form. However, extensive model diagnostics are currently being examined.

Staff did a simulation study of Jackknife (delete one unit) and Random group (with 10 groups) variance estimators of the child poverty ratio, along with the usual Taylor series approximation of the variance estimator using random samples of various sample sizes drawn from artificial populations constructed from various Bernoulli and poisson distributions and from 2005 ACS data of the five largest counties in Maryland. We examined

simulation results to see whether the distribution of the variance estimates of poverty ratio can be approximated by a scaled chi-squared distribution and, if so, with what degrees of freedom. We obtained various measures (Kolmogorov-Smirnov statistic, degrees of freedom, coefficient of variation, relative bias) to evaluate the distribution of the variance estimates of poverty ratio from the simulated samples. It shows that the measures vary by the population parameters and the sample sizes.

All three variance estimators are significantly underestimated for very small sample sizes. The Jackknife and the Random group variance estimates are also overestimated for some sample sizes. The degrees of freedom of the Jackknife variance estimates and the Taylor series variance estimates are about the same magnitude and in general, larger than the degrees of freedom of the Random group estimator. The degrees of freedom of the Random group variance estimates are generally less than 9 (the number of the random groups minus one) for large sample sizes.

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: Staff finalized a new approach to estimating lower level (state level, county level) estimates using a linear transformation of the lower level parameters and drawing new parameters using the transformed conditional distribution and a random variable drawn from the higher level distribution. This new method seems intuitive and has good theoretical justification. However, questions remain concerning the practicality of implementing this new procedure. Simulations are currently being conducted using single linear constraints benchmarking state estimates to national estimates. If this is successful, the full model will be tested with multiple linear constraints and multiple levels.

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 applied selective editing to the full data set earlier in the editing process, which leads to a reduction in edit rejects. Staff continued working on a paper describing the application and testing of selective editing to these data.

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

1.17 DISCLOSURE AVOIDANCE METHODS (Economic Project 2470051)

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

Highlights: Staff members examined and documented more issues on statistical inferences under multiplicative noise and began documenting results from an application to a Census Bureau data set.

Most work for the economic directorate was concentrated on a single major project: cell suppression modernization. A Research and Methodology (R&M) team was formed to determine 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 will be done by a contractor to be hired soon.

Two of the short term topics have already been discussed in R&M team meetings. They are the handling of negative values and the handling of rounded values. Staff wrote short drafts on these two topics, as did others on the R&M team. On some topics, a consensus will be reached on how to solve a problem. On other topics, two or more solutions will be developed. The Disclosure

Review Board and survey managers may need to be consulted on the suitability of these solutions.

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 Franklin Templeton Investimentos (Brasil) Ltda., Baker & McKenzie Consulting LLC, Fortress Investment Group, Ford Corporation Apple Corporation, Morgan Stanley, SAS Institute, Bank of Japan, Bank of England, FHFA, Bureau of Labor Statistics, Bureau of Economic Analysis, South Carolina Department of Commerce, Statistical Research Institute (Korea), Swiss Statistical Agency, Australian Bureau of Statistics, Ministry of Economy, Trade and Industry (Japan), Statistics Sweden, Government of Turkey, Office of National Statistics (UK), Nankai University, and Universidad Nacional de Tucumán (Argentina).

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

Staff developed subroutines in Fortran and C++ for the new signal diagnostic routines in X-13ARIMA-SEATS. HTML output for X-13ARIMA-SEATS is also under development.

Staff added an option to allow users to specify whether a series is a flow or stock series into the series and composite specs of X-13ARIMA-SEATS.

Staff: 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 estimation of seasonal long memory models, (c) studied Markov Chain Monte Carlo techniques for Bayesian estimation of seasonal time series models, (d) studied properties of parameter estimates for mis-specified models when sampling error or cyclical effects are present, (e) compared performance of two competing benchmarking procedures on time series, (f) studied performance of new distribution theory for Ljung-Box statistics, (g) developed a C++ simulation environment for running X-12-ARIMA subroutines, (h) studied how seasonal adjustment methods facilitate short-term forecasting of cyclical and trend dynamics, and (i) implemented 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), Brian Monsell, Christopher Blakely, Ekaterina Sotiris, Natalya Titova, Bill 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 revised the X-12-ARIMA website to match the Census Bureau's new Look and Feel Guidelines, added new content related to new users of X-12-ARIMA, and provided listings of the Seasonal Adjustment Papers website by Author and Topics.

Staff updated the Genhol utility to include the capability to generate change-of-regime holiday regressors. The documentation for the X-13ARIMA-SEATS prototype was also updated.

Staff: Brian Monsell (x31721), Tucker McElroy, Natalya Titova, David Findley (consultant)

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 implemented the bootstrap described by Shao to evaluate the variance of estimator derived from longitudinally imputed data. The longitudinal imputation method accounts for cross-sectional nonignorability by taking advantage of information available longitudinally.

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

1.20 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 have tested and evaluated one of the confidentiality routines in the MAS known as the Drop q Rule. The Drop q Rule is a universe subsampling rule within the MAS, which is meant to protect against differencing attack disclosures. A differencing attack disclosure occurs when a data intruder forms two similar universes on the MAS:

U(n): a universe with n total observations
U(n-1): a universe with the exact same n observations as U(n) less one unique observation.

The difference U(n) – U(n-1) = U(1), where U(1) is a universe that contains only one unique observation. Let T[ ] denote an m-way contingency table. Suppose a data intruder defines both U(n) and U(n-1) on the MAS, then obtains the following two similar m-way contingency tables, T[U(n)] and T[U(n-1)], then performs the following matrix subtraction attack on the cells of T[U(n)] and T[U(n-1)]:

(1) $T[U(n)] - T[U(n-1)] = T[U(1)]$

The differencing attack (1) shown above yields a new m-way table, T[U(1)], that contains a count of one in the

cell that represents the unique partial microdata record for the one unique observation contained in U(1).

To prevent data intruders from performing differencing attacks like the one shown above, the MAS implements a universe subsampling routine called the Drop q Rule: for every universe U(n) defined on the MAS, remove a fixed q number of observations at random from U(n) to obtain a new subsampled universe data set U(n-q). If the same U(n) was selected again by the same user, or by a different user, then the exact same q observations would be removed from U(n) to yield the exact same U(n-q) as before. Therefore if a data intruder attempted the differencing attack in (1), they would actually perform the differencing attack in (2):

(2) $T[U(n-q)] - T[U(n-1-q)] = T[U(1)]$

where T[U(n-q)] and T[U(n-1-q)] are two similar m-way contingency tables based on two independently subsampled universe data sets: U(n-q) and U(n-1-q). Note that the resulting table T[U(1)] from (2) may or may not yield a successful disclosure of the partial microdata record for the one unique observation contained in U(1).

Staff members analyzed the Drop q Rule and found a function that modeled the probability of obtaining a successful disclosure from the resulting table T[U(1)] from the differencing attack shown in (2). Using the concepts of majorization and Schur-convexity, staff members determined that the probability of obtaining a successful disclosure from (2) was dependent on the distribution of the cell proportions within the original m-way table T[U(n)] fitted on the original non-subsampled universe data set U(n), and found that the probability of obtaining a successful disclosure from (2) was always bounded between a minimum value and 1. As a result, staff members determined that the original Drop q Rule did not provide as strong protection against differencing attack disclosures as was once though.

Staff members then developed a modified version of the Drop q Rule called the modified Drop qv Rule: for every universe U(n) defined on the MAS, first draw a random Q = qv from a discrete uniform distribution with probability mass function P(Q = qv) = 1/(k-1), where qv = {2,…,k}. Then, subsample the U(n) data set by removing the Q = qv observations at random from U(n) to yield U(n-qv). Again, if the same U(n) is selected by the same user or by a different user, the same Q = qv will be drawn again, and the same exact qv observations will be removed from U(n) to yield the exact same U(n-qv) as before.

Staff members then evaluated the effectiveness of the Modified Drop qv Rule in preventing the rebuilding of confidential microdata records for the differencing attack:

(3) $T[U(n-q1)] - T[U(n-1-q2)] = T[U(1)]$

where Q = q1 and Q = q2 are independently drawn from the same discrete uniform distribution as defined above. Staff members found another function to model the approximate probability of obtaining a successful disclosure from the resulting table T[U(1)] from (3). Staff members determined that the probability of obtaining a successful disclosure from (3) was bounded between some minimum value and 1/(k-1), which was a vast improvement compared to the probability of obtaining a successful disclosure from (2) when q was fixed for all universe data sets.

Staff members worked with other members of our division to design a user interface for the MAS for variable selection and universe formation. This interface will be documented and forwarded to members of DID to aid in the redesign of the user interface of the current MAS beta prototype.

Staff members are in the process of writing a memo that documents how to reconstruct an m-way table of counts through the results obtained through two separate queries on the MAS: one query for binary or multinomial logistic regression, and a second query for an (m-1) dimensional contingency table. The results obtained through these two queries can be combined to reconstruct the full m-way contingency table.

Future Plans: Staff will continue to work closely with staff members of DID to develop the beta AQS/MAS prototype. Staff will work with members of DID to determine confidentiality rule testing algorithms for universes created by linking observed variables using OR statements. Staff will begin thinking of possible ways to test the confidentiality rules and routines within the new beta AQS/MAS prototype for universe formation and ordinary least squared regressions as well as binary and multinomial logistic regressions. 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 possible ways to develop matrix scatter plots based on synthetic residual methods.

Staff: Laura Zayatz (x34955), Asoka Ramanayake, Jason Lucero, Paul Massell, Lisa Singh

1.21 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), Tina Arbogast, 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: The Computer Services Division built and delivered four identical IBM LS42 blade servers, each with four quad-core processors and 128 GB of RAM, for testing and customer acceptance. Computing throughput did not meet our expectations. In order to achieve throughput comparable to the existing SGI Altix 3700, three additional blades will be needed. Additionally there were problems related to GFS2 file locking when files are accessed by multiple nodes. These issues have been reported to RedHat Engineering. The current plan is to work through these problems and proceed with the migration.

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

The Certification and Accreditation of CEN14 was completed and interim authority to operate has been granted for the system. There were several findings that require action in the coming months. The goal is to address all of these findings by the end of FY2010.

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 members are currently researching the use of rank swapping as a possible way to fix the age-sex ratios of elderly men to elderly women in the CPS PUMS files. Staff members are researching ways to measure data quality vs. disclosure risk for this rank swapping technique.

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

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

Staff: 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: No significant progress this quarter.

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 in the Economic area to refine the design of a 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: 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 second quarter of FY2010, two task orders were modified. 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-five task orders have been completed and one task order terminated, leaving 20 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 second quarter of FY2010, two task orders were awarded and one was modified. To date, there have been 3 task order awarded under the R&D2014 contracts, with a monetary value of $2.9 million ($2.5 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: A method for the estimating design effect to be used for a semiparametric model for small area census coverage has been developed. Staff brought in data from different sources and provided preliminary estimates of variance and design effects. Staff are currently researching methods of calculating simultaneous equations in SAS.

Staff also modeled Census Coverage residuals to account for only mis-specification effects that are not part of a model and provided preliminary estimates of variance and miss-specification effects.

Staff: Don Malec (x31718), Aaron Gilary, Elizabeth Huang, Partha Lahiri, Jerry Maples

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 reviewed seven papers on record linkage from Statistics Canada. Staff sent papers, software, and advice on record linkage to Statistics Canada.

Staff continued investigating generalizing of the Birthday and Collision Problems.

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: There are two SAS versions of the genbnds software for generating a complete set of ratio edits for a given set of explicit ratio edits. Staff wrote a FORTRAN version of the code implementing a variation of the shortest path algorithm (no cost transformation) similar to the one we previously implemented in the SAS/IML code. The advantage is that this new version of the genbnds program can be embedded in the general package for data editing/imputation to handle continuous data under ratio edit constraints.

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 implemented infrastructure derived from edit programs written by William Winkler. The infrastructure support software for model-based imputation. Staff integrated in the software code for data synthesis. The integrated code was adapted for the purpose of imputing missing items based on a model.

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: Staff provided ongoing VPLX support to the Economic Directorate. Staff were contacted by the Manufacturing and Construction Division on the topic of calculating variance estimates for medians. Staff were contacted by members of the Economic Statistical Methods and Programming Division on the topic of creating a new VPLX executable for a new server in their division.

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 produced, installed, and executed new production code for ACS group-quarters disclosure avoidance, which generated new synthetic data for three collection years (2006-2008). Execution of code for 2009 data is forthcoming.

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: Staff continued work on generalized nonresponse-variance estimation routines for discrete data.

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.

Census 2010 Web Site (Systems Support Division)
Description: This Web site provides information via text, audio, and video to respondents about the 2010 Census.

Highlights: Staff continued to support SSD in their development of the 2010 Census Web site. The following aspects of the site were evaluated:

1) On the main page, the "Navigate to" combo box was accessible, but automatically goes to the first selection when using the down arrow. The remedy is to add the instruction "Use the alt and down arrow key to open the list" as title text to select an option.
2) Although accessible, a combo box was too narrow to display all text, so a recommendation was given to widen the field so text could be read by all users.
3) The links to access video stories in English and Spanish were not accessible. Staff recommended labeling and placing these links in the tab order so they would be made accessible to screen-reader users.
4) The media player provided to play video stories was not accessible. Users could tab to the image of the person, but cannot play the video. The visual focus did not follow the visual order of moving from upper left to lower right and no labels for player controls were spoken by the JAWS screen-reader.
5) The interactive Census form was determined to be inaccessible because all questionnaire snapshots were graphics with no detectable text. Staff recommended an accessible alternative, such as plain text.

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

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

Highlights: No significant progress this quarter.

Staff: Lawrence Malakhoff (x33688), Brian Monsell

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

Highlights: No significant progress this quarter.

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

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

Highlights: Staff evaluated the NotifyMe Web application and found and made recommendations for the following issues:

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

Staff: Lawrence Malakhoff (x33688)

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

Highlights: No significant progress this quarter.

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 & Programming Division)

Highlights: This application will require a low to moderate effort to correct issues identified in this accessibility review. Buttons, links, and combo boxes do not show focus when accessed by the tab key throughout the application. Instructions to use a combo box directed users to select more than one option which is not possible. Staff recommended using extended list-boxes which do support multiple option selection. A link to Adobe Reader is needed for users to read PDF content. We recommend another review to ensure corrections were made properly. This project is currently on the waiting list for testing in the usability lab.

Staff: Lawrence Malakhoff

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

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

Highlights: Staff performed accessibility evaluations for 26 documents for the CIS program. These included two documents for grades K-8 English Language Learners; three documents for Puerto Rico grades 9-12; thirteen documents for Diversity training; four documents for English as a Second Language Adult learners; and four documents for the Census Road tour.

Staff: Lawrence Malakhoff (x33688)

A.3. Census.gov Template Development
Description: The purpose of this study is to develop a set of templates with a consistent and usable look and feel for the Census.gov website. The template is intended to be used by both the demographic and economic domains of Census.gov. Some of the techniques to develop the template include card sorting, low-fidelity prototype testing, and usability testing.

Highlights: No significant progress this quarter.

Staff: Erica Olmsted-Hawala (x34893)

A.4. 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 worked on the final report of Iterations 1 and 2 of the AFF study, in addition to preparing tasks and conducting a dry run for Iteration 3 of the AFF web site. Staff ran seven participants through the study, notating usability issues with each session. Staff met with the sponsor and discussed the study in an ongoing basis during the testing.

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

A.5. 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.6. 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.7. 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 worked on the final draft of the ACS Website Baseline Usability report for submission to the division's *Research Report Series.*

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

A.8 Usability Testing of Iteration 3 of the New American FactFinder
Highlights: We conducted a round of testing on the new American FactFinder, which is currently in production. This usability study followed an earlier low-fidelity usability study of the new AFF Web site, which was carried out solely with paper versions of the interface and supporting materials and examined if participants understood the conceptual design of the site (Romano, Olmsted-Hawala, & Murphy, 2009), and two rounds of medium-fidelity usability tests, which were carried out on a computer (Romano, Chen, Olmsted-Hawala, & Murphy, in press). This version of the prototype had some functionality, and tasks were designed to enable participants to use the available functions. In this iteration, a high-fidelity prototype was presented to participants on a computer screen, and eye tracking was conducted. Eight people participated between March 2 and 17, 2010. The report and recommendations are currently in progress.

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

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: Ten submissions were sent to OMB for approval under the generic clearance for questionnaire pretesting research: focus groups to guide the development of relationship and marital status questions for the census and other demographic surveys, usability interviews on the data reliability indicator for American Community Survey (ACS) data tables, cognitive interviews on the translations of the ACS language guides in Chinese and Korean, cognitive interviews on translations of the ACS survey letters and a brochure in English, Spanish, Chinese, Korean, Russian, and Vietnamese; cognitive interviews on the Spanish translation of the ACS CATI/CAPI instrument; usability

interviews on the National Survey of College Graduates; usability interviews on the ACS website; cognitive and usability interviews on the ACS Internet questionnaire; respondent debriefing on the post-census-enumeration coverage measurement questionnaire for residence halls; usability interviews for the Medical Expenditure Panel Survey – Insurance Component; and in-depth interviews and focus groups for group quarters coverage measurement in 2020.

Staff sent the pre-submission notice for the extension of the generic clearance for questionnaire pretesting research to ACSD for publication in the Federal Register. Staff made a presentation on pretesting as part of the Demographic Surveys Training Course.

Future Plans: Staff will consult on generic clearance submissions as needed.

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 continued working on a paper comparing this random digit dialing study to a field pretest. In addition, staff worked on completing a final, comprehensive report for this study.

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

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

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

Highlights: Our staff worked collaboratively with researchers in 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, and 5) creation of best practices for the management of non-English language cognitive testing research.

During this quarter, we continued to collaborate with researchers at the National Cancer Institute to work on examining the effectiveness of cognitive interviewing techniques for English and Chinese. We will conduct interviews with four groups of respondents: monolingual English interviewed in English, bilingual Chinese interviewed in English, bilingual Chinese interviewed in Chinese, and monolingual Chinese interviewed in Chinese. With this design we will be able to dissect language effect vs. cultural effect.

In addition, staff members continued to be actively involved with the Federal Government Interagency Roundtable on Languages (IRL) to work on the development of the Census Bureau language assessment tool. Staff conducted a literature review on language proficiency testing and compiled a list of experts for organizing a panel of experts to help the Census Bureau work on language proficiency tests. We were also working on a series of conference papers to examine the effectiveness of translation-review process and cognitive interview probes. We also analyzed ACS data to identify patterns of nonresponse among Chinese speakers.

During the first quarter, in collaboration with university researchers, we continued to work on two book projects on discourse analysis and politeness communication.

Staff: Yuling Pan (x34950), Patricia Goerman, Anna Chan, Virginia Wake Yelei, George Higbie, Matthew Clifton

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: No significant progress this quarter.

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: No significant progress this quarter.

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: Staff continued development of the Q-Bank database and began investigating eTraining to use the Q-Bank system.

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

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 2010 field test was launched in late March, immediately following production CPS ASEC interviewing. Final preparation tasks for the field test included preparing for, conducting, and analyzing three phases of inter-divisional comprehensive testing of the instrument – specifically the users, systems, and verifications test. The advance letter was written, specifications for daily interviewing progress reports were written and tested, and training materials were finalized. Also, due to an unexpected response from the Office of Management and Budget (OMB) to our request to conduct this study under our division's generic clearance, staff prepared an OMB emergency clearance package. With regard to evaluations, staff held a focus group with four experienced telephone interviewers in the Tucson telephone center to develop alternative methods of gathering question-level feedback on the questionnaire. Staff then developed a hybrid closed/open ended question-by-question log for interviewers to record notes as the interviewing progressed. Staff conducted a general training with all 21 interviewers and 14 monitors and supervisors, and then conducted individual training sessions for each of the three groups assigned to each questionnaire treatment. Staff has been monitoring interviewing progress since the study's launch and making adjustments as needed. Staff also presented a paper on the pretest findings at the Survey Methodology Brown Bag series.

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: A paper documenting interim results for this project was accepted for presentation at the American Sociological Association Annual Meeting in August, 2010. It is entitled "Variations in Household and Family Structure by Race/Ethnicity and Geography: Census Coverage Implications."

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: No significant progress this quarter.

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 completed data collection and final data analysis. The results indicate that thinking about past successful or unsuccessful knowledge assessments does not significantly affect people's overall view of what they know (although the results are marginally significant). However, there were effects on the accuracy of people's predictions about their memory. People are most accurate at predicting what they know when they find it easy to think about past unsuccessful memory performance. They have a more realistic view of their knowledge. People are least accurate at predicting what they know when they think about past successful memory performance, even when they find it difficult to think about those past successes. They are over-confident in their knowledge. These results have implications for how people answer survey questions. The results will be presented at a poster session at the Association for Psychological Science convention.

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: Tina Arbogast, Matthew Clifton, Matt Gore (HRD), George Higbie, Temika Holland, Gloria Prout, Lorraine Randall, Kelly Taylor

3. PUBLICATIONS

3.1 JOURNAL ARTICLES, PUBLICATIONS

Fitzsimmons, P. and McElroy, T. (In press). "On Joint Fourier-Laplace Transforms," *Communications in Statistics*.

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

Herzog, T. H., Scheuren, F., and Winkler, W. E. (In press). "Data Quality," *Wiley Interdisciplinary Reviews : Computational Statistics*.

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

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

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

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

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

3.2 BOOKS/BOOK CHAPTERS

3.3 PROCEEDINGS PAPERS

3.4 STATISTICAL RESEARCH DIVISION RESEARCH REPORTS

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

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

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

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

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

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

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

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

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

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

3.5 STATISTICAL RESEARCH DIVISION STUDIES

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

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

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

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

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

3.6 OTHER REPORTS

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

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

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

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

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

4. TALKS AND PRESENTATIONS

Annual Conference of the Linguistic Society of America, Baltimore, Maryland, January 9-11, 2010.

- Yuling Pan, "Language and Measurement Research at the U.S. Census Bureau." (Poster presentation).

National Institute of Statistical Sciences/Education Statistics Service Institute Workshop, Washington, D.C., February 16, 2010.

- Laura Zayatz, "U.S. Census Bureau Disclosure Avoidance Practices and Research: An Update for NISS/ESSI."

FedCASIC 2010, Bureau of Labor Statistics, Washington, D.C., March 16-18, 2010.

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

5. STATISTICAL RESEARCH DIVISION SEMINAR SERIES

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

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

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

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

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

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*.
- Representative, Capital Area Social Psychological Association (CASPA) to the American Association for the Advancement of Science's (AAAS) Science and Human Rights Coalition.

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

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

Donald Malec
- Refereed papers for *Computational Statistics and Data Analysis* and the *Journal of the American Statistical Association*.

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

Tucker McElroy
- Publications Officer, Business and Economic Statistics Section, American Statistical Association.
- Reviewed papers for *Journal of the American Statistical Association* and *Statistics and Probability Letters.*

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

Mary H. Mulry
- Associate Editor, *Journal of Official Statistics*.

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

Erica Olmsted-Hawala
- Reviewed a paper for *Interacting with Computers.*

Yuling Pan
- Refereed manuscripts for the *Asian Journal of Communication* and the *Journal of Chinese Language and Discourse*.
- 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.
- Co-organizer, Session on Cross-Cultural Studies, 2010 Joint Statistical Meetings.

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.

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

Bill Winkler
- Refereed papers for *Survey Methodology* and the *Journal of Official Statistics.*
- Reviewed a proposal for the National Science Foundation.
- 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
- 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
- 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.
- Organizer, Privacy in Statistical Databases 2010.
- Member, Advisory Board, *Journal of Empirical Research on Human Research Ethics*.

6.3 PERSONNEL NOTES

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

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