Federal Committee on Statistical Methodology Research Conference

Final Program and Abstracts

Washington D.C Convention Center
November 4-6, 2013

Major Sponsors
Bureau of Economic Analysis
Bureau of Justice Statistics
Economic Research Service
Energy Information Administration
National Agricultural Statistics Service
National Center for Education Statistics
National Center for Health Statistics
Statistics of Income Division, Internal Revenue Service
U.S. Census Bureau
U.S. Environmental Protection Agency

Also sponsored by:
Abt-SRBI
American Institutes for Research
Nielsen
Sabre Systems
Westat

Hosted by:
Council of Professional Associations on Federal Statistics

FCSM RESEARCH CONFERENCE PROGRAM COMMITTEE

Pamela McGovern, Co-Chair, National Agricultural Statistics Service

Michael Planty, Co-Chair, Bureau of Justice Statistics

Wendy Barboza, National Agricultural Statistics Service

Benjamin Bridgman, Bureau of Economic Analysis

Louise Camalier, U.S. Environmental Protection Agency

Wan-Ying Chang, National Center for Science and Engineering Statistics

Richard Chard, Social Security Administration

Thesia Garner, Bureau of Labor Statistics

Jeffrey Gonzalez, Agency for Healthcare Research and Quality

Karen Hamrick, Economic Research Service

Lauren Harris-Kojetin, National Center for Health Statistics

Kim Henry, Statistics of Income Division, Internal Revenue Service

David Kashihara, Agency for Healthcare Research and Quality

Lynn Langton, Bureau of Justice Statistics

Preston McDowney, Energy Information Administration

Dawn Nelson, U.S. Census Bureau

Amy Newman-Smith, U.S. Census Bureau

Joy Sharp, Bureau of Transportation Statistics

Julie Weeks, National Center for Health Statistics

Andrew White, National Center for Education Statistics

Andrew Zukerberg, National Center for Education Statistics

CONFERENCE COORDINATORS

Anna Holaus, U.S. Census Bureau

Stacia Henderson Martin, U.S. Census Bureau

Lowanda Rivers, U.S. Census Bureau

Elaine Russell, U.S. Census Bureau

FEDERAL COMMITTEE ON STATISTICAL METHODOLOGY

Brian Harris-Kojetin (Chair), Office of Management and Budget

Nancy Bates, U.S. Census Bureau

Jonaki Bose, Substance Abuse and Mental Health Services Administration

Chris Chapman (CDAC Chair) Bureau of Labor Statistics

Christopher Chapman, National Center for Education Statistics

Steven B. Cohen, Agency for Healthcare Research and Quality

Stephen H. Cohen, National Center for Science and Engineering Statistics

John Eltinge, Bureau of Labor Statistics

Dennis Fixler, Bureau of Economic Analysis

Barry Graubard, National Cancer Institute

Deborah Griffin, U.S. Census Bureau

Arthur B. Kennickell, Federal Reserve Board

Jennifer Madans, National Center for Health Statistics

Rochelle Martinez, Office of Management and Budget

Renee Miller, Energy Information Administration

Jeri Mulrow, National Center for Science and Engineering Statistics

Grace O'Neill (Secretary), Substance Abuse and Mental Health Services Administration

William Sabol, Bureau of Justice Statistics

Nathaniel Schenker, National Center for Health Statistics

Rolf Schmitt, Bureau of Transportation Statistics

Marilyn McMillen Seastrom, National Center for Education Statistics

G. David Williamson, Agency for Toxic Substances and Disease Registry

Gordon Willis, National Institutes of Health

COUNCIL OF PROFESSIONAL ASSOCIATIONS ON FEDERAL STATISTICS (COPAFS)

Katherine "Kitty" Smith, Executive Director

Mae Pattison, Executive Assistant

2013 FCSM Research Conference

The 2013 Federal Committee on Statistical Methodology (FCSM) Research Conference was initiated by the FCSM. The FCSM is an interagency committee dedicated to improving the quality of federal statistics. The committee's major goals are to:

- Communicate and disseminate information on statistical practice among all federal statistical agencies.
- Recommend the introduction of new methodologies in federal statistical programs to improve data quality.
- Provide a mechanism for statisticians in different federal agencies to meet and exchange ideas.

The 2013 FCSM Research Conference provides a forum for experts from around the world to discuss and exchange current research and methodological topics relevant to federal government statistical programs. Each day of the conference will offer papers on a wide range of topics including: adaptive survey design, address-based sampling, administrative records, Bayesian statistical methods, coding, confidentiality and disclosure, crowdsourcing, data quality, economic statistics, editing, frame development, imputation, mobile devices, multi-mode and web data collection, nonresponse and measurement error, paradata and metadata, questionnaire and survey design, record linkage, small area estimation, survey redesign, variance estimation, and weighting.

Technical demonstrations will run concurrently on the second day of the conference during the first morning session. This session will include demonstrations on "Advances in Utilizing Maps and Visualizing Data."

Sessions feature papers and demonstrations by government, private sector, and academic researchers from five countries. All paper sessions will include an open discussion and some sessions will include a formal discussion. Papers will be made available on the FCSM Web site (http://www.fcsm.gov/) in March 2014.

Final Program

Federal Committee on Statistical Methodology Research Conference

Washington, DC – November 4-6, 2013

Monday, November 4

7:30 a.m.
Registration
(Concourse)

Continental Breakfast
(Concourse)

8:30 - 9:30 a.m.
Welcoming Remarks and PLENARY SESSION
(Rooms 207A and 207B)

9:30 - 10:00 a.m.
Break
(Concourse)

10:00 - 11:30 a.m.
CONCURRENT SESSION
A-1: Room 206
A-2: Room 207A
A-3: Room 207B
A-4: Room 202B

11:30 a.m. - 1:00 p.m.
Lunch on Your Own

1:00 - 2:30 p.m.
CONCURRENT SESSION
B-1: Room 206
B-2: Room 207A
B-3: Room 207B
B-4: Room 202B

2:30 - 2:45 p.m.
Break
(Concourse)

2:45 - 4:15 p.m.
CONCURRENT SESSION
C-1: Room 206
C-2: Room 207A
C-3: Room 207B
C-4: Room 202B

Tuesday, November 5

7:30 a.m.
Registration
(Concourse)

Continental Breakfast
(Concourse)

8:30 - 10:00 a.m.
CONCURRENT SESSION
D-1: Room 206
D-2: Room 207A
D-3: Room 207B
D-4: Room 202B

10:00 - 10:15 a.m.
Break
(Concourse)

10:15 - 11:45 a.m.
CONCURRENT SESSION
E-1: Room 206
E-2: Room 207A
E-3: Room 207B
E-4: Room 202B

11:45 a.m. - 1:15 p.m.
Lunch on Your Own

1:15 - 2:45 p.m.
CONCURRENT SESSION
F-1: Room 206
F-2: Room 207A
F-3: Room 207B
F-4: Room 202B

2:45 - 3:00 p.m.
Break
(Concourse)

3:00 - 4:30 p.m.
CONCURRENT SESSION
G-1: Room 206
G-2: Room 207A
G-3: Room 207B
G-4 Room 202B

Wednesday, November 6

7:30 a.m.
Registration
(Concourse)

Continental Breakfast
(Concourse)

8:30 - 10:00 a.m.
CONCURRENT SESSION
H-1: Room 206
H-2: Room 207A
H-3: Room 207B
H-4: Room 202B

10:00 - 10:15 a.m.
Break
(Concourse)

10:15 - 11:45 a.m.
CONCURRENT SESSION
I-1: Room 206
I-2: Room 207A
I-3: Room 207B
I-4: Room 202B

11:45 a.m. - 1:15 p.m.
Lunch on Your Own

1:15 - 2:45 p.m.
CONCURRENT SESSION
J-1: Room 206
J-2: Room 207A
J-3: Room 207B
J-4 Room 202B

Meeting Rooms:
All rooms located on the Mt. Vernon Place side of the Convention Center on the second floor.

Medical Assistance available in Room 208B.

MONDAY, NOVEMBER 4

7:30 a.m. - 4:15 p.m.
Registration
(Concourse)

7:30 - 8:30 a.m.
Continental Breakfast
(Concourse)

8:30 - 8:40 a.m.
Welcoming Remarks
Location: Rooms 207A and 207B

8:40 - 9:30 a.m.
PLENARY SESSION
Location: Rooms 207A and 207B

Confronting the Challenges of Household Surveys by Mixing Modes
Roger Tourangeau (Westat)

9:30 - 10:00 a.m.
Break
(Concourse)

10:00 - 11:30 a.m.
CONCURRENT SESSION A-1:
CONSIDERATIONS WHEN ADOPTING ADDRESS-BASED SAMPLING FOR FEDERAL SURVEYS

Location: Room 206
Organizer and Chair: Jonaki Bose (Substance Abuse and Mental Health Services Administration)

On Moving to the Use of a Hybrid Sampling Frame in the National Survey on Drug Use and Health: Motivations and Challenges
Arthur Hughes (Substance Abuse and Mental Health Services Administration)
Jonaki Bose (Substance Abuse and Mental Health Services Administration)
Bonnie Shook-Sa (RTI International)
Katherine Morton (RTI International)

Current and Future Uses of Address Lists for the National Health Interview Survey Sample
Chris Moriarity (National Center for Health Statistics)

A Tale of Two Surveys: Learning From the Application of Address-Based Sampling in Federal Surveys
Eileen O'Brien (Energy Information Administration)

Calculating Eligibility Rates in an Address-Based National Household Survey
Cameron McPhee (American Institutes for Research)
Amber Noel (American Institutes for Research)
Danielle Battle (American Institutes for Research)

Discussant: Clifford Loudermilk (U.S. Census Bureau)

10:00 - 11:30 a.m.
CONCURRENT SESSION A-2:
DEVELOPING SMALL AREA ESTIMATION MODELS

Location: Room 207A
Organizer: Kim Henry (Internal Revenue Service)
Chair: Tracy Haines (Internal Revenue Service)

Comparison of State-level Estimates of Adoption of Electronic Medical/Health Records Systems by Using Area-level Models and Direct Estimates
Vladislav Beresovsky (National Center for Health Statistics)
Janey Hsiao (National Center for Health Statistics)

Statistical Modeling of National Survey of College Graduates and American Community Survey Variables
Michael Larsen (The George Washington University)

Variance Modeling Research for the Small Area Health Insurance Estimates Program
Mark Bauder (U.S. Census Bureau)
Sam Szelepka (U.S. Census Bureau)
Donald Luery (U.S. Census Bureau)

Progress in Developing Small Area Estimates of Crime Based on the National Crime Victimization Survey
Robert Fay (Westat)
Mamadou Diallo (Westat)
Michael Planty (Bureau of Justice Statistics)

10:00 - 11:30 a.m.

CONCURRENT SESSION A-3: DESIGN AND TESTING STRATEGIES FOR WEB SURVEYS AND WEBSITES

Location: Room 207B
Organizer and Chair: Lauren Harris-Kojetin (National Center for Health Statistics)

The Impact of Survey Communications on Response Rates and Response Quality
Frances Barlas (ICF International)
Amy Falcone (ICF International)
Nikki Bellamy (Substance Abuse and Mental Health Services Administration)
Amy Mack (ICF International)

Developing an Inclusive Web Survey Design for Respondents with Disabilities
Jessica Jagger (U.S. Marine Corps)
Ashley Schaad (ICF International)
Ashleigh Davis (ICF International)
Amy Falcone (ICF International)

Improving Government Websites and Surveys with Usability Testing: A Comparison of Methodologies
Jennifer Romano Bergstrom (Fors Marsh Group, LLC)
Jonathan Strohl (Fors Marsh Group, LLC)

Discussant: Jean Fox (Bureau of Labor Statistics)

10:00 - 11:30 a.m.

CONCURRENT SESSION A-4: DATA QUALITY: ADMINISTRATIVE RECORDS VERSUS CENSUS AND SURVEY DATA

Location: Room 202B
Organizer and Chair: David Kashihara (Agency for Healthcare Research and Quality)

Evaluating Race and Hispanic Origin in Administrative Records Relative to the 2010 Census
Sonya Rastogi (U.S. Census Bureau)
James Noon (U.S. Census Bureau)
Amy O'Hara (U.S. Census Bureau)
Ellen Zapata (U.S. Census Bureau)

2010 American Community Survey Match Study: Comparing Administrative Records and ACS Data
Adela Luque (U.S. Census Bureau)
Renuka Bhaskar (U.S. Census Bureau)

Coverage and Agreement of Administrative Records and 2010 ACS Demographic Data
Renuka Bhaskar (U.S. Census Bureau)
Adela Luque (U.S. Census Bureau)
James Noon (U.S. Census Bureau)
Sonya Rastogi (U.S. Census Bureau)

Medicaid Undercount in the American Community Survey
Kathleen Call (University of Minnesota – SHADAC)
Brett O'Hara (U.S. Census Bureau)
Joanna Turner (University of Minnesota – SHADAC)
Michel Boudreaux (University of Minnesota – SHADAC)
Brett Fried (University of Minnesota – SHADAC)

The Medicaid Undercount in the 2007-2009 Current Population Survey
James Noon (U.S. Census Bureau)
Leticia Fernandez (U.S. Census Bureau)
Sonya Rastogi (U.S. Census Bureau)

11:30 a.m. - 1:00 p.m.
Lunch on Your Own

1:00 - 2:30 p.m.

CONCURRENT SESSION B-1: BUILDING BLOCKS OF ADAPTIVE SURVEY DESIGN

Location: Room 206
Organizer: Peter Miller (U.S. Census Bureau)
Chair: Gianna Dusch (U.S. Census Bureau)

How Much Is Enough? Moving Toward Smart Stopping Rules for Data Collection
Wan-Ying Chang (National Center for Science and Engineering Statistics)
Lynn Milan (National Center for Science and Engineering Statistics)

Planning an Adaptive Design Treatment in 2020 Census Tests
Gina Walejko (U.S. Census Bureau)
Peter Miller (U.S. Census Bureau)
Gianna Dusch (U.S. Census Bureau)
Kevin Deardorff (U.S. Census Bureau)

Adaptive Curtailment of Survey Followup Based on Contact History Data
Eric Slud (U.S. Census Bureau)
Chandra Erdman (U.S. Census Bureau)

Using Substantive Variables and Paradata to Inform Responsive Design Models: Results From a Nationally Representative Longitudinal Study
Ted Socha (National Center for Education Statistics)

Adaptive Design Strategies: Who Do You Target? And What Do You Do With Them?
Jaki McCarthy (National Agricultural Statistics Service)
Kathy Ott (National Agricultural Statistics Service)
Melissa Mitchell (National Agricultural Statistics Service)

1:00 - 2:30 p.m.

CONCURRENT SESSION B-2: SOLVING DATA PROBLEMS WITH BAYESIAN METHODS: MODELING, COMPLIANCE CONTROL, AND CLIMATE CHANGE

Location: Room 207A
Organizer and Chair: Andrew White (National Center for Education Statistics)

Sample Design and Analysis for Assessing Compliance Control Settings
Bhojnarine Rambharat (Office of the Comptroller of the Currency)

Log Linear Mixed With Additive Random Components Modeling for Small Area Estimation: A Bayesian-Frequentist Integrated Approach
Avinash Singh (NORC at the University of Chicago)

Spatio-Temporal Modeling of the Impact of Climate Change on Road Traffic Accidents – A Case Study of New Brunswick
Md Shohel Reza Amin (Concordia University, Canada)
Alireza Zareie (McGill University, Canada)
Luis Amador-Jiménez (Concordia University, Canada)

Weight Smoothing With Laplace Prior and Its Application in GLM Model
Xi Xia (University of Michigan)
Michael Elliott (University of Michigan)

1:00 - 2:30 p.m.

CONCURRENT SESSION B-3: ECONOMIC STATISTICS

Location: Room 207B
Organizer and Chair: Benjamin Bridgman (Bureau of Economic Analysis)

Fair Value Accounting and Measures of Corporate Profits in the U.S. National Accounts
Dylan Rassier (Bureau of Economic Analysis)

Forecasting Retail Food Expenditures for American Households
Richard Volpe (Economic Research Service)
Annemarie Kuhns (Economic Research Service)
Timothy Park (Economic Research Service)

Supplemental Poverty Measure: A Comparison of Geographic Adjustments With Regional Price Parities vs. Median Rents from the American Community Survey
Trudi Renwick (U.S. Census Bureau)
Bettina Aten (Bureau of Economic Analysis)
Troy Martin (Bureau of Economic Analysis)

New Firm Size JOLTS Data Provide Insight to the U.S. Labor Market
Charlotte Oslund (Bureau of Labor Statistics)

1:00 - 2:30 p.m.

CONCURRENT SESSION B-4: ON THE STATISTICAL VALIDITY OF INTERNET OPT-IN PANEL SURVEYS

Location: Room 202B
Organizer and Chair: Satvinder Dhingra (Northrop Grumman)

Investigating Novel Approaches for State-Based Behavioral Risk Surveillance
Carol Gotway Crawford (Centers for Disease Control)
Catherine Okoro (Centers for Disease Control)
Satvinder Dhingra (Northrop Grumman)

Sources of Comparability Between Probability Sample Estimates and Nonprobability Web Sample Estimates
William Riley (National Cancer Institute)
Bob Kaplan (National Institutes of Health)
David Cella (Northwestern University of Feinberg School of Medicine)
Catherine Okoro (Centers for Disease Control)
Satvinder Dhingra (Northrop Grumman and Centers for Disease Control)

Testing Ignorable Selection Across Multiple Domains: Economic, Politics, and Health
Stephen Ansolabehere (Harvard University)
Douglas Rivers (Stanford University and YouGov)

Discussant: Sandro Galea (Columbia University)

2:30 - 2:45 p.m.
Break
(Concourse)

2:45 - 4:15 p.m.

CONCURRENT SESSION C-1: SUPPLEMENTING THE COVERAGE OF ADDRESS-BASED SAMPLING FRAMES FOR IN-PERSON HOUSEHOLD SURVEYS

Location: Room 206
Organizer and Chair: Rachel Harter (RTI International)

Surveying Young Adults Using an Address List-Based Frame
Luciano Viera, Jr. (Fors Marsh Group, LLC)
Ricardo Carvalho (Fors Marsh Group, LLC)
Sean Marsh (Fors Marsh Group, LLC)

Correcting Coverage Deficiencies in Address-Based Frames: The Use of Enhanced Listing
Ned English (NORC at the University of Chicago)
Colm O'Muircheartaigh (Irving B. Harris Graduate School of Public Policy at the University of Chicago)
Katie Dekker (NORC at the University of Chicago)

Using an Area Linkage Method to Improve the Coverage of ABS Frames for In-Person Household Surveys
Sylvia Dohrmann (Westat)
Richard Sigman (Westat)

The CHUM: An Adaptation of the Half-Open Interval Procedure for Use with ABS Frames
Joe McMichael (RTI International)
Bonnie Shook-Sa (RTI International)
Jamie Ridenhour (RTI International)

Discussant: Allen McCutcheon (Gallup Research Center, University of Nebraska-Lincoln)

2:45 - 4:15 p.m.

CONCURRENT SESSION C-2: PRODUCTION AND EVALUATION OF SUBNATIONAL ESTIMATES

Location: Room 207A
Organizer: Michael Planty (Bureau of Justice Statistics)
Chair: William Sabol (Bureau of Justice Statistics)

Small Area Estimation: New Developments and Directions for Health and Human Services Data
John Czajka (Mathematica Policy Research)
Amang Sukasih (Mathematica Policy Research)
Susan Queen (Department of Health and Human Services)

Challenges of Using Prediction Models to Produce Nationally Representative Estimates of Mental Illness
Sarra Hedden (Substance Abuse and Mental Health Services Administration)
Joseph Gfroerer (Substance Abuse and Mental Health Services Administration)
Jonaki Bose (Substance Abuse and Mental Health Services Administration)
Phillip Kott (RTI International)
Dan Liao (RTI International)
Lisa Colpe (The National Institute of Mental Health)

Small Area Modeling of County Estimates for Corn and Soybean Yields in the U.S.
Matthew Williams (National Agricultural Statistics Service)

A Comparison of Small Area Models Used in the Quality Indicator Program Sponsored by Agency for Healthcare Research and Quality
Robert Baskin (Agency for Healthcare Research and Quality)
Pamela Owens (Agency for Healthcare Research and Quality)
Christopher Sroka (Battelle)

2:45 - 4:15 p.m.

CONCURRENT SESSION C-3: NOVEL APPROACHES TO CODING, EDITING, AND IMPUTING DATA

Location: Room 207B
Organizer and Chair: Jeffrey Gonzalez (Agency for Healthcare Research and Quality)

Issues in Coding Hospital Intensive Care for Data Analysis
Adeline Wilcox (Department of Veterans Affairs)

Statistical Analysis of Text in Survey Records
Wendy Martinez (Bureau of Labor Statistics)
Alex Measure (Bureau of Labor Statistics)

Evaluation of a New Edit Methodology for the Common Core of Data Nonfiscal Surveys
Elizabeth Goldberg (U.S. Census Bureau)
Robert Stillwell (National Center for Education Statistics)
Jeffrey Little (U.S. Census Bureau)

Banff Automated Edit and Imputation Applied to the U.S. Hog Inventory Survey
James Johanson (National Agricultural Statistics Service)

Boosting Algorithms for Edit and Imputation of Multiple-response Variables
Ping Li (Cornell University)
John Abowd (Cornell University)

2:45 - 4:15 p.m.

CONCURRENT SESSION C-4:
EVALUATING THE IMPACTS OF MODE SELECTION IN SURVEYS

Location: Room 202B
Organizer and Chair: Louise Camalier (U.S. Environmental Protection Agency)

An Analysis of the Mixed Collection Modes for Business Surveys at the U.S. Census Bureau
Broderick Oliver (U.S. Census Bureau)
Katherine Jenny Thompson (U.S. Census Bureau)

Truth or Consequences: The Use of Benchmarks in Calibrating Health Research
Victor Lange (Marketing, Inc.)
Steven Gittelman (Marketing, Inc.)
Elaine Trimarchi (Marketing, Inc.)

Mode Selection and Recruitment Strategies for Survey Respondents with Disabilities
Ashley Schaad (ICF International)
Amy Falcone (ICF International)
Jessica Jagger (U.S. Marine Corps)
Ashleigh Davis (ICF International)

Great Expectations: Changing Mode of Survey Data Collection in Military Populations
Ronald Szoc (ICF International)
Jacqueline Pflieger (ICF International)
Frances Barlas (ICF International)
Randall Thomas (GfK Custom Research, LLC)

TUESDAY, NOVEMBER 5

7:30 a.m. – 4:30 p.m.
Registration
(Concourse)

7:30 - 8:30 a.m.
Continental Breakfast
(Concourse)

8:30 - 10:00 a.m.

CONCURRENT SESSION D-1:
COMPARISONS OF VARIANCE ESTIMATION TECHNIQUES

Location: Room 206
Organizer and Chair: David Kashihara (Agency for Healthcare Research and Quality)

Investigation of Variance Estimators for the Survey of Business Owners (SBO)
Marilyn Balogh (U.S. Census Bureau)
Sandy Peterson (U.S. Census Bureau)

Inference of Domain Parameters by Using a Dynamic Jackknife Variance Estimator
Sixia Chen (Westat)
Tom Krenzke (Westat)

Using Reimputation Methods to Estimate the Variances of Estimates of the ACS Group Quarters Population with the New Group Quarters Imputation Methodology
Michael Beaghen (U.S. Census Bureau)

Comparing Generalized Variance Functions to Direct Variance Estimation for the National Crime Victimization Survey
Bonnie Shook-Sa (RTI International)
Lance Couzens (RTI International)
Rick Williams (RTI International)
Marcus Berzofsky (RTI International)

8:30 - 10:00 a.m.

CONCURRENT SESSION D-2:
IMPROVING SURVEY DESIGN AND ESTIMATES USING EXTERNAL SOURCES

Location: Room 207A
Organizer: Preston McDowney (Energy Information Administration)
Chair: Marilyn Worthy (Energy Information Administration)

Analysis of Housing Square Footage Estimates Reported by the AHS and SOC
Angela Delano (U.S. Census Bureau)

Household Estimates Conundrum: Effort to Develop More Consistent Household Estimates Across Current Surveys
Arthur Cresce Jr. (U.S. Census Bureau)
Yang Cheng (U.S. Census Bureau)
Andrew Zbikowski (U.S. Census Bureau)
Christopher Grieves (U.S. Census Bureau)

Estimating Substance Abuse Treatment: A Comparison of Data from a Household Survey, a Facility Survey, and an Administrative Data Set
Joe Gfroerer (Substance Abuse and Mental Health Services Administration)
Jonaki Bose (Substance Abuse and Mental Health Services Administration)
Deborah Trunzo (Substance Abuse and Mental Health Services Administration)
Alex Strashny (Substance Abuse and Mental Health Services Administration)
Kathy Batts (RTI International)
Michael Pemberton (RTI International)

A Comparison of Consumer Expenditure Surveys
Nhien To (Bureau of Labor Statistics)
Brett McBride (Bureau of Labor Statistics)

Comparing BLS's Price Indexes with Industry Sources
Rachel Soloveichik (Bureau of Economic Analysis)
Paul Sullivan (Bureau of Labor Statistics)

8:30 - 10:00 a.m.

CONCURRENT SESSION D-3:
RECENT FINDINGS AND INNOVATIONS IN UTILIZING PARADATA AND METADATA

Location: Room 207B
Organizer and Chair: Karen Hamrick (Economic Research Service)

Metadata Standards and Technology Development for the NSF Survey of Earned Doctorates
Kimberly Noonan (National Center for Science and Engineering Statistics)
Pascal Heus (Metadata Technology North America)
Tim Mulcahy (NORC at the University of Chicago)

Using Paradata to Model Nonresponse in the Current Population Survey
John Dixon (Bureau of Labor Statistics)

Using Paradata to Understand Business Survey Reporting Patterns
Eric Fink (U.S. Census Bureau)
Joanna Fane Lineback (U.S. Census Bureau)

What Can Paradata Tell Us About Business Reporting in the 2012 Economic Census?
Elizabeth Hoeffel (Census Bureau)
Gianna Dusch (U.S. Census Bureau)
Joanna Fane Lineback (U.S. Census Bureau)

8:30 - 10:00 a.m.

CONCURRENT SESSION D-4:
TECHNICAL DEMONSTRATIONS
ADVANCES IN UTILIZING MAPS AND VISUALIZING DATA

Location: Room 202B
Organizer: Wendy Barboza (National Agricultural Statistics Service)
Chair: Jaki McCarthy (National Agricultural Statistics Service)

A Demonstration Occupational Employment Statistics' Web-Based Chart and Map Visualization Tools
Tanner Beam (Bureau of Labor Statistics)

Visualizing Historical Agricultural Data: The Current State of the Art
Irwin Anolik (National Agricultural Statistics Service)

Hurricane Category Establishment Identification Project (HCat)
Peter Smith (Bureau of Labor Statistics)

Map Matching and Real World Integrated Sensor Data Warehousing
Evan Burton (National Renewable Energy Lab)
Jeff Gonder (National Renewable Energy Lab)
Adam Duran (National Renewable Energy Lab)
Eric Wood (National Renewable Energy Lab)

10:00 - 10:15 a.m.
Break
(Concourse)

10:15 - 11:45 a.m.

CONCURRENT SESSION E-1:
WEIGHTING METHODS IN THE PRESENCE OF NONRESPONSE

Location: Room 206
Organizer: Kim Henry (Internal Revenue Service)
Chair: Nicholas Mountjoy (Internal Revenue Service)

Evaluation Study of Calibration Estimation for the Survey of Local Government Finance
Elizabeth Love (U.S. Census Bureau)
Bac Tran (U.S. Census Bureau)

Calibration Methods in the Quarterly Summary of State and Local Tax Revenue
Justin Nguyen (U.S. Census Bureau)
Bac Tran (U.S. Census Bureau)

Performance of Generalized Regression Estimator and Raking Estimator in the Presence of Nonresponse
Daifeng Han (Westat and University of Maryland College Park)
Richard Valliant (University of Maryland)
Jill Montaquila (Westat)
Keith Rust (Westat)

Patterns of Biomarker Participation in L.A. Family and Neighborhood Survey
Bonnie Ghosh-Dastidar (The RAND Corporation)
Narayan Sastry (University of Michigan)

10:15 - 11:45 a.m.

CONCURRENT SESSION E-2: ADMINISTRATIVE RECORDS AND SURVEY RESEARCH

Location: Room 207A
Organizer and Chair: Sonya Rastogi (U.S. Census Bureau)

The Nature of the Bias When Studying Only Linkable Person Records: Evidence from the American Community Survey
J. David Brown (U.S. Census Bureau)
Adela Luque (U.S. Census Bureau)
Amy O'Hara (U.S. Census Bureau)

Social Security Income Measurement in Two Surveys
Howard Iams (Social Security Administration)
Patrick Purcell (Social Security Administration)

Changes in EITC Eligibility and Participation, 2005-2009
Maggie Jones (U.S. Census Bureau)

Profile of Supplemental Nutrition Assistance Program (SNAP) Usage at the State and County Levels: Evidence from Texas and New York SNAP Administrative Records and the American Community Survey
Benjamin Harris (U.S. Census Bureau)
Erik Scherpf (Economic Research Service)
Constance Newman (Economic Research Service)
Graton Gathright (U.S. Census Bureau)

10:15 - 11:45 a.m.

CONCURRENT SESSION E-3: DATA DISCLOSURE ISSUES

Location: Room 207B
Organizer: Michael Planty (Bureau of Justice Statistics)
Chair: Daniela Golinelli (Bureau of Justice Statistics)

A Disclosure Avoidance Research Agenda
Paul Massell (U.S. Census Bureau)

Evaluating Noise Infusion for Disclosure Protection for Two Time Periods
Jared Martin (U.S. Census Bureau)

Is an On-Line Microdata Tool That Uses American Community Survey Transportation Data Feasible?
Tom Krenzke (Westat)
Penelope Weinberger (American Association of State Highway and Transportation Officials)
Jianzhu Li (Westat)
Elaine Murakami (Federal Highway Administration)

The Census Bureau's New Cell Suppression System
Philip Steel (U.S. Census Bureau)

Managing Confidentiality and Provenance Across Mixed Private and Publicly-Accessed Data and Metadata
Carl Lagoze (University of Michigan)
William Block (Cornell Institute of Social and Economic Research)
Jeremy Williams (Cornell Institute of Social and Economic Research)
John Abowd (Cornell University)
Lars Vilhuber (Cornell University)

10:15 - 11:45 a.m.

CONCURRENT SESSION E-4: MEASURING, MONITORING, AND ASSESSING THE QUALITY OF SURVEY DATA

Location: Room 202B
Organizer: Jeffrey Gonzalez (Agency for Healthcare Research and Quality)
Chair: Scott Fricker (Bureau of Labor Statistics)

The Quality Assurance Reviews at Statistics Canada
Laurie Reedman (Statistics Canada)
Claude Julien (Statistics Canada)

Distribution of Quality Control to the Point of Data Collection in the Field: Impacts on Cost, User Experience, and Security
Andrea Johnson (U.S. Census Bureau)
Sandy Dyer (U.S. Census Bureau)
Michael Ratcliffe (U.S. Census Bureau)
Jonathan Krentel (Gunnison Consulting Group, Inc.)

An Approach to Measurement Error Assessment for a Household Expenditure Survey: A Review and Application
Brandon Kopp (Bureau of Labor Statistics)
Roger Tourangeau (Westat)
Lucilla Tan (Bureau of Labor Statistics)
Scott Fricker (Bureau of Labor Statistics)

A Model for Matching Error for Triple System Estimation
Richard Griffin (U.S. Census Bureau)

11:45 a.m. - 1:15 p.m.
Lunch on Your Own

1:15 - 2:45 p.m.

CONCURRENT SESSION F-1: EVALUATING WEIGHTING AND VARIANCE ESTMATION METHODS

Location: Room 206
Organizer: Preston McDowney (Energy Information Administration)
Chair: Carrie Hughes-Cromwick (Energy Information Administration)

Effects of Poststratification and Raking Adjustments on Precision of MEPS Estimates
Sadeq Chowdhury (Agency for Healthcare Research and Quality)

Model-Assisted Domain Estimation
Dan Liao (RTI International)
Phillip Kott (RTI International)

Variance Estimation for Calibration to Estimated Totals
Siyu Qing (The George Washington University)
Michael Larsen (The George Washington University)

Impact of Influential Observations on Enumeration and Variance Estimation in the National Crime Victimization Survey
Michael Planty (Bureau of Justice Statistics)
Lynn Langton (Bureau of Justice Statistics)

1:15 - 2:45 p.m.

CONCURRENT SESSION F-2: INTERNATIONAL ECONOMIC STATISTICS

Location: Room 207A
Organizer: Benjamin Bridgman (Bureau of Economic Analysis)
Chair: Tani Fukui (U.S. International Trade Commission)

Estimating Foreign-Born Emigration from the United States Using Data from the American Community Survey
Mark Leach (U.S. Census Bureau)

The Use of Theory and Model Averaging for Population Prediction: An Application to the U.S. Overseas Population
Sidney Carl Turner (Fors Marsh Group, LLC)
Joseph Luchman (Fors Marsh Group, LLC)
Andrew Therriault (Lightbox Analytics)
Brian Griepentrog (Fors Marsh Group, LLC)
Kinsey Gimbel (Fors Marsh Group, LLC)
Fritz Scheuren (University of Chicago)
Ali Mushtaq (Consultant)
Paul Drugan (Federal Voting Assistance Program)

Effects of the Eurozone Economy on Economic Indicators
Tracey Kyckelhahn (Bureau of Justice Statistics)

Offshore Outsourcing, Portfolio of Intangible Capital, and Performance: Theory and Evidence in the IT and Pharmaceutical Industries
Wendy Li (Bureau of Economic Analysis)

Benchmarking Monthly Indicator Series with Quarterly Information Derived From Administrative Data
Pieter Vlag (Statistics Netherlands)

1:15 - 2:45 p.m.

CONCURRENT SESSION F-3:
ASSESSING THE EFFECTIVENESS OF STRATEGIES TO IMPROVE SURVEY RESPONSE RATES

Location: Room 207B
Organizer and Chair: Lynn Langton (Bureau of Justice Statistics)

Testing Contact and Response Strategies to Improve Response in the 2012 Economic Census
Erica Marquette (U.S. Census Bureau)
Michael Kornbau (U.S. Census Bureau)

Use of Smart Phones/Text Messaging to Increase Response Rates
Piper DuBray (ICF International)
Naomi Freedner (ICF International)
Kisha Bailly (ICF International)
Kristie Healey (ICF International)

Do Names Matter? Experiments Comparing Different Branding and Levels of Personally Identifiable Information in a Mail Questionnaire
Sarah Hastedt Carroll (National Center for Education Statistics)
Andrew Zukerberg (National Center for Education Statistics)

Time Use, Response Rates, and Data Quality by Time of Day
Rose Woods (Bureau of Labor Statistics)
Laura Wronski (Bureau of Labor Statistics)

1:15 - 2:45 p.m.

CONCURRENT SESSION F-4:
ADAPTIVE APPROACHES TO SAMPLING AND SUBSAMPLING

Location: Room 202B
Organizer and Chair: Dawn Nelson (U.S. Census Bureau)

Adaptive Design Features for Using Address-Based Sampling in a National CATI Survey of Households with Children
Charles DiSogra (Abt SRBI)
David Finkelhor (University of New Hampshire)
Heather Hammer (Abt SRBI)
Stas Kolenikov (Abt SRBI)
Heather Turner (University of New Hampshire)

Adaptive Adjustment of the Multicriteria Optimal Allocation of a Hard to Reach Population Survey
Benjamin Phillips (Abt SRBI)
Stas Kolenikov (Abt SRBI)

Optimizing Unit Nonresponse Adjustment Procedures After Subsampling Nonrespondents in the Economic Census
Laura Bechtel (U.S. Census Bureau)
Katherine Jenny Thompson (U.S. Census Bureau)

Determining Allocation Requirements for Subsampling Nonrespondents From the Annual Survey of Manufactures
Daniel Whitehead (U.S. Census Bureau)
Stephen Kaputa (U.S. Census Bureau)
Katherine Jenny Thompson (U.S. Census Bureau)

2:45 - 3:00 p.m.
Break
(Concourse)

3:00 - 4:30 p.m.

CONCURRENT SESSION G-1: ISSUES ASSOCIATED WITH USING ADMINISTRATIVE RECORDS AND LINKAGE

Location: Room 206
Organizer and Chair: Julie Weeks (National Center for Health Statistics)

Assessment of Quality, Cost and Risk Factors in Statistical Work with Administrative Record Data: A Review
John Eltinge (Bureau of Labor Statistics)

Riddles Wrapped in Mysteries Inside Enigmas: Issues with Getting and Using Administrative Data for Impact Evaluations
Alisha Creel (ICF International)
Ronald Szoc (ICF International)

Administrative Data Initiatives at Statistics Canada
Julie Trépanier (Statistics Canada)
Jean Pignal (Statistics Canada)
Don Royce (Statistics Canada)

Discussant: Dean Judson (National Center for Health Statistics)

3:00 - 4:30 p.m.

CONCURRENT SESSION G-2: ASSESSING ALTERNATIVE WEIGHTING METHODOLOGIES

Location: Room 207A
Organizer: Wendy Barboza (National Agricultural Statistics Service)
Chair: Daniel Adrian (National Agricultural Statistics Service)

Performance of Weighted and Non-weighted Estimators in a Cell Phone Based Electoral Poll: An Academic Study of the 2012 Presidential Elections in Mexico
Olivia Carrillo-Gamboa (Tecnológico de Monterrey, Mexico)
Rosa Isela Hernández-Zamora (Tecnológico de Monterrey, Mexico)
Jesús Cantú-Escalante (Tecnológico de Monterrey, Mexico)

The Use of Signal Filtering for Hog Inventory Estimation
Stephen Busselberg (National Agricultural Statistics Service)

Depicting Time-Dependent Changes in Environmental Data for Prospective Uses in Personalized Medicine
Turkan Gardenier (Pragmatica Corp)

Prediction Performance of Single Index Principal Fitted Component Models
Jia-Ern Pai (National Highway Traffic Safety Administration)
Kofi Placid Adragni (University of Maryland, Baltimore County)

An Alternative Estimation Method for the Current Population Survey
Daniel Bonnéry (University of Maryland, College Park, and U.S. Census Bureau)
Partha Lahiri (University of Maryland, College Park)
Yang Cheng (U.S. Census Bureau)

3:00 - 4:30 p.m.

CONCURRENT SESSION G-3: HITTING THE TARGET IN HOSPITAL PROFILING: THE AHRQ QUALITY INDICATORS

Location: Room 207B
Organizer: Frank Yoon (Mathematica Policy Research)
Chair: Hali Hambridge (Mathematica Policy Research)

Variation in Quality by Hospital Characteristics: True or False?
Eric Schone (Mathematica Policy Research)
Alex Bohl (Mathematica Policy Research)
David Jones (Mathematica Policy Research)
Dmitriy Poznyak (Mathematica Policy Research)
Jessica Ross (Mathematica Policy Research)
Samantha Stalley (Mathematica Policy Research)
Frank Yoon (Mathematica Policy Research)
Joe Zickafoose (Mathematica Policy Research)

The Role of Hospital Characteristics in Setting Appropriate Yardsticks for Quality Measurement
Frank Yoon (Mathematica Policy Research)
Alex Bohl (Mathematica Policy Research)
David Jones (Mathematica Policy Research)
Dmitriy Poznyak (Mathematica Policy Research)
Jessica Ross (Mathematica Policy Research)
Eric Schone (Mathematica Policy Research)
Samantha Stalley (Mathematica Policy Research)
Joe Zickafoose (Mathematica Policy Research)

Hospital Peer Groups, Reliability, and Stabilization: Shrinking to the Right Mean
Alex Bohl (Mathematica Policy Research)
David Jones (Mathematica Policy Research)
Dmitriy Poznyak (Mathematica Policy Research)
Jessica Ross (Mathematica Policy Research)
Eric Schone (Mathematica Policy Research)
Samantha Stalley (Mathematica Policy Research)
Frank Yoon (Mathematica Policy Research)
Joe Zickafoose (Mathematica Policy Research)

Discussant: Pam Owens (Agency for Healthcare Research and Quality)

3:00 - 4:30 p.m.

CONCURRENT SESSION G-4:
NEW APPROACHES TO SURVEY DESIGN, IMPLEMENTATION, AND USE

Location: Room 202B
Organizer: Richard Chard (Social Security Administration)
Chair: Barbara Smith (Social Security Administration)

The Consumer Expenditure Survey's New Design and Implementation Plans
Laura Paszkiewicz (Bureau of Labor Statistics)
Jennifer Edgar (Bureau of Labor Statistics)
Brett McBride (Bureau of Labor Statistics)
Dawn Nelson (U.S. Census Bureau)
Adam Safir (Bureau of Labor Statistics)

Race and Ethnicity Measurement: Effects of Response Format
Randall Thomas (GfK Custom Research, LLC)
Frances Barlas (ICF International)
Wendy Gross (GfK Custom Research, LLC)

Backtranslation vs. Committee Approach: An Experiment Comparing How They Perform in Questionnaire Translation
Alisú Schoua-Glusberg (Research Support Services)
Ana Villar (City University, United Kingdom)

Getting the Most out of a Limited Sample Size Field Test: Experiences from the National Survey on Drug Use and Health
Jonaki Bose (Substance Abuse and Mental Health Services Administration)
Dicy Painter (Substance Abuse and Mental Health Services Administration)
Doug Currivan (RTI International)
Larry Kroutil (RTI International)
Kevin Wang (RTI International)

7:30 a.m. - 1:15 p.m.
Registration
(Concourse)

7:30 - 8:30 a.m.
Continental Breakfast
(Concourse)

8:30 - 10:00 a.m.

CONCURRENT SESSION H-1: CROWDSOURCING METHODOLOGIES THROUGHOUT THE SURVEY LIFECYCLE

Location: Room 206
Organizer: Michael Keating (RTI International)
Chair: Jennifer Edgar (Bureau of Labor Statistics)

Crowdsourcing in the Cognitive Interviewing Process
Joe Murphy (RTI International)
Michael Keating (RTI International)
Jennifer Edgar (Bureau of Labor Statistics)

Case Study Comparing Data Collected via Crowdsourcing vs. Trained Data Collectors
Annice Kim (RTI International)
Allie Lieberman (RTI International)
Daniel Dench (RTI International)

The Census Bureau Mail Return Rate Challenge: Crowdsourcing to Develop a Hard to Count Score
Chandra Erdman (U.S. Census Bureau)
Nancy Bates (U.S. Census Bureau)

A Methodological Framework for Crowdsourcing in Research
Michael Keating (RTI International)
Robert Furberg (RTI International)

8:30 - 10:00 a.m.

CONCURRENT SESSION H-2: LINKED DATA FOR ANALYSIS AND SURVEY EVALUATION

Location: Room 207A
Organizer and Chair: Karen Hamrick (Economic Research Service)

A Comparison of Person-Reported Industry to Employer-Reported Industry in Survey and Administrative Data
Emily Isenberg (U.S. Census Bureau)
Liana Christin Landivar (U.S. Census Bureau)
Esther Mezey (Axioma, Inc.)

The National Household Survey: Evaluating Data Quality in a Large Voluntary Survey
Sander Post (Statistics Canada)

Enhancing the Medical Expenditure Panel Survey Through Data Linkages
Lisa Mirel (Agency for Healthcare Research and Quality)
Steven Machlin (Agency for Healthcare Research and Quality)

Characteristics of Medicare Beneficiaries with Part D Coverage: NHANES 2003-2004 Linked to 2006 Medicare Part D Data
Ryne Paulose (National Center for Health Statistics)
Hannah Day (National Center for Health Statistics)
Steven Frenk (National Center for Health Statistics)

8:30 - 10:00 a.m.

CONCURRENT SESSION H-3: HEALTH INSURANCE AND HOUSEHOLD INCOME IN THE CURRENT POPULATION SURVEY (CPS) REDESIGN AND PRODUCTION

Location: Room 207B
Organizer: Brett O'Hara (U.S. Census Bureau)
Chair: Thesia Garner (Bureau of Labor Statistics)

The 2013 Annual Social Economic Supplement Income and Health Insurance Questionnaire Test Sample Design
David Hornick (U.S. Census Bureau)

Non-response Bias in the March 2013 Content Test
Matthew Brault (U.S. Census Bureau)

Evaluating the Income and Health Insurance Questions 2013 CPS ASEC Content Test Using Timer Data
Aaron Cantu (U.S. Census Bureau)
Adam Bee (U.S. Census Bureau)

Evaluating the 2013 CPS ASEC Income Redesign Content Test
Jessica Semega (U.S. Census Bureau)

Health Insurance in the Current Population Survey: Redesign and Production
Carla Medalia (U.S. Census Bureau)
Amy Steinweg (U.S. Census Bureau)
Brett O'Hara (U.S. Census Bureau)
David Lee (U.S. Census Bureau)
Jessica Smith (U.S. Census Bureau)
Joanne Pascale (U.S. Census Bureau)
Jonathan Rodean (U.S. Census Bureau)

8:30 - 10:00 a.m.

CONCURRENT SESSION H-4:
INNOVATIONS IN FRAME DEVELOPMENT

Location: Room 202B
Organizer and Chair: Andrew Zukerberg (National Center for Education Statistics)

Sample Design Considerations for the Occupational Requirements Survey
Brad Rhein (Bureau of Labor Statistics)
Chester Ponikowski (Bureau of Labor Statistics)
Erin McNulty (Bureau of Labor Statistics)

Exploratory Research on the Use of Google Earth to Create a Sampling Frame of Buildings
Katie Lewis (Energy Information Administration)

Surveying Community Stakeholders: Exploring the Viability of a National Sampling Frame
Barbara Robles (Board of Governors of the Federal Reserve System)

Redesigning National School Surveys: Coverage and Stratification Improvement Using Multiple Datasets
William Robb (ICF International)
Kate Flint (ICF International)
Ronaldo Iachan (ICF International)

10:00 - 10:15 a.m.
Break
(Concourse)

10:15 - 11:45 a.m.

CONCURRENT SESSION I-1:
MOBILE DEVICES AND APPLICATIONS FOR SURVEY DATA COLLECTION

Location: Room 206
Organizer: Joy Sharp (Bureau of Transportation Statistics)
Chair: Kenneth Steve (Bureau of Transportation Statistics)

Bring Your Own Device and the 2020 Census Research and Testing
Keith Evan Moffett (U.S. Census Bureau)
Jay Occhiogrosso (U.S. Census Bureau)
Scott Williams (U.S. Census Bureau)
Ryan King (U.S. Census Bureau)

Field Data Collection in Area Frame Surveys Utilizing iPads® - USDA's June Area Survey
Michael Gerling (National Agricultural Statistics Service)
Eric Wilson (National Agricultural Statistics Service)
Linda Lawson (National Agricultural Statistics Service)
Alan Dotts (Iowa State University)
Andrew Vardeman (Iowa State University)

The iPad® Computer-Assisted Personal Interview System - A Revolution for In-Person Data Capture?
Heather Driscoll (ICF International)
James Dayton (ICF International)

Innovative Retention Methods in Panel Research: Can SmartPhones Improve Long-term Panel Participation?
James Dayton (ICF International)
Andrew Dyer (ICF International)

10:15 - 11:45 a.m.

CONCURRENT SESSION I-2:
INTEGRATED APPROACHES TO DATA EDITING, IMPUTATION, AND DISCLOSURE CONTROL

Location: Room 207A
Organizer: Jerome Reiter (Duke University)
Chair: Alan Karr (National Institute of Statistical Science)

Simultaneous Edit-imputation for Continuous Microdata
Hang Joon Kim (Duke University and National Institute of Statistical Science)
Jerome Reiter (Duke University)
Alan Karr (National Institute of Statistical Science)
Larry Cox (National Institute of Statistical Science)

Simultaneous Edit-imputation for Categorical Microdata
Daniel Manrique-Vallier (Indiana University)
Jerome Reiter (Duke University)

Statistical Disclosure Limitation and Edit-imputation
Alan Karr (National Institute of Statistical Science)
Hang Joon Kim (Duke University and National Institute of Statistical Science)
Jerome Reiter (Duke University)
Larry Cox (National Institute of Statistical Science)

Discussant: William Winkler (U.S. Census Bureau)

10:15 - 11:45 a.m.

CONCURRENT SESSION I-3:
EXAMINING THE IMPACTS OF SAMPLE DESIGN: COST, BIAS, VARIANCE AND OTHER SURVEY CONCERNS

Location: Room 207B
Organizer: Amy Newman-Smith (U.S. Census Bureau)
Chair: Carma Hogue (U.S. Census Bureau)

Optimal Cutoff Sampling for the Annual Survey of Public Employment and Payroll
Brian Dumbacher (U.S. Census Bureau)
Carma Hogue (U.S. Census Bureau)

Numerical Impact of a Simple Random Subsample on Consumer Spending for Children
Daniel Yang (Bureau of Labor Statistics)
Jeffrey Gonzalez (Agency for Healthcare Research and Quality)

Effects of Rotation Group Bias on Estimation of Smoking Prevalence
Younghwan Song (Union College)

The Impact of Cellphone Sample Representation on Variance Estimates in a Dual-Frame Telephone Survey
A. Elizabeth Ormson (NORC at the University of Chicago)
Kennon Copeland (NORC at the University of Chicago)
Stephen Blumberg (National Center for Health Statistics)
Nadarajasundaram Ganesh (NORC at the University of Chicago)

10:15 - 11:45 a.m.

CONCURRENT SESSION I-4:
USE OF CAPTURE-RECAPTURE METHODS IN FEDERAL CENSUSES

Location: Room 202B
Organizer: Vincent Thomas Mule Jr. (U.S. Census Bureau)
Chair: Steven Klement (U.S. Census Bureau)

Reducing Nonsampling Error in the Census Coverage Measurement
Tamara Adams (U.S. Census Bureau)
Magdalena Ramos (U.S. Census Bureau)
Gia Donnalley (U.S. Census Bureau)

Accounting for Missing Data in the Census Coverage Measurement Survey
Vincent Thomas Mule Jr. (U.S. Census Bureau)
Scott Konicki (U.S. Census Bureau)

Unresolved Matched Records in Capture-Recapture Methodology
Andrea Lamas (National Agricultural Statistics Service)
Linda Young (National Agricultural Statistics Service)
Denise Abreu (National Agricultural Statistics Service)
Shu Wang (University of Florida)
Daniel Adrian (National Agricultural Statistics Service)

Incorporating Misclassification into Capture-Recapture Methodology in the 2012 Census of Agriculture
Daniel Adrian (National Agricultural Statistics Service)
Linda Young (National Agricultural Statistics Service)
Denise Abreu (National Agricultural Statistics Service)
Shu Wang (University of Florida)
Andrea Lamas (National Agricultural Statistics Service)

11:45 a.m. - 1:15 p.m.
Lunch on Your Own

1:15 - 2:45 p.m.

CONCURRENT SESSION J-1:
ANALYTIC CONSIDERATIONS OF DATA QUALITY IN RECORD LINKAGE

Location: Room 206
Organizer: Mary Layne (U.S. Census Bureau)
Chair: Amy O' Hara (U.S. Census Bureau)

The Impact of Name Quality on Record Linkage
Mary Layne (U.S. Census Bureau)
Deborah Wagner (U.S. Census Bureau)
Cynthia Rothhaas (U.S. Census Bureau)

Analytic Challenges with National Survey Data Linked to a State-Level Cancer Registry
Eric Miller (National Center for Health Statistics)
Dean Judson (National Center for Health Statistics)
Yulei He (National Center for Health Statistics)
Jennifer Parker (National Center for Health Statistics)
Hannah Day (National Center for Health Statistics)

Clean-up and Statistical Analysis of Sets of National Files
William Winkler (U.S. Census Bureau)

Evaluating Hispanic and Asian Name Algorithms as a Tool for Improving Linkage to the National Death Index
Dean Judson (National Center for Health Statistics)
Eric Miller (National Center for Health Statistics)
Jennifer Parker (National Center for Health Statistics)
Hannah Day (National Center for Health Statistics)

1:15 - 2:45 p.m.

CONCURRENT SESSION J-2:
UNDERSTANDING AND ADDRESSING SURVEY NONRESPONSE

Location: Room 207A
Organizer and Chair: Lynn Langton (Bureau of Justice Statistics)

On Checking Whether Response is Ignorable or Not
Michail Sverchkov (Bureau of Labor Statistics)

The Utility of the Integrated Design of the Medical Expenditure Panel Survey to Inform Trends in MEPS Nonresponse
Frances Chevarley (Agency for Healthcare Research and Quality)
Karen Davis (Agency for Healthcare Research and Quality)

The Impact of Efforts to Increase Response Rates on Survey Estimates
Karen Wessels (ICF International)
Frances Barlas (ICF International)

Changing the Way We Look at Survey Nonresponse
Deborah Griffin (U.S. Census Bureau)
Dawn Nelson (U.S. Census Bureau)

1:15 - 2:45 p.m.

CONCURRENT SESSION J-3:
INNOVATIVE APPROACHES FOR SAMPLING SPECIAL POPULATIONS

Location: Room 207B
Organizer and Chair: Joy Sharp (Bureau of Transportation Statistics)

A New Source of Local Health Data: Facebook Likes
Steven Gittelman (Marketing, Inc.)
Elaine Trimarchi (Marketing, Inc.)

A New Online Sample Frame Concept: Passive Data Collection
Elaine Trimarchi (Marketing, Inc.)
Steven Gittelman (Marketing, Inc.)
Philip Garland (Survey Monkey)

The Use of Non-Probability Samples to Characterize Rare Conditions
John Boyle (ICF International)

Discussant: Rob Santos (Urban Institute)

1:15 - 2:45 p.m.

CONCURRENT SESSION J-4:
NEW MEASUREMENT SOLUTIONS: AGREEMENT, INTRACLASS CORRELATION, AND OUTCOME PATTERNS

Location: Room 202B
Organizer and Chair: Andrew White (National Center for Education Statistics)

State Based Intraclass Correlation Values for Planning Group-Randomized Trials in Education: Within and Between District Estimates
Eric Hedberg (NORC at the University of Chicago)
Larry Hedges (Northwestern University)

Agree or Disagree? A Demonstration of An Alternative Statistic to Cohen's Kappa for Measuring the Extent and Reliability of Agreement Between Observers
Qingshu Xie (MacroSys, LLC)

Statistical Tests of Agreement
Elizabeth Stanwyck (University of Maryland, Baltimore County)
Bimal Sinha (University of Maryland, Baltimore County)
Barry Nussbaum (U.S. Environmental Protection Agency)

Measuring Health and Healthcare Disparities
James Scanlan (James P. Scanlan, Attorney at Law)

Abstract Booklet

This section represents abstracts received as of August 2013.

The following abstracts have not been edited for content.

CONCURRENT SESSION A-1

CONSIDERATIONS WHEN ADOPTING ADDRESS-BASED SAMPLING FOR FEDERAL SURVEYS

On Moving to the Use of a Hybrid Sampling Frame in the National Survey on Drug Use and Health: Motivations and Challenges

Arthur Hughes (Substance Abuse and Mental Health Services Administration), Jonaki Bose (Substance Abuse and Mental Health Services Administration), Bonnie Shook-Sa (RTI International), and Katherine Morton (RTI International)

The sampling frame currently used in the National Survey on Drug Use and Health (NSDUH) is constructed by Listers who visit each selected area segment and conduct a complete enumeration of dwelling units that are eligible for selection. The half-open interval (HOI) procedure is used by Field Interviewers during the data collection phase to identify and include dwelling units that were missed during the listing operation. The use of address-based sampling (ABS) is increasingly becoming a viable alternative to traditional approaches for developing sampling frames and has the potential to significantly reduce costs associated with the listing operation. ABS rely solely on the use of a computerized listing of addresses (maintained and updated by the U.S. Postal Service) instead of field enumerated listings to select samples of dwelling units.

Research was conducted recently to evaluate the coverage, cost, and implementation of a hybrid sampling frame for NSDUH, where ABS would be employed in most area segments while the traditional listing method would continue in areas where ABS coverage is low. A new Check for Housing Units Missed (CHUM) procedure was proposed to identify missed dwelling units to serve as a replacement for the HOI. This conference paper will summarize the current NSDUH sample design including the sampling frame procedure, present potential alternative hybrid sampling frames based on prior research, and discuss motivations, challenges, and current decisions associated with implementing a hybrid frame in NSDUH.

Current and Future Uses of Address Lists for the National Health Interview Survey Sample

Chris Moriarity (National Center for Health Statistics)

The National Health Interview Survey (NHIS) is a continuous survey sponsored by the National Center for Health Statistics (NCHS) that has collected health data using personal interviews since 1957. The Census Bureau has been the NHIS field collection agent since the beginning of the NHIS. Beginning with the 1985 NHIS, Census Bureau personnel have created address lists via canvassing sample areas, and these lists were then used to select sample addresses for inclusion in the survey.

The current address list creation procedure is expensive, and it will become even more expensive within a few years as other federal agencies currently sharing field canvass costs with NCHS move to using the Census Bureau's Master Address File (MAF) as their primary sample address source. The MAF is not an option for the NHIS, because the Census Bureau cannot release MAF addresses to NCHS. NCHS moved to field canvassing in 1985 so the NHIS sample addresses could be released to NCHS for followback surveys, record linkage to administrative data, etc.

NHIS will begin using an address list in the 2013 NHIS as part of a sample augmentation in select states. NHIS also will transition to using one or more address lists as the primary sample address source for the next NHIS sample design period beginning in 2016. I discuss our planning processes for current and future uses of address lists for the NHIS sample.

A Tale of Two Surveys: Learning From the Application of Address-Based Sampling in Federal Surveys

Eileen M O'Brien (Energy Information Administration)

Choosing one sample frame (or frames) over another requires considering many features of a survey's design and purpose. Initially, the statistical suitability of the frame depends on how well the frame covers the population of interest and what potential for bias exists when frame missingness is correlated with key survey measures. In recent years, however, pressure to stem rising survey costs has motivated data collectors to test and adopt innovative sampling methods. This paper discusses what factors contribute to a decision to change frames in a federal survey environment and what challenges may result. They may include murkier relationships between frame elements and the survey's unit of analysis; varying presence and quality of stratification variables, which are key to reducing the effect of sample clustering in area probability designs; and different capacities to observe, monitor and adjust for errors of observation and non-observation. This paper concludes with suggestions for balancing the budget-quality tradeoffs given fixed features and unique mandates of a federal survey program, including mixed sample designs, new and multiple modes of data collection, and new technologies for monitoring and adjusting for errors in design, implementation and enumeration.

Calculating Eligibility Rates in an Address-Based National Household Survey

Cameron McPhee (American Institutes for Research), Amber Noel (American Institutes for Research), and Danielle Battle (American Institutes for Research)

The calculation of accurate survey response rates is crucial to the rigorous assessment of data quality and the representativeness of the results to the target population. In addition to the number of respondents, the response rate calculation requires knowing the number of eligible cases in the sample. Oftentimes, the eligibility status of some sampled cases is unknown. These cases pose a challenge to the calculation of response rates; underestimating the number of ineligible cases will lower reported response rates, while overestimating the number of ineligible cases incorrectly inflates them.

AAPOR standards provide several methods for estimating eligibility under various survey designs and data collection methodologies. However, little has been published regarding the calculation of eligibility rates for mail surveys *without* specifically named respondents or other household-level eligibility criteria. While the postal service provides some information that can be used to classify ineligible cases through the return of non-deliverable mail, relying exclusively on this information may underestimate the number of ineligible sampled cases. Similarly, while address-based sampling frame vendors often provide appended information about the occupancy status of certain addresses, this information is not always current and accurate.

This paper will explore several methods for calculating the eligibility (and subsequent response rates) for a household mail survey. Data for this analysis come from the 2012 National Household Education Survey sponsored by the National Center for Education Statistics. The survey used an address-based sample of 160,000 addresses and includes appended information about address type and occupancy status. Analyses will compare response rates calculated under various eligibility assumptions overall and within address type and will examine the accuracy of the frame information as well as the consistency of postal-service returned mail delivery within different address types. Results will provide researchers better methodological guidelines for calculating eligibility for mail surveys in the future.

CONCURRENT SESSION A-2

DEVELOPING SMALL AREA ESTIMATION MODELS

Comparison of State-level Estimates of Adoption of Electronic Medical/Health Records Systems by Using Area-level Models and Direct Estimates

Vladislav Beresovsky (National Center for Health Statistics) and Janey Hsiao (National Center for Health Statistics)

Literature suggests that use of electronic medical record /electronic health record (EMR/EHR) Systems can be important in improving quality of healthcare. Reliable estimates of the proportion of physicians using EMR/EHR systems by state are important in monitoring the progress in adopting this technology in the US. The National Electronic Health Records Survey, a mail survey to office-based physicians, collects information on EMR/EHR systems and has been used to report official design-based estimates on EMR/EHR adoption at the state-level. In this analysis we apply model-based and model-assisted small area estimation methods to improve precision of state-level estimates. We conduct a simulation study to compare mean-squared errors of model-based and randomization-based estimates. Similar model-based methods were previously investigated and used by the Census Bureau to estimate poverty rates among children in small places.

Statistical Modeling of National Survey of College Graduates and American Community Survey Variables

Michael Larsen (The George Washington University)

The National Survey of College Graduates (NSCG) has been conducted by the Census Bureau for the National Science Foundation (NSF) since the 1960s. It is the nation's only source of detailed statistics on the science and engineering (S&E) labor force. The NSCG uses a rotating panel design and selects its sample on a biennial basis from the American Community Survey to allow both cross-sectional and longitudinal analysis of education, employment, and demographic characteristics of the S&E labor force. Under this design, the NSCG data is collected and released on a biennial or triennial schedule. The 2010 survey cycle marked the initial use of the ACS as a sampling frame for the NSCG. The 2010 NSCG responses allow NSF the ability to produce estimates of the S&E labor force as of the 2010 calendar year. The next NSCG survey cycle is scheduled for 2013 and will allow NSF the ability to produce estimates of the S&E labor force as of the 2013 calendar year. This paper discusses issues of modeling data between waves of the NSCG and between NSCG and ACS surveys. Statistical modeling is critical for producing estimates in off-years of the survey, updating estimates for currently unused ACS waves, and small area estimation.

Variance Modeling Research for the Small Area Health Insurance Estimates Program

Mark Bauder (U.S. Census Bureau), Sam Szelepka (U.S. Census Bureau), and Donald Luery (U.S. Census Bureau)

The Small Area Health Insurance Estimates (SAHIE) program uses survey estimates from the American Community Survey (ACS), along with administrative data, to estimate health insurance coverage for small domains. The component models for the ACS survey estimates include the sampling variances. A common practice for small-area estimation has been to use the direct survey estimates as if they were measured without error. Because the domains of interest often have very small sample sizes, direct estimates of these ACS variances are too unreliable to use directly in the model. Instead, the current SAHIE model assumes that ACS sampling variances follow a functional form, with parameters that are estimated jointly with the other unknowns in the model.

An alternative to the current SAHIE methodology, and to using direct ACS variance estimates in the SAHIE model, is to use a generalized variance function (GVF) to estimate the ACS variances outside of the SAHIE model. In this approach, direct sampling variance estimates are used to estimate parameters in a GVF, and the GVF is then used to produce smoothed sampling variance estimates. In this paper, we

report results from research into developing a GVF model for the purpose of incorporating the resulting smoothed variance estimates into the SAHIE model.

Progress in Developing Small Area Estimates of Crime Based on the National Crime Victimization Survey

Robert Fay (Westat), Mamadou Diallo (Westat), and Michael Planty (Bureau of Justice Statistics)

The Bureau of Justice Statistics (BJS) is expanding the geographic detail available from the National Crime Victimization Survey (NCVS) in order to increase its usefulness for its users. For example, a supplement to the NCVS sample begun in July 2013 will enable publication of direct estimates of 3-year crime rates in the 11 most populous states. BJS has also supported research on possible application of small area estimates based on the NCVS data. At the previous FCSM Research Conference, Li, Diallo, and Fay (2012) outlined initial strategies to address this objective.

A number of small area applications have developed cross-sectional area-level models for application to ongoing repeated cross-sectional surveys. In many cases, modeling is restricted to the sample data for the current time period, ignoring possible information from sample estimates during previous periods. The Rao-Yu model initially seemed appropriate for the NCVS application, and subsequent research proposed and applied a modified version of this model to the NCVS at the state-level. To take into account the types of crime representing the components of the overall violent or property crime rate, a multivariate extension of the model has been developed and applied.

In this paper we summarize the application of these methods to produce state-level estimates. We then present recent work to extend the approach to large counties and possibly to large cities. As the model is extended to smaller geographic areas where less NCVS data are available, its performance in the context of increasingly sparse data will be examined. We also reflect on how the small area estimates may find a place in BJS's overall statistical program.

CONCURRENT SESSION A-3

DESIGN AND TESTING STRATEGIES FOR WEB SURVEYS AND WEBSITES

The Impact of Survey Communications on Response Rates and Response Quality

Frances Barlas (ICF International), Amy Falcone (ICF International), Nikki Bellamy (Substance Abuse and Mental Health Services Administration), Amy Mack (ICF International)

This paper investigates how efforts to improve response rates affect response quality in web surveys by assessing a number of indicators of quality. Response rates for all modes of survey administration have been declining in recent years and web-based surveys tend to have lower response rates than other administration modes. We conducted a survey to evaluate the services and supports provided by the Substance Abuse and Mental Health Services Administration's (SAMHSA) Disaster Technical Assistance Center (DTAC). This study extends previous research that examined how the design of survey communication emails and landing page interact to affect survey response. The previous study found highest survey completion rates among those sent a long versus short email *and* presented with the web survey link at the top of the landing page versus the bottom. This may suggest respondents perceive survey burden only within the web survey interface and not the invitation email, and therefore may tolerate more information in email than on a landing page. The current study tests this by creating a 2x2 factorial design, where sample members are either presented the survey information needed for informed consent on a landing page (receiving a short email) or in an email (receiving a long email and bypassing the landing page), and either have the survey link/button presented at the top or bottom of the email communication. Maximizing survey response rates decreases the risk of nonresponse bias, helping to ensure that the survey will provide valid results. We compare response and completion rates and investigate whether the differences translate to variation in response quality by assessing response times, item non-response, patterns of systematic responses to grid questions, and use of the not applicable option when available. In addition we compared survey estimates across the four conditions to assess the impact on study findings.

Developing an Inclusive Web Survey Design for Respondents with Disabilities

Jessica Jagger (U.S. Marine Corps), Ashley Schaad (ICF International), Ashleigh Davis (ICF International), and Amy Falcone (ICF International)

Although a deep body of literature exists on the design of web surveys of the general population, little or no studies have examined the design of such surveys for individuals with disabilities. Well-intentioned researchers often design web surveys that fall short of compliance with Section 508 of the U.S. Rehabilitation Act and rarely go beyond compliance to consider maximizing the accessibility of their data collection tools. The current study presents the design challenges faced and strategies developed in creating and testing a web survey targeting those eligible for the Talking Book and Braille Program through the Library of Congress National Library Service for the Blind and Physically Handicapped (NLS). Many of the standard practices in designing web surveys, such as minimizing scrolling on web pages, are not effective in designing surveys for respondents with visual disabilities. Furthermore, taking a web survey with a screen reader often takes longer than taking it without one, increasing respondent burden for those using screen readers. Testing web surveys for accessibility with multiple screen readers can also be challenging for researchers who are not familiar with the technologies and do not have access to the numerous different technologies available. Commercial and free screen reader software is available, but these can often be difficult for non-users to master which can create challenges in testing for ease of use and response burden. Also important is testing the survey with those who will be using other methods of web accessibility, such as magnification software. Data sources include participant reactions captured during a pilot test of the web survey and anecdotally reported throughout the survey period, as well as a comparative analysis of average survey length among respondents using the web survey versus average survey length among those completing the same survey over the phone.

Improving Government Websites and Surveys with Usability Testing: A Comparison of Methodologies

Jennifer Romano Bergstrom (Fors Marsh Group) and Jonathan Strohl (Fors Marsh Group)

In order to ensure an optimal user experience and quality data, usability testing of websites and surveys is necessary. Usability testing with typical end users can be conducted in many ways, including remotely, in-the-field, and in-the-lab. In addition to usability tests, other modes of collecting user feedback include focus groups, one-on-one interviews, and surveys. In this paper, we describe several methods we have used to evaluate the usability of numerous government websites and surveys. For example, we will detail how we have used:

- Remote user experience testing to evaluate Internal Revenue Service forms
- Focus groups, one-on-one interviews, and In-lab usability tests to evaluate U.S. Department of Veterans Affairs websites and web-based tools
- In-the-field one-on-one interviews to evaluate Department of Defense social media sites
- In-lab cognitive and usability studies to evaluate Department of Defense paper and web-based surveys
- In-lab password usability studies on tablets and smartphones for the National Institute of Standards and Technology
- In-lab usability tests to evaluate National Cancer Institute websites, booklets, and videos.

In the presentation, we will describe the tools we typically use to conduct the different types of usability tests, including eye-tracking technology, satisfaction and emotion surveys, participant activities, and remote testing software that allows for screen sharing. We will discuss the various types of data each method yields, including qualitative (e.g., verbalizations, facial expressions, errors made) and quantitative (e.g., accuracy, time and number of steps to complete tasks, error rate) data. Attendees will learn specific issues that government agencies and contractors need to be aware of and the technical requirements and pros and cons of each method.

CONCURRENT SESSION A-4

DATA QUALITY: ADMINISTRATIVE RECORDS VERSUS CENSUS AND SURVEY DATA

Evaluating Race and Hispanic Origin in Administrative Records Relative to the 2010 Census

Sonya Rastogi (U.S. Census Bureau), James Noon (U.S. Census Bureau), Amy O'Hara (U.S. Census Bureau), and Ellen Zapata (U.S. Census Bureau)

The U.S. Census Bureau acquired administrative records data from federal agencies and commercial vendors to conduct the 2010 Census Match Study, an official Census Bureau evaluation that assessed the quality and coverage of administrative records relative to the 2010 Census (Rastogi and O'Hara 2012). The 2010 Census Match Study was a proof of concept for using administrative records data in a U.S. decennial census context. The results of the study indicate that while administrative records cannot replace a decennial census, they can enhance decennial census operations. This paper presents an overview of the study's descriptive evaluation on administrative records demographic coverage and quality relative to the 2010 Census, and it extends the evaluation by utilizing regression analysis to assess housing and person characteristics that may be associated with the agreement of race or ethnicity between administrative records and 2010 Census persons. These results reveal gaps in administrative records quality and coverage of race and Hispanic origin reporting. The work informs future data acquisitions and research to operationalize administrative records in decennial and survey designs.

2010 American Community Survey Match Study: Comparing Administrative Records and ACS Data

Adela Luque (U.S. Census Bureau) and Renuka Bhaskar (U.S. Census Bureau)

Using administrative records data from federal government agencies and commercial sources, the 2010 ACS Match Study measures administrative records coverage of 2010 ACS addresses, persons, and persons at addresses at different levels of geography as well as by demographic characteristics and response mode. The 2010 ACS Match Study represents a continuation of the research undertaken in the 2010 Census Match Study, the first national-level evaluation of administrative records data coverage.

Preliminary results indicate that administrative records provide substantial coverage for addresses and persons in the 2010 ACS (92.7 and 92.1 percent respectively), and less extensive though substantial coverage, for person-address pairs (74.3 percent). In addition, some variation in address, person and/or person-address coverage is found across demographic and response mode groups. This research informs future uses of administrative records in survey and decennial census operations to address the increasing costs of data collection and declining response rates.

Coverage and Agreement of Administrative Records and 2010 ACS Demographic Data

Renuka Bhaskar (U.S. Census Bureau), Adela Luque (U.S. Census Bureau), James Noon (U.S. Census Bureau), and Sonya Rastogi (U.S. Census Bureau)

The U.S. Census Bureau is researching possible uses of administrative records in decennial census and survey operations. The 2010 Census Match Study and American Community Survey (ACS) Match Study represent efforts by the Census Bureau to evaluate the extent to which administrative records provide data on persons and addresses in the 2010 Census and 2010 ACS. The Census Match Study also examined the quality and coverage of demographic response data collected in administrative records. This paper seeks to build on this analysis by evaluating the quality and coverage of demographic information from administrative records relative to the ACS. This study matches data from the 2010 ACS to federal and commercial administrative records as well as to previous census data to examine age, sex, race, and Hispanic origin responses. Coverage and quality of demographic response data are evaluated for each of the different federal and commercial administrative records source files by the 2010 ACS mode of data collection.

Medicaid Undercount in the American Community Survey

Kathleen Call (University of Minnesota – SHADAC), Brett O'Hara (U.S. Census Bureau), Joanna Turner (University of Minnesota – SHADAC), Michel Boudreaux (University of Minnesota – SHADAC), and Brett Fried (University of Minnesota – SHADAC)

Surveys are the only source of estimates of the distribution of health insurance in the population, representing a critical source for evaluating the impact of the Patient Protection and Affordable Care Act (ACA). However, measuring health coverage is challenging and virtually every survey is said to undercount Medicaid enrollment. In surveys such as the National Health Interview Survey (NHIS), Medical Expenditure Panel Survey (MEPS) and the Current Population Survey (CPS) Medicaid enrollment counts are always lower than counts available from enrollment data. If enrollees do not report Medicaid, estimates of other coverage or being uninsured will be biased upwards and Medicaid estimates will be biased downwards.

We extend work from the SNACC team, a multi-phase project examining the Medicaid undercount in these other federal surveys, to the American Community Survey (ACS). We use linked 2008 ACS data, the first year health insurance variables were available, and 2008 Medicaid Statistical Information System (MSIS) data to examine the extent to which Medicaid enrollment is misreported. We compare the magnitude of the undercount and factors associated with misreporting in the ACS to other federal surveys. From previous research we know that measuring health coverage is prone to some level of error and is worse in surveys with extended recall periods; yet bias to uninsured estimates is minimal. This work provides the first look at the Medicaid undercount in the ACS and is part of a research agenda to further explore patterns of misreporting and its effect on coverage estimates.

The Medicaid Undercount in the 2007-2009 Current Population Survey

James Noon (U.S. Census Bureau), Leticia Fernandez (U.S. Census Bureau), and Sonya Rastogi (U.S. Census Bureau)

As Medicaid enrollment is expected to increase with the Patient Protection and Affordable Care Act, understanding differences in estimates of Medicaid coverage between survey and administrative data will be important for better assessing and tracking changes in enrollment over time. Previous research has shown that estimates of Medicaid coverage based on survey data consistently produce an undercount of beneficiaries when compared to actual Medicaid enrollment records. This paper extends past work by examining the Medicaid undercount in the 2007-2009 Current Population Survey (CPS) as compared to enrollment data from the National Medicaid Statistical Information System (MSIS) for calendar years 2006-2008.

Linking individuals across these datasets, we update previous estimates of two types of disagreement regarding Medicaid enrollment rates. First, some persons have Medicaid coverage but report no coverage in CPS (false negative error). For these individuals, we examine the details of their Medicaid benefits and their personal characteristics. Second, we also examine the characteristics of persons who report Medicaid coverage in the CPS but are not linked to Medicaid enrollment data (false positive errors).

Consistent with earlier findings, preliminary analysis suggests that the rate of false negative error increases with individuals' income and declines with the number of days enrolled in Medicaid. In contrast, the rate of false positive error tends to decline with individuals' income, and a larger component can be explained by imputation or editing procedures relative to false negative error. We discuss possible reasons for differences in our estimates with respect to previous research.

CONCURRENT SESSION B-1

BUILDING BLOCKS OF ADAPTIVE SURVEY DESIGN

How Much Is Enough? Moving Toward Smart Stopping Rules for Data Collection

Wan-Ying Chang (National Center for Science and Engineering Statistics) and Lynn Milan (National Center for Science and Engineering Statistics)

For large-scale national surveys, the mandate to achieve a high response rate can result in a prolonged data collection period. However, as timeliness of survey data becomes a higher priority, improving survey efficiency must focus not only on cost reduction but also on accelerating the delivery of survey results. The Survey of Doctorate Recipients (SDR), a biennial longitudinal survey sponsored by the National Science Foundation and the National Institutes of Health, is built on a sampling frame rich in auxiliary information useful in determining areas for survey efficiency considerations. For past SDRs, data collection typically lasted more than 40 weeks, to achieve the OMB-suggested response rate of 80 percent or higher. The recent developments on responsive design and data quality monitoring indicators offer promising tools to assess the tradeoff between quality, cost, and timely productivity. In this study, we analyze the weekly progress of the 2010 SDR in terms of (1) balance between the response set and the full sample, (2) non-response bias on auxiliary variables known to be highly correlated with key survey outcomes, and (3) the data quality at the item level. We then develop non-response weighting adjustment and imputation procedures for estimation. Estimates of key survey outcomes are tracked weekly for their stability and estimates of selected auxiliary variables are tracked weekly for benchmarks to evaluate effects of weighting adjustment on bias correction. We consider several stopping rules and evaluate the tradeoff between shortened data collection and increased bias and variance in estimation.

Planning an Adaptive Design Treatment in 2020 Census Tests

Gina Walejko (U.S. Census Bureau), Peter Miller (U.S. Census Bureau), Gianna Dusch (U.S. Census Bureau), and Kevin Deardorff (U.S. Census Bureau)

Adaptive design is usually considered in the context of sample surveys. Here we describe the use of adaptive design in a census environment. First, we discuss why and how researchers may choose to implement adaptive design in a household census. Next, we describe an adaptive design experiment planned for an upcoming U.S. Census Bureau decennial nonresponse follow-up test. The test compares an adaptive design method for case management in nonresponse follow-up with a standard enumeration approach. We describe how cases in the treatment group will be assigned contact mode and enumeration priority. We illustrate the frame data and paradata resources used to inform evolving business rules for contact attempts during enumeration. Finally, we discuss data quality and cost measurements to be employed in evaluating the adaptive design treatment.

Adaptive Curtailment of Survey Followup Based on Contact History Data

Eric Slud (U.S. Census Bureau) and Chandra Erdman (U.S. Census Bureau)

Several major Census Bureau household surveys employ a common form of Contact History paradata called the Contact History Instrument (CHI). At each attempt at household contact, computer-aided telephone (CATI) or personal (CAPI) interviewers enter various codes for interviewee reluctance in a standard format to aid in planning for subsequent contact attempts.

The American Community Survey (ACS) in particular, in its CATI phase, schedules household call attempts with regard to previous nonproductive calls and reluctance. Analyses of ACS contact histories have recently been conducted (Zelenak and Davis 2013, ACS Memorandum #ACS13-RER-08, Griffin and Hughes 2013) to explore CATI curtailment policies which can minimize the perceived harassment of households due to multiple contact attempts while losing as few interviews as possible.

This paper describes simple methods which can support adaptive design to yield optimal interview rates subject to constraints on survey follow-up. These methods make minimal use of models and consist of the

following primary elements. **(1)** Interview yields per contact attempt are constructed as functions of the number of contact attempts, on mutually exclusive subgroups of the survey population. **(2)** Subgroups defined from contact-history variables are chosen to achieve maximal separation of the curves defined in (1). **(3)** Policies designed to curtail follow-up differentially in the separate subgroups found in (2) are compared with respect to an array of costs or loss-functions, including measures of perceived harassment which themselves vary in separate paradata-defined groups, under simple models based on tabulated groupwise response behavior. No policy can simultaneously minimize all costs, but this kind of analysis allows us to identify decision-theoretically admissible policies and to choose policies which most efficiently trade off the most essential cost elements.

The steps described above are illustrated using ACS CATI data, in light of current ACS design aimed at minimizing perceived harassment.

Using Substantive Variables and Paradata to Inform Responsive Design Models: Results From a Nationally Representative Longitudinal Study

Ted Socha (National Center for Education Statistics)

The Baccalaureate and Beyond Longitudinal Study (B&B), conducted by the Department of Education, examines students' education and work experiences after they complete a bachelor's degree. The 2012 B&B included a field test experiment to evaluate the use of response propensity modeling (based on Mahalanobis distancing) to target cases with low likelihood of response, with the goal of improving weighted response rates and thereby minimizing nonresponse bias. This presentation will include results from this experiment, including: (1) choice of RD model, (2) application of the model, (3) implications of using RD for monetary and non-monetary incentives, (4) inclusion of substantive and paradata model variables, (5) dealing with missingness in substantive variables, and (6) discussion of nonresponse bias reduction experiment results.

Adaptive Design Strategies: Who Do You Target? And What Do You Do With Them?

Jaki McCarthy (National Agricultural Statistics Service), Kathy Ott (National Agricultural Statistics Service), and Melissa Mitchell (National Agricultural Statistics Service)

With limited time or resources, the best way to handle data collection for a given sample depends heavily on the survey in question. Optimal procedures for a sample may differ depending on the sample design and estimation scheme, data collection period and procedures, paradata available, information known about the sample units, quality standards for the survey estimate, and costs. Two different NASS surveys where data collection procedures were altered to improve survey results given cost constraints will be discussed.

The Quarterly Agricultural Survey (QAS) produces crop statistics at the state and national level. NASS has developed predictive models to identify, prior to data collection, QAS sample units most likely to be nonrespondents. Data collection strategies for these units can be different from those used for "easy" respondents – for example using specific interviewers, or in person interviews. However, high cost strategies are likely only necessary for potential nonrespondents who can substantially impact survey estimates. Records that have both high nonresponse propensities AND potentially large impact on survey estimates were identified as "impact" records for special handling.

NASS also has a predictive model to pre-identity highly likely nonrespondents in the Agricultural Resource Management Survey (ARMS) which collects farm finance and cost of production data. ARMS uses calibration weighting to adjust for nonresponse and undercoverage. The likely nonrespondent group is quite large and special handling strategies cannot be applied to all of them. In this case, "impact" records were those that will contribute the most in coverage for the calibration targets.

While records can be identified for special handling in data collection, it is not always clear what special handling may be effective in increasing response rates for these records. "Special handling" procedures were applied to half of the identified likely nonrespondent "impact" records compared to half treated

normally in each of these surveys. Whether our special handling was effective and the potential effect on the survey estimates will be discussed.

CONCURRENT SESSION B-2

SOLVING DATA PROBLEMS WITH BAYESIAN METHODS: MODELING, COMPLIANCE CONTROL, AND CLIMATE CHANGE

Sample Design and Analysis for Assessing Compliance Control Settings

Bhojnarine Rambharat (Office of the Comptroller of the Currency)

The design and analysis of a sample from a population is important in establishing controls to mitigate false negatives in a given study. An error (or exception) in a sample is essentially a false negative or Type II error. A financial institution needs to ensure that the appropriate settings are established and tested as these controls helps to enhance compliance functions. Sampling techniques allow compliance analysts to test those units in a population that are not flagged as errors but could potentially be errors, or it could allow an analyst to better understand a population of interest so that false negative rates could be minimized. However, a given sampling design may not account for certain nuances of the population where it may be initially expected to contain few or no errors. Attribute sampling is often used to understand the error rate in such a population. However, there is typically not enough attention given to the total number of errors even though this could be an important quantity to compliance functions or to policy-makers in the regulatory sphere. A sampling paradigm is discussed where Bayesian methodologies are leveraged to estimate the posterior distribution of the total number of errors in a population. Subsequently, we utilize the distribution of the total number of errors in a population to decide how to conduct more in-depth, or ``deep-dive,'' research about the errors in the population. Furthermore, the results of the deep dive analysis could provide insights on how to adjust important compliance controls. We discuss two stages of sample design and analysis: i) to estimate the total number of errors in a given population, and ii) to design samples for deep dive studies aimed at understanding the root cause of errors in a population.

The views expressed in the abstract are solely those of the author and do not, in any way, reflect the opinions of the Office of the Comptroller of the Currency or the U.S. Department of the Treasury.

Log Linear Mixed with Additive Random Components Modeling for Small Area Estimation: A Bayesian-Frequentist Integrated Approach

Avinash Singh (NORC at the University of Chicago)

In the context of SAHIE application, we propose a hierarchical Bayes (HB) model termed log linear mixed with additive random components-LLMARC which starts with a building block model so that a single multivariate model can be used for both county and state level estimates. The observation is a combined 10-variate (5 Income-to-Poverty Ratio by 2 Insurance Coverage categories) vector of counts. The expected count is modeled by the inverse-log link function while allowing for normal random effects to be part of the nonlinear mean function. Modeling estimated domain total counts is preferable to estimated domain means as domain counts can be collapsed to obtain a reduced model with the same parameters but with a stable observed variance-covariance (V-C) matrix. This observed V-C matrix can be further smoothed by generalized covariance functions, but instead of using the customary simple random sample covariance as an initial stable working covariance matrix which is known to be inadequate for capturing positive covariances, a suitable alternative but stable working covariance matrix is suggested in order to better capture the underlying design features. Under LLMARC, it turns out that a closed form expression of the marginal mean after integrating out the random effect can be obtained which makes its interpretation somewhat analogous to linear mixed models-LMM. This gives rise to frequentist LMM-type parameter estimates in a Bayesian-Frequentist Integrated approach which can be used to perform usual frequentist-type diagnostics and test HB results. Having an analytical form of the marginal mean (i.e., without random effects) allows for introducing new covariates for approximate built-in benchmarking of HB estimates to key higher level direct estimates. In addition, LLMARC models allow for simple ways of

introducing spatial correlations through simultaneous auto regression and temporal correlations through state space modeling as a means to take advantage of additional internally available

Spatio-Temporal Modeling of the Impact of Climate Change on Road Traffic Accidents - A Case Study of New Brunswick

Md Shohel Reza Amin (Concordia University, Canada), Alireza Zareie (McGill University, Canada), and Luis Amador-Jiménez (Concordia University, Canada)

The objective of this research is to study the impact of climate change on the hazardous weather-related road collisions. The Canadian province of New Brunswick is considered in a case study. The study uses road collision data for both single and multiple collisions, from police accident reports. Thirty-years of weather records were used in the analysis, this included daily data for seven different zones of New Brunswick, National Centers for Environmental Prediction (NCEP) re-analysis dataset, and large-scale simulation data from the Canadian Global Circulation Model, General Climate Model, and General Circulation Model (CGCM3). The large-scale simulation data from Canadian GCM under SRES-A2 scenario during 21st century are used to model the climate in the future. This study develops an Exposure to Weather-Accident Severity (EWAS) index and estimate the relationship between EWAS index and weather-related explanatory variables of road accidents by applying negative binomial regression and Poisson regression models. The regression models find out that surface-weather condition, weather-driver's gender, weather-driver's age, weather-driver's experience and weather-vehicle's age have strong positive correlation with weather-accident severity (EWAS index). The surface-road alignment and surface-road characteristics have negative relationship with EWAS index. The spatial pattern of EWAS index with respect to weather-related explanatory variables is examined for the fifteen census divisions of New Brunswick, which derives similar results. The climate change modeling estimates that the number of rainy days may increase for all of the climate zones and the number of snowy days and freezing days may decrease or stay the same for most of the zones in the province during three different future periods in the 21st century (i.e. 2011-2040, 2041-2070, 2071-2100). More hazardous weather will result in increased accident severity. This study suggests that the Road Safety Strategy 2015 of Transport Canada should not only adopt the holistic approaches based on the impaired driving, speed and aggressive driving, occupant protection, it should also take protective measures for the hazardous weather conditions in order to reduce accident severity.

Weight Smoothing with Laplace Prior and Its Application in GLM Model

Xi Xia (University of Michigan) and Michael Elliot (University of Michigan)

When analyzing data sampled with unequal inclusion probabilities, correlations between the probability of selection and the sampled data can induce bias. Weights equal to the inverse of the probability of selection are commonly used to correct this possible bias. When weights are uncorrelated with the sampled data, or more specifically the descriptive or model estimators of interest, highly disproportional sample design resulting in large weights can introduce unnecessary variability, leading to an overall larger root mean square error (RMSE) comparing to the unweighted or Winsorized methods.

We describe an approach we term weight smoothing that models the interactions between the weights and the estimators of interest as random effects, reducing the overall RMSE by shrinking interactions toward zero when such shrinkage is supported by data. This manuscript adapts a more flexible Laplace prior distribution for the hierarchical Bayesian model in order to gain more robustness against model misspecification. We consider this approach in the context of a generalized linear model, and evaluate it in a simulation setting.

CONCURRENT SESSION B-3

ECONOMIC STATISTICS

Fair Value Accounting and Measures of Corporate Profits in the U.S. National Accounts

Dylan G. Rassier (U.S. Bureau of Economic Analysis)

Under U.S. financial and tax accounting rules, fair value accounting (FVA) is a practice in which an asset or liability is treated as sold at fair value even when no sale takes place. As a result, related holding gains and losses may be recognized in firms' financial statements and tax returns. Given the size of financial assets and liabilities subject to FVA at some finance and insurance firms, potentially misleading effects of FVA on measured performance of the firms is an area of increased scrutiny since the 2008-2009 financial crisis (Bhat, et al., 2011; Laux and Leuz, 2010).

The international guidelines for national accounting require statisticians to exclude holding gains and losses from national income and product statistics because holding gains and losses reflect changes in prices rather than economic production. The U.S. Bureau of Economic Analysis (BEA) uses financial- and tax-based source data to measure corporate profits. In addition, BEA generally measures output with survey-based source data, which are designed to be consistent with national accounting concepts but are often based on financial accounting records. While BEA adjusts the source data to remove holding gains and losses, FVA practices call into question the completeness of the adjustments (Rassier, 2012).

This paper evaluates the source data underlying measures of U.S. corporate profits and U.S. output for 2006 to 2010 in the finance and insurance industries. The goal is to determine whether the data are adjusted completely for the effects of FVA or whether corporate profits and output may include related holding gains and losses. While the extent to which the measures are affected is difficult to ascertain, preliminary results suggest corporate profit measures are more vulnerable than output measures. The paper also offers alternatives to provide information in the U.S. national accounts on adjustments for holding gains and losses.

Forecasting Retail Food Expenditures for American Households

Richard Volpe (Economic Research Service), Annemarie Kuhns (Economic Research Service), and Timothy Park (Economic Research Service)

The Food Markets Branch of the Food Economics Division in the USDA Economic Research Service studies the economics of food once it has passed the farm gate. One of the major responsibilities of the branch is to forecast retail food prices, based on the CPI, a data product of the Bureau of Labor Statistics. ERS publishes a topic page that provides monthly updates on the expected changes in the prices for a number of food categories along with a discussion of the latest changes and emerging trends in the CPI. The forecasting approach is undergoing an extensive revision to incorporate error-correction methodologies, endogenous structural breaks, and smoothing techniques to utilized quarterly data when no higher frequency is available.

One useful extension of this forecasting procedure is to estimate the expected changes in food expenditures for American households. The Consumer Expenditure Survey, which is used to calculate relative importance weights for the CPI, reveals important differences across socioeconomic groups in terms of food expenditures. Notably, the share of food expenditures attributable to various categories differ considerably when comparing low-income households to high-income households. Therefore, the expected changes in food expenditures can be very different according to income groups depending on the price forecasts. For example, the 2012 drought is expected to have the largest impacts on meat prices in 2013, meaning those prices are forecasted to rise more than most others. The lowest-income households spend over 16% of their food dollars on meats, while the highest-income spends 11%. This leads to important differences in both absolute food expenditures and in food expenditures as a share of household income. Our paper describes the ERS forecasting methodology in detail and illustrates important distributional impacts of food price forecasts across a wide survey of socioeconomic groups.

Supplemental Poverty Measure: A Comparison of Geographic Adjustments With Regional Price Parities vs. Median Rents From the American Community Survey

Trudi Renwick (U.S. Census Bureau), Bettina Aten (Bureau of Economic Analysis), and Troy Martin (Bureau of Economic Analysis)

Drawing on the recommendations of the 1995 report of National Academy of Sciences (NAS) Panel on Poverty and Family Assistance and the subsequent extensive research on poverty measurement, in 2010 an Interagency Technical Working Group (ITWG) issued a series of suggestions to the Census Bureau and the Bureau of Labor Statistics (BLS) on how to develop a new Supplemental Poverty Measure. In 2011 and 2012, the Census Bureau issued the first Research Supplemental Poverty Measure reports with poverty estimates for 2009, 2010 and 2011.

The ITWG suggested that the poverty thresholds be adjusted for price differences across geographic areas using the best available data and statistical methodology. The estimates in the Census Bureau reports use American Community Survey (ACS) data to adjust the housing portion of the poverty thresholds for differences in housing costs. This geographic cost index uses median outlays of renters for rent and utilities for two-bedroom apartments.

One shortcoming of this geographic cost adjustment mechanism is that it does not account for geographic differences in the cost of other elements of the poverty threshold. Both the 1995 NAS report and the 2010 ITWG suggestions concluded that while adjustment of the entire market basket may be desirable, adequate data on price differences for other elements did not exist.

Over the past few years, the Regional Price Branch of the Bureau of Economic Analysis has developed regional price parities (RPPs) that combine data from the BLS Consumer Price Index program with Census Bureau multi-year rents. The RPPs provide estimates of price level differences across regions for various consumption expenditure classes.

This paper will compare state-level SPM poverty rates using RPPs to adjust the poverty thresholds to rates using the index based on median rents using the 2012 Current Population Survey Annual Social and Economic Supplement.

New Firm Size JOLTS Data Provide Insight to the U.S. Labor Market

Charlotte Oslund (Bureau of Labor Statistics)

The Job Openings and Labor Turnover Survey (JOLTS) program at the Bureau of Labor Statistics began producing experimental size class estimates for the private sector in 2010. These estimates are based on the size of the establishment, or single location, since the JOLTS sample unit is the establishment. These estimates are available upon request.

The demand is even greater, however, for size class estimates based on the size of the firm, which includes all locations of a business. It is the firm that makes the upper level business decisions based on the economic climate. The JOLTS program has developed firm-based size class estimates for release in 2013 on an experimental basis. The methodology includes post-stratifying establishments to firm size. One challenge was finding a method for independent population controls on employment by firm size since there are currently no monthly employment estimates by size of firm. The JOLTS firm size estimates are for the private sector by three size classes and include all JOLTS data items: job openings, hires, quits, layoffs and discharges, other separations.

Studying the JOLTS data by size class provides a better understanding of the labor market in different sized firms throughout the business cycle. Job openings, hires, and layoffs and discharges aggregated by firm size provide insight into how businesses of different sizes change their behavior throughout the business cycle. Quits and other separations aggregated by firm size provide information about how employees of different sized firms change their behavior according to the state of the economy. The preliminary data show that in the recent recession, medium and large firms experienced the largest drops and post-recession, large firms have recovered the most, followed closely by medium firms. Small firms

were also hit hard by the recession and have shown the least post-recession improvement.

CONCURRENT SESSION B-4

ON THE STATISTICAL VALIDITY OF INTERNET OPT-IN PANEL SURVEYS

Investigating Novel Approaches for State-Based Behavioral Risk Surveillance

Carol Gotway Crawford (Centers for Disease Control), Catherine Okoro (Centers for Disease Control), and Satvinder Dhingra (Northrop Grumman)

Between traditional dual-frame random-digit dialing (RDD) surveys, which are getting more difficult and expensive, and cheaper Internet surveys that may be simply convenient or potentially unrepresentative, the time for serious re-thinking and re-engineering of survey platforms has arrived. While both types of surveys may provide nationally-representative samples, both are often infeasible at the county- and community-levels and may not adequately characterize specific subpopulations of interest.

New, cost-effective survey and analytical methods that provide reliable estimates that meet local public health requirements are needed. Online opt-in panels have been the staple of market research for several years, but the demands of the public health community for granularity and feasibility may outstrip the capabilities of current panels.
In the face of austere budgets, the need to balance competing demands of lower costs, higher quality (coverage and non-response), and more timely data will continue to be imperative. Most door-to-door face-to-face surveys using multi-stage address-based samples, still considered the gold standard, gave way to random digit dialed (RDD) phone surveys because of cost and time. Now RDD phone surveys are facing considerable challenges.
Different sampling frames, modes and analytical methods that may overcome these challenges and assist state public health professionals to continue to collect affordable quality and timely data that are representative of their respective populations are being evaluated. Novel approaches to health and behavioral surveillance include single and blended non-probability opt-in panels, and new statistical estimation methods.

This presentation covers some of the novel approaches and results from pilot studies conducted by the Division of Behavioral Surveillance, Centers for Disease Control and Prevention in support of the Behavioral Risk Factor Surveillance System. This effort represents a wide-ranging public-private collaborations with states, academic researchers, and private companies.

Sources of Comparability Between Probability Sample Estimates and Nonprobability Web Sample Estimates

William Riley (National Cancer Institute), Bob Kaplan (National Institutes of Health), David Cella (Northwestern University Feinberg School of Medicine), Catherine Okoro (Centers for Disease Control), and Satvinder Dhingra (Northrop Grumman and Centers for Disease Control)

The data needs for producing population estimates for various subgroups at varying geographic levels in a timely manner are on the rise. Because it is difficult to satisfy those needs with traditional probability samples due to their high resource requirements, survey practice has turned to data collection using Internet opt-in panels. This practice, however, does not provide data with desirable unbiased properties due to the nonprobablistic nature of the sample, yet the use of Internet opt-in panels has outpaced the effort to understand the biases produced by this approach and to develop statistical methods to correct for them.

Methods: Sample selection using matching methodology is a two-stage process for selection of “representative” samples from non-randomly selected pools of respondents. For the pilot surveys, sample matching started with an enumeration of all adults through the use of the American Community Survey. First, a true random sample was drawn from the target population. Second, for each member of the target sample, one or more matching members from the pool of opt-in Internet panelists were selected to constitute a matched sample.

In this study, we will use the data from the Patient-Reported Outcomes Measurement Information System (PROMIS) Global Health items which used an Internet opt-in sample while targeting the general population. PROMIS included a number of health-related items that were also measured in well-established probability-based national surveys. The analysis will focus on the comparison on these items across data sources. The comparison includes three types of statistics: 1) point estimates of the common variables for the general population, 2) point estimates for the population subgroups (e.g., gender, age, race/ethnicity, education, geography), and 3) relationships across variables through regression modeling. PROMIS data will be used with and without weights in these comparisons.

The findings from this study will enhance empirical evidence to understanding and using Internet opt-in panel data.

Testing Ignorable Selection Across Multiple Domains: Economic, Politics, and Health

Stephen Ansolabehere (Harvard University) and Douglas Rivers (Stanford University and YouGov)

Attempts to remove selection bias in sampling depend upon the identification of a set of covariates which, when conditioned upon, make the measurements made in a survey independent of sample selection. This condition is often referred to as "ignorable selection." The validity of ignorable selection depends upon the particular survey measurements of interest and the chosen set of covariates. There is no reason for a set of covariates sufficient to make selection ignorable in one domain should necessarily work in another. When overlapping data are available for the survey measurements and covariates from a sample without selection bias, it is possible to test for ignorability. We propose two statistical tests for ignorability and apply these to economic, political and health surveys from an opt-in internet panel.

CONCURRENT SESSION C-1

SUPPLEMENTING THE COVERAGE OF ADDRESS-BASED SAMPLING FRAMES FOR IN-PERSON HOUSEHOLD SURVEYS

Surveying Young Adults Using an Address List-Based Frame

Luciano Viera, Jr. (Fors Marsh Group), Ricardo Carvalho (Fors Marsh Group), and Sean Marsh (Fors Marsh Group)

The Department of Defense (DoD) Youth Poll is conducted to track attitudes, impressions, and behavioral intentions as they relate to military enlistment among young adults. It has been conducted twice a year beginning in 2002, and used a random-digit-dial (RDD) telephone-based methodology until the 2010 fall administration when it transitioned to a mail-based methodology. One of the primary reasons for this switch was the decreasing coverage of RDD telephone samples. The current Youth Poll frame is an address-based list of youth compiled from multiple publicly available sources.

This purpose of the present study is to examine how well the Youth Poll frame captures 16- to 24-year-old young adults living in the United States. In order to do this, frame counts were compared to estimates provided by Census counts by gender, age and geography. Preliminary results show that the Youth Poll frame covers the vast majority of the target population (92%). With respect to gender, females were better covered than males. With respect to age, coverage was highest in the adult age range (18- to 24-year-olds) and relatively lower in the younger 16- to 17-year-old age groups. By geography, analyses showed that the South and the Midwest were the best-covered regions, followed closely by the Northeast region with West being the least covered region.

Although not perfect, the address-list based frame represents a marked improvement in coverage of the young adult population over existing RDD methodologies, which do not capture the steadily growing number of cell phone–only households. Indeed, recent estimates indicate that 49% of 18- to 24-year-olds live in households with only a cell phone, a group that is not covered by traditional RDD survey methodologies (Blumberg & Luke, 2012). Results and implications for existing survey practice along directions for future research will be discussed.

Correcting Coverage Deficiencies in Address-Based Frames: The Use of Enhanced Listing

Ned English (NORC at the University of Chicago), Colm O'Muircheartaigh (Irving B. Harris Graduate School of Public Policy Studies at the University of Chicago), and Katie Dekker (NORC at the University of Chicago)

Sampling frames based on the United States Postal Service delivery-sequence file (DSF or CDSF) have become standard for single- and multi-mode address-based studies, at least in urban and suburban areas. One limitation of the DSF is the presence of non-city-style addresses in rural areas, which reduces its utility as a frame of dwellings. Researchers therefore need to decide which areas have sufficient coverage from the DSF vs. those that require another method of frame construction prior to sample selection. Fortunately, most areas in the USA have some presence of geocodable city-style addresses from the DSF which may be used as a basis for sampling frame construction. At NORC we employ a method known as "enhanced" listing, where we have staff called "listers" augment and edit address lists that have been determined to be insufficient a priori. Our paper first describes our method of determining if an area is sufficiently covered by the DSF so that it may serve as a sampling frame alone. We then detail the "enhanced listing" approach, and the technology employed in the field to enable the process. Lastly we present results of a comparison between enhanced and traditional listing methods across environments.

Using an Area Linkage Method to Improve the Coverage of ABS Frames for In-Person Household Surveys

Sylvia Dohrmann (Westat) and Richard Sigman (Westat)

This paper presents an approach for using USPS lists in multistage-area-sample surveys of households. This approach tackles coverage and geocoding issues by defining two types of segments: area segments and list segments. An area segment is a parcel of land defined by Census geographic boundaries. A list segment is the set of addresses that geocode to a particular area segment, using street-level geocoding. Using the list segments as the survey's secondary sampling units avoids the issue of inaccurate geocoding although while street-level geocoding may not place every address accurately, it does assign every address on the USPS list to some list segment, and thus gives all USPS addresses a chance of selection.

The issue of coverage is addressed by adding an address coverage enhancement procedure. Within a subsample of area segments linked to the survey's sampled list segments addresses present in the area segment but not in the associated list segment are identified. After confirming that the identified addresses are not anywhere on the USPS list, they are given a chance of selection for the survey. The aim is to give the addresses identified and sampled by the address coverage enhancement procedure the same selection probabilities as addresses sampled from the USPS list. However, practical aspects of the procedure may result in deviations from equal probability to balance the cost of implementation. We describe the address coverage enhancement procedure's sampling concepts and field procedures.

The CHUM: An Adaptation of the Half Open Interval Procedure for Use with ABS Frames

Joe McMichael (RTI International), Bonnie Shook-Sa (RTI International), and Jamie Ridenhour (RTI International)

To compensate for the undercoverage of housing units by Address Based Sampling (ABS) frames used for in-person surveys, RTI developed and tested the Check for Housing Units Missed (CHUM) procedure (McMichael et al. 2008). The CHUM is similar in concept to the familiar half-open interval (HOI) procedure in that the interviewers search the selected HU and the prescribed area up to the next HU on the frame. Interviewers also check a subset of sample census blocks so that housing units in blocks with no city-style addresses on the CDS have a chance of selection. The CHUM systematically identifies HUs missing from the ABS frame, giving each one a chance of selection with known probability whether it is missing due to non-city-style addresses, geocoding error, or new construction. CHUM is most effective when monitored and conducted in a separate field visit from the survey interviewing. It is less costly than enhanced listing because only small portions of the geographical areas are searched, while still giving all housing units a known chance of selection through the corresponding sample HUs and subsampled blocks. Because it is conducted after HUs are selected, not at the frame-building stage, the coverage is as up-to-date as the data collection period itself.

CONCURRENT SESSION C-2

PRODUCTION AND EVALUATION OF SUBNATIONAL ESTIMATES

Small Area Estimation: New Developments and Directions for Health and Human Services Data

John Czajka (Mathematica Policy Research), Amang Sukasih (Mathematica Policy Research), and Susan Queen (Department of Health and Human Services)

Small area estimation (SAE) methods have advanced substantially in the past three decades, yet the number of ongoing applications in the federal government is quite small. Within the Department of Health and Human Services (HHS), there are fewer than half a dozen individual applications that produce annual estimates. Yet SAE methods represent an important tool in developing new data sources and enhancing the value of existing data sources. To inform staff in the HHS operating divisions about the potential of

SAE methods for addressing data needs and to identify ways to facilitate greater use of such methods within the department, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) convened an expert panel meeting in the spring of 2013. The panelists included representatives of the HHS operating divisions and selected statistical agencies with active SAE programs, representatives from the Government Accountability Office and the Office of Statistical Policy in the Office of Management and Budget, and academic researchers. Panelists were invited to share their insights from the application of SAE methods, including the identification of suitable techniques, the choice of data sources, communication of findings, and issues encountered in implementing the methods. This paper summarizes the panel discussion and outlines a number of steps that ASPE and other federal entities can take to facilitate more effective use of SAE methods within the federal government. The paper is based on a report that was prepared in conjunction with the panel meeting.

Challenges of Using Prediction Models to Produce Nationally Representative Estimates of Mental Illness

Sarra Hedden (Substance Abuse and Mental Health Services Administration), Joseph Gfroerer (Substance Abuse and Mental Health Services Administration), Jonaki Bose (Substance Abuse and Mental Health Services Administration), Phillip Kott (RTI International), Dan Liao (RTI International), and Lisa Colpe (The National Institute of Mental Health)

Using the National Survey on Drug Use and Health (NSDUH), SAMHSA has been producing model-based prevalence estimates of past year serious mental illness (SMI) and any mental illness (AMI) among adults aged 18 or older since 2008. These model-based estimates are calculated from individual responses to two short scales (past year K6 distress scale and past year abbreviated WHODAS disability scale) in the NSDUH. Specifically, a prediction model of mental illness was developed using adult respondents' K6 and WHODAS scores. This model was then applied to all NSDUH adult respondents to produce annual SMI and AMI estimates. The methodology allows SAMHSA to take advantage of the large NSDUH sample size for estimating SMI and AMI, using the small clinical interview sample to create a prediction model that links the short scale data from the full NSDUH sample to the actual diagnostic assessments. This presentation will the focus on the methods developed to produce the NSDUH model-based estimates of mental illness, and the issues we considered when developing these methods. Particularly, the challenges of producing unbiased estimates and trends for the overall adult population as well as for subpopulations, and the error incurred when producing model-based estimates will be discussed.

Small Area Modeling of County Estimates for Corn and Soybean Yields in the U.S.

Matthew Williams (National Agricultural Statistics Service)

The National Agricultural Statistics Service (NASS) conducts the County Agricultural Production Survey (CAPS) to establish annual county-level acreage and production in the U.S. The data collected from this survey are combined with the year-end production surveys to improve the accuracy and defensibility. The estimates produced from these surveys are extremely important to administering programs within the United States Department of Agriculture as well as to the general public. The goal of this research is to evaluate the utility of a class of modeling techniques known as *Small Area Models* to add value to the county estimation process at NASS.

We apply several of these methods to estimate corn and soybean yields for 3 states in 2010 and for 10 states in 2011 and 2012. Comparing performance of estimators to the official published yield, the models show little improvement over the survey indications. Possible explanations include model misspecification, poor or noisy covariate information, the shrinkage nature of the model estimates, and the use of the official published yields as the standard for comparison. Current plans are to compare the performance of the estimators to results from the 2012 Census of Agriculture. Related research into benchmarking and alternative covariate and variance structures continues in preparation for this comparison.

A Comparison of Small Area Models Used in the Quality Indicator Program Sponsored by Agency for Healthcare Research and Quality

Robert Baskin (Agency for Healthcare Research and Quality), Pamela Owens (Agency for Healthcare Research and Quality), and Christopher Sroka (Batelle)

The Agency for Healthcare Research and Quality (AHRQ) Quality Indicators™ include a set of measures that can be calculated from hospital administrative data for quality improvement and comparative reporting. The Prevention Quality Indicators (PQIs), one of four modules of the AHRQ QIs, identify hospital admissions in a geographic area that evidence suggests may have been potentially avoided through access to high-quality outpatient care. The current PQIs adjust to general populations in counties but not to domain specific populations. For example, diabetes or asthma PQIs are adjusted to the general population of the county instead of the condition-specific populations since those estimates are not currently available at the necessary level of geographic detail.

The objective of this study is to evaluate domain specific models for estimating county-level counts of persons with specific conditions measured by the PQIs (e.g., diabetes or asthma). Two sources of data are available for model building but neither covers all geography consistently. The Behavioral Risk Factor Surveillance System (BRFSS), conducted by the Centers for Disease Control and Prevention, is a telephone survey of health risk factors has been used to produce some county level estimates. The Medical Expenditure Panel Survey – Household Component (MEPS-HC), conducted by the Agency for Healthcare Research and Quality, is a household level survey of medical conditions, usage, and expenditures that has also been used to produce some county level estimates. Small Area Models will be produced on the BRFSS data and compared to available information in the MEPS-HC. These models will be fit using a hierarchical Bayesian model and estimated using both MCMC sampling and in some cases Laplace approximation of the posterior. A further investigation of benchmarking the models to national estimates will be begun.

CONCURRENT SESSION C-3

NOVEL APPROACHES TO CODING, EDITING, AND IMPUTING DATA

Issues in Coding Hospital Intensive Care for Data Analysis

Adeline Wilcox (Department of Veterans Affairs)

Part of the American Recovery and Reinvestment Act (ARRA) of 2009, the Health Information Technology for Economic and Clinical Health Act (HITECH), offers incentive funding for Meaningful Use (MU) of electronic health records (EHRs). Reporting of Clinical Quality Measures (CQMs) to the federal government is a MU of EHRs. For the measure called ICU Venous Thromboembolism (VTE) Prophylaxis, ICU admissions must be identified. The Veterans Health Administration (VHA) Corporate Data Warehouse (CDW) holds several columns describing intensive care. These data are loaded from its legacy EHR called the Veterans Health Information Systems and Technology Architecture (VistA). Analysis of these unstructured free text data is all but impossible.

Descriptive analyses show data from CDW columns named Specialty, BedSection, WardLocation and others. Relevant metadata for these columns are summarized. Coding obstacles are illustrated. Several ontologies with codes for intensive care are discussed, the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT) among them. Characteristics of hospitals, wards, rooms, beds, and physician specialty have been used to classify intensive care. In contrast, the National Health Service Data Model and Dictionary for England gives definitions describing the care provided to the patient. None of the ontologies lay out a ready means for coding ICU data from the VHA CDW.

Statistical Analysis of Text in Survey Records

Wendy Martinez (Bureau of Labor Statistics) and Alex Measure (Bureau of Labor Statistics)

This presentation highlights ongoing research to facilitate human coding of data collected through the Survey of Occupational Injuries and Illnesses (SOII). This is an annual survey of U.S. establishments, where data are collected on work-related injuries. At least 5 codes are extracted from each case; these include occupation, nature of the injury, event that caused the injury, part of the body that was affected, and object or substance that directly caused the injury. This coding is currently being done by humans; so, it can be very time-consuming. Additionally, the text we receive is often confusing and unclear, making the job more difficult.

The BLS is exploring ways to assist human coders by using statistical methods for document classification and clustering. In classification (or supervised learning), one starts with a training set, where class labels are available. These training data are used to build a classifier for coding future documents. Clustering (or unsupervised learning) is concerned with putting unlabeled documents into groups with like topics (or events).

We discuss two bag-of-words approaches for converting text documents to numbers, which would then allow us to employ statistical learning methods like classification and clustering. These are the term-document matrix and the bigram proximity matrix. We demonstrate how these bag-of-words methods can be applied to SOII injury reports. We then describe model-based clustering and show how this can be used to find groups of documents with similar characteristics. The hope is that these clusters can be used to facilitate human coding of injury reports.

Evaluation of a New Edit Methodology for the Common Core of Data Nonfiscal Surveys

Elizabeth Goldberg (U.S. Census Bureau), Robert Stillwell (National Center for Education Statistics), and Jeffrey Little (U.S. Census Bureau)

The Common Core of Data (CCD) nonfiscal surveys consist of data submitted annually to the National Center for Education Statistics (NCES) by state education agencies (SEAs) in the 50 states, territories, and other agencies. CCD survey staff edits the data to produce a clean data file, which NCES uses to construct general-purpose publications and specialized reports. The principal users of CCD nonfiscal data are the federal government, the education research community, state and local government officials, including school boards and LEA administrators; and the general public. With the 2010-11 survey year, public concerns were raised about the potential for extreme erroneous results to not be flagged in the editing process. As a result, a new methodology of editing based on comparing multiple years of data for an edited data element vs. the current year/prior year editing process previously used. This paper discusses the results of this new edit methodology on the data quality for the 2011-12 CCD survey year, as well as the potential for this method to be applied in other ways moving forward.

Banff Automated Edit and Imputation Applied to the U.S. Hog Inventory Survey

James Johanson (National Agricultural Statistics Service)

The National Agricultural Statistics Service (NASS) conducts surveys within the U.S. Department of Agriculture. Traditionally, editing and imputation are done manually for many surveys via Blaise, which is labor intensive. To reduce costs, NASS purchased Statistics Canada's Banff system for automated statistical editing and imputation. This system was applied to the U.S. hog survey for December 2012. After data collection and before any manual editing, the original data were processed through the Banff editing code, consisting of only commodity-specific edits.

To promote transparency and acceptance by analysts, the Banff imputed values were then incorporated into selective editing methodology and displayed in a GUI interface. Observations can be sorted by selective edit unit score and reviewed. Researchers reviewed the largest 20% of unit scores and made changes where appropriate.

In the operational environment, analysts will review and edit records, during data collection. To simulate this practice, cumulative daily batches were run through Banff. Once a record passed Banff edits without error, that record was not changed again and was eligible to be a donor in subsequent batches.

Records still in error were given the values derived from the manual editing for calculating estimates to mimic procedures, if Banff was implemented into the operational procedures. Four sets of estimates were calculated and compared from the traditionally editing data, the Banff only data, the final daily batch, and the Banff plus selective editing.

Boosting Algorithms for Edit and Imputation of Multiple-response Variables

Ping Li (Cornell University) and John Abowd (Cornell University)

Many practical classification datasets naturally have multiple responses of the form of "check all that apply" variables. Such variables are also known as "multi-label." A prominent example is the multi-label race variable mandated by the OMB standards for race and ethnicity reporting that is used by the Census Bureau and many other federal agencies. This response is frequently reported as "white alone," "black or Afro-American alone," "white and some other race," *etc.* An additional example occurs in text categorization where a document may belong to several pre-specified topics, such as sports and news. In medical diagnosis, a disease could belong to multiple categories. In general, predicting multiple-response variables is a more challenging task than regular single-response classification. In this paper, we propose a simple statistical framework that generalizes the standard multinomial logistic regression model to accommodate multiple-response variables. We do this without artificially expanding the classification variable to all feasible, as opposed to observed, permutations of the multiple responses. This is the basis for the prediction performance gains we achieve. Our algorithm, implemented based on boosting and trees, is tested on artificial data and on several standard multi-label datasets, including public-use decennial census data. The principle advantage of our approach is that it allows for a more accurate prediction model that can be used as part of the edit and imputation processing for surveys that contain multiple-response variables.

CONCURRENT SESSION C-4

EVALUATING THE IMPACTS OF MODE SELECTION IN SURVEYS

An Analysis of the Mixed Collection Modes for Business Surveys at the U.S. Census Bureau

Broderick Oliver (U.S. Census Bureau) and Katherine Jenny Thompson (U.S. Census Bureau)

Business surveys conducted by the U.S. Census Bureau generally make initial contacts with the respondents by mail and provide a variety of response options (mail, fax, secure internet collection, and phone). There are many hypotheses about the merits of each collection method: mail collection can be more expensive for the data collection agency than telephone, fax, and internet collection. Telephone and internet collection can provide better quality data than mail and fax collection because the opportunity for checking data quality during the data collection process reduces the needs for later data edit and revision. Using historic data from three ongoing programs, we examine unit response rates by collection mode as well as associated quality measures and indicators for key items by collection model over the same time periods.

Truth or Consequences: The Use of Benchmarks in Calibrating Health Research

Victor Lange (Marketing, Inc.), Steven Gittelman (Marketing, Inc.), and Elaine Trimarchi (Marketing, Inc.)

The general question of perceived quality of health is likely to be a core question in evaluating health programs of the future. Three "gold standard" studies lend their hand at measurement and yet produce different between them. The NHIS shows health improving over time, the BRFSS shows health declining, and the CPS shows a path right down the middle. The modes of these studies are different. Thus understanding the impact of modes is again brought to the forefront.

Social desirability bias, acts independently of a properly distributed sample frame. The interviewer-respondent interaction may make some question types and subject verticals particularly vulnerable. Here we provide an analysis of critical questions drawn from a battery of national health studies collected through both online and telephone. Our sample is drawn from data collected in four target states during December 2012. Many measures of health, particularly emotional health, are so different between the two sources that post-stratification has no chance of correcting the disparities. A traditional viewpoint would be to discredit online data in favor of phone. However, some research exists (e.g. Link & Mokdad, 2005) to suggest that in cases where social desirability is a factor, online data may be more representative of true values due to the absence of a human interviewer.

In order to establish this, we treat a well-conducted mail survey as the arbiter, lacking either the coverage bias of online research or the desirability bias of phone. With a more complete replication of the BRFSS having now been conducted using this mode, we can now begin to answer conclusively which areas online truly produces less representative data, and in which our impressions of American health have been biased by modal effects. No mode appears to be the perfect fit for all questions, suggesting that multi-mode studies should be executed where questions are fitted to the mode.

Mode Selection and Recruitment Strategies for Survey Respondents with Disabilities

Ashley Schaad (ICF International), Amy Falcone (ICF International), Jessica Jagger, (U.S. Maine Corps) and Ashleigh Davis (ICF International)

Several general population studies have examined the impact of survey mode options offered on response rates, but little or no research has examined this among individuals with disabilities. As mode preference in this population may at least in part be dictated by the relative accessibility of a particular mode, and individuals with disabilities are likely present in the target population of the majority of survey studies, it is important to better understand which data collection mode(s) are preferred in this population. The use of multi-mode surveys may increase accessibility, as self- and interviewer-administered survey modes each present unique accessibility challenges and benefits. The current study examines mode selection, either web or phone, among respondents in a survey for the Library of Congress National Library Service for the Blind and Physically Handicapped (NLS). Mode selection based on user characteristics (e.g., age, primary disability, internet usage) is also examined. Additionally, this study explores the effectiveness of multiple survey recruitment strategies among individuals with disabilities, particularly the blind and low vision population. Recruitment for the survey utilized traditional print outreach, as well as online and social media outreach. The use of both print and online recruitment techniques allows for an analysis of the effects of survey recruitment strategy on mode selection as well. All analyses employ frequency and cross-tabulation statistics. The findings provide much-needed data on survey mode preference and survey recruitment strategies among researchers whose target population includes the 18.7% of Americans with disabilities (Brault, 2012).

Great Expectations: Changing Mode of Survey Data Collection in Military Populations

Ronald Z. Szoc (ICF International), Jacqueline Pflieger (ICF International), Frances M. Barlas (ICF International) and Randall Thomas (GfK Custom Research, LLC)

As difficulties and costs of "in person" surveys have escalated, many studies have migrated to new modes of data collection. Most research concerning the differences in prevalence rates for health-related behaviors, such as alcohol use, as a function of data collection mode, has been done on civilian populations. We report results of a study of the effects of mode on such prevalence rates for a military population comprised of National Guard and Reserve units from all four military Department of Defense branches. Since 1980, the DoD Health Related Behaviors (HRB) surveys have been conducted onsite at military installations worldwide, with respondents completing paper questionnaires in group-administered, in-person sessions. For the most recent HRB National Guard and Reserve survey, we conducted a parallel study to assess the feasibility of conducting the HRB as a web-based, individually administered survey - 16,079 completed onsite and 18,895 completed online (response rates of 64% and 33% respectively). Controlling for sample differences, this afforded us the ability to compare self-reported prevalence estimates for 15 key risk and protective measures, such as alcohol use and BMI, between onsite and online modes. Two primary research questions were: (1) What differences were observed in

behavioral health estimates as a result of mode of administration? and (2) Will changing modes retain the ability to make comparisons across services and still maintain trends over time? The results, from hierarchical regression analyses, were different from what might be expected from similar studies with civilian populations - online respondents reported significantly lower prevalence rates for risky behaviors than their onsite counterparts, while prevalence rate rank order was similar across modes. Thus, contrary to civilians, military members may be less likely to report engaging in risky behaviors in an online mode, and may require additional assurances of data anonymity for online surveys.

CONCURRENT SESSION D-1

COMPARISONS OF VARIANCE ESTIMATION TECHNIQUES

Investigation of Variance Estimators for the Survey of Business Owners (SBO)

Marilyn Balogh (U.S. Census Bureau) and Sandy Peterson (U.S. Census Bureau)

The Survey of Business Owners and Self-Employed Persons (SBO) provides, on a quinquennial basis, selected economic and demographic characteristics for businesses and business owners by gender, ethnicity, race, and veteran status. The SBO sample is a stratified systematic sample that employs the random group variance estimator to estimate sampling error. This paper reports the results of a simulation study conducted to compare the random group variance estimator to the delete-a-group jackknife variance estimator and the stratified jackknife variance estimator. The methods were compared by examining the mean squared error (MSE) of each variance estimate for five states (Florida, Georgia, Kansas, New York, and North Dakota) for different characteristic data (public, race, ethnicity, sex, and veteran status). Another factor taken into consideration was computer processing time, which is not trivial for the stratified jackknife applied to a large survey. The results of this study suggest that the delete-a-group jackknife variance estimator may provide an improvement over the random group estimator in terms of MSE, with little differences in processing time.

Inference of Domain Parameters by Using a Dynamic Jackknife Variance Estimator

Sixia Chen (Westat) and Tom Krenzke (Westat)

In this paper, we explore adjustments that can be made with small domain sizes to obtain confidence intervals with coverage probabilities closer to the nominal values. A jackknife replication variance estimator is evaluated for the inference of domain parameters. A dynamic subset of the original replication weights is used to construct the variance estimator for the estimates. The degrees of freedom are adjusted according to the membership of sampling units. The proposed method is compared with Taylor series linearization and traditional jackknife replication variance estimators that do not take the dynamic structure into consideration. We further extend the proposed method under two-stage sampling design. Two limited simulation studies based on 2011 National Health Interview Survey (NHIS) public use data are presented in the paper, which evaluates the proposed method in terms of coverage rate and confidence interval length.

Using Reimputation Methods to Estimate the Variances of Estimates of the ACS Group Quarters Population with the New Group Quarters Imputation Methodology

Michael Beaghen (U.S. Census Bureau)

The Census Bureau has implemented a new imputation program to enhance American Community Survey (ACS) estimates of the group quarters (GQ) population for small areas starting with the 2011 ACS 1-year, 2009-2011 ACS 3-year, and 2007-2011 ACS 5-year estimates. This method presents novel challenges for variance estimation because it is a mass imputation of whole-person records with roughly as much imputed data as sampled data.

In previous years the ACS used successive differences replication to estimate variances. However, naively applying to this method to data augmented by imputation leads to a serious underestimation of variances. Hence, we implemented a method that applies inflation factors to the replicate weights. This method is understood to be coarse but is considered adequate.

The Census Bureau has been searching for a more suitable variance estimation methodology. Reimputation with replication methods of variance estimation may yield improved estimates of variance in the context of the current methodology. The basic method of reimputation is that for each replicate or pseudo-replicate one reimputes for the missing data independently from the restricted donor pool defined by the replicate or pseudo replicate. We choose to reimpute with random group replicates, as forming random groups is tractable with the ACS GQ data.

In this study we assess the feasibility and soundness of using reimputation methods to estimate variances for the 2009-2013 ACS 5-year, 2011-2013 ACS 3-year, and 2013 ACS 1-year estimates of the GQ population. We also evaluate the current variance estimation methodology, successive differences replication with inflation factors. We compare the variances for ACS production and alternative variance estimation methods for the 1-year ACS estimates for the states and the 5-year ACS estimates for counties and tracts.

Comparing Generalized Variance Functions to Direct Variance Estimation for the National Crime Victimization Survey

Bonnie Shook-Sa (RTI International), Lance Couzens (RTI International), Rick Williams (RTI International), and Marcus Berzofsky (RTI International)

Currently, the National Crime Victimization Survey (NCVS) relies on generalized variance functions (GVFs) for the calculation of standard errors and for significance testing. However, GVFs developed for the NCVS are cumbersome, do not allow for complex analyses such as regression modeling, and the accuracy of GVF estimates for outcomes not included in developing the GVF parameters is unknown. Use of GVFs requires knowledge about the correct GVF parameters and formulas to use, and these decisions are dependent on the outcome of interest.

Direct variance estimation techniques such as Taylor Series Linearization (TSL) and Balanced Repeated Replication (BRR) allow variances to be calculated using existing software packages, making estimation more straight forward for most users. Both estimation techniques require study design data (i.e. stratification variables and primary sampling units) in either the creation of the weights (BRR) or in the variance estimation itself (TSL), so resulting estimates accurately reflect the complex survey design. While the NCVS public use file contains some design variables, the full set of variables are not publically available due to disclosure concerns. This paper presents the first evaluation of the accuracy of direct variance estimates based on the available design variables and addresses logistical challenges imposed by direct estimation techniques, specifically those encountered when estimating victimization rates based on multiple input files and sampling weights.

We discuss the complexities associated with calculating direct variance estimates for the NCVS and compare direct variance estimates (TSL and BRR) to estimates produced using GVFs. We evaluate these methods for multiple outcome types (e.g. totals and rates), subgroups of interest (e.g. gender, race, and age), and for single and multi-year estimates. Additionally, we develop recommendations for users of the NCVS public use files regarding NCVS variance estimation.

CONCURRENT SESSION D-2

IMPROVING SURVEY DESIGN AND ESTIMATES USING EXTERNAL SOURCES

Analysis of Housing Square Footage Estimates Reported by the AHS and SOC

Angela Delano (U.S. Census Bureau)

Government agencies, policy analysts, financial institutions, and manufacturers, to name a few, use data collected by the Census Bureau to measure and evaluate size, composition, and change occurring within the inventory of residential housing units available in the United States. This endeavor uses estimates from two surveys - the American Housing Survey (AHS) and the Survey of Construction (SOC).

The AHS is a longitudinal survey conducted every two years in odd numbered years. This survey interviews occupants of existing housing units regarding housing characteristics and demographics. The SOC is a survey conducted on a monthly basis. This survey interviews builders or owners of new constructed housing units from start of construction through completion or occupancy of the housing unit. The SOC collects information concerning construction timelines – construction start dates, completion dates, and sale dates if applicable – and housing characteristics. Both the AHS and the SOC collect similar housing characteristics such as the house size, or square footage, number of stories, and number of bedrooms.

This paper will examine housing square footage estimates reported by the AHS and the SOC at varying geographies. The data sources used to obtain these estimates will consist of the 2011 AHS National Microdata file and the 2002-2011 SOC Microdata files. Highlights from this paper will include analysis of issues encountered with collection of housing square footage between the AHS and the SOC. These issues include differences related to data collection, sample sizes, and square footage definitions. These issues may suggest reasons for square footage differences reported by the AHS and the SOC. Additional discussion will take place related to post data collection issues specifically related to the SOC. Results from this paper will be used to propose future enhancements to the collection and handling of housing square footage estimates by the SOC.

Household Estimates Conundrum: Effort to Develop More Consistent Household Estimates Across Current Surveys

Arthur Cresce Jr. (U.S. Census Bureau), Yang Cheng (U.S. Census Bureau), Andrew Zbikowski, (U.S. Census Bureau) and Christopher Grieves (U.S. Census Bureau)

The Census Bureau over the last several decades has had to deal with the problem of producing inconsistent estimates of households from its current surveys. This is a particularly difficult problem given that some surveys produce estimates of households based on responses to a population question (relationship to householder) and others produce estimates of occupied housing units (households) based on housing data. As a result, survey estimates of households from our current surveys for the same time period can differ by several million. These differences, in turn, raise questions among data users as to which represents the “true” estimate of households. This paper examines the sources of these differences and makes recommendations for arriving at more consistent estimates.

Estimating Substance Abuse Treatment: A Comparison of Data from a Household Survey, a Facility Survey, and an Administrative Data Set

Joe Gfroerer (Substance Abuse and Mental Health Services Administration), Jonaki Bose (Substance Abuse and Mental Health Services Administration), Deborah Trunzo (Substance Abuse and Mental Health Services Administration), Alex Strashny (Substance Abuse and Mental Health Services Administration), Kathy Batts (RTI International), and Michael Pemberton (RTI International)

This conference presentation and paper will discuss the comparison of data on treatment for alcohol and illicit drug abuse problems obtained from the National Survey on Drug Use and Health (NSDUH), Treatment Episode Data Set (TEDS) and the National Survey of Substance Abuse Treatment Services (N-SSATS). The comparisons are done at the national and state levels and reasons for similarities and differences are discussed. The paper also discusses the implication of the differences for data analysis and making inferences. In addition to providing information about the results of this study, this paper also intends to highlight methodological challenges in comparing data from studies that have differing purposes, different data collection methods and different units of analysis, e.g., persons receiving treatment versus admissions versus single-day census.

A Comparison of Consumer Expenditure Surveys

Nhien To (Bureau of Labor Statistics) and Brett McBride (Bureau of Labor Statistics)

The Consumer Expenditure (CE) Survey is in the initial stages of a major redesign. The overall mission of the redesign is to improve data quality through a verifiable reduction in measurement error. As part of this effort, the CE program studied the design features of consumer expenditure surveys conducted in other countries. The survey design elements were gathered from a number of sources including other countries’ consumer expenditure program websites, methodology reports, and e-mail correspondence with representatives from the survey programs. The surveys included were selected based on the diversity of their characteristics and the accessibility of information about their design features.

This presentation will provide key findings from the research, including an overview of the key design features of these surveys and, in particular, noted unique features that other surveys successfully implemented. One interest of the study was to identify which other countries had recently redesigned their

surveys and how. Obtaining information on their experiences and design features provided valuable guidance as the United States CE program considered which features would be beneficial for a redesigned CE survey.

This study noted that, although there are many shared features between the United States CE survey and those of other countries, the CE program should consider moving from independent samples for in-person interview and self-administered diary surveys to a single sample for both, which was a common feature of other countries' surveys. Based on the initial research done by this study, it was further recommended that the CE program expand its consultations with representatives from survey programs that use design features such as individual diaries, incentives, new data collection technologies, and multi-mode options. By doing so, the CE program could learn more about their experiences designing and testing the features, and their lessons learned about how to successfully implement them.

Comparing BLS's Price Indexes with Industry Sources

Rachel Soloveichik (Bureau of Economic Analysis) and Paul Sullivan (Bureau of Labor Statistics)

We compare Bureau of Labor Statistics (BLS) price indexes for entertainment with prices for the same products reported by industry sources. We find that industry sources often differ dramatically from BLS's price indexes. For example, BLS's price index for DVD's (CUUR0000SS31022) dropped by 24% from 2006 to 2011. Over the same time period, The-numbers.com reports steady prices for newly released DVD's. Furthermore, BLS's price index for purchased music like CD's (CUUR0000SERA06) dropped 8% from 1998 to 2011. Over the same time period, the Recording Industry Association of America reports a 4% rise. In contrast, we find that BLS's price index for cable (CUSR0000SERA02) reports higher inflation rates than the National Cable and Telecommunications Association.

We are currently studying BLS's confidential micro-data to determine why industry sources behave differently than BLS's price indexes. We will have preliminary results by July.

CONCURRENT SESSION D-3

RECENT FINDINGS AND INNOVATIONS IN UTILIZING PARADATA AND METADATA

Metadata Standards and Technology Development for the NSF Survey of Earned Doctorates

Kimberly Noonan (National Science Foundation), Pascal Heus (Metadata Technology North America), and Tim Mulcahy (NORC at the University of Chicago)

Generically, metadata consist of a variety of information about the data, providing essential knowledge to maximize data usefulness. Survey metadata contain information about the survey instrument, such as question wording and answer categories; microdata variables, such as name, format, valid values; imputation, such as flags, methods; data sets, names, dates, owners; and much more. Collecting, documenting, organizing, and storing metadata is essential to the appropriate use and interpretation of survey data. These activities can however require significant resources and optimizing the related processes is therefore highly desirable.

The National Center for Science and Engineering Statistics (NSCES) at the National Science Foundation (NSF), NORC, and Metadata Technology North America (MTNA) are collaborating on a project to develop new technologies to capture comprehensive metadata, automate the production of essential documentation, and generate an archival package for the annual NSF Survey of Earned Doctorates (SED). The work on this project links to a larger project where NCSES is developing a unified data system to store and disseminate all NCSES survey data.

The SED Metadata project builds on the Data Documentation Initiative (DDI)© metadata specification, globally recommended practices, and industry standard information technologies to deliver metadata content in a consistent, timely, cost efficient, automated process. In this presentation, we will give an

overview of the project, lessons learned, potential reuse of the tools and procedures, next steps, and a brief demonstration of the application.

Using Paradata to Model Nonresponse in the Current Population Survey

John Dixon (Bureau of Labor Statistics)

This study examines nonresponse bias using paradata from the Current Population Survey (CPS). The CPS is administered to a household for 4 consecutive months, followed by a break of 8 months, and then interviewed another 4 consecutive months. These 8 interviews (panels) form the basis of this analysis. Estimates of unemployment may be affected by nonresponse, sample attrition, and movers. If those households which initially don't participate (either refuse or couldn't be contacted) are less likely to be unemployed, then estimates would be inflated. If those who drop out of the survey are less likely to be unemployed, then the estimates would also be inflated. If households which move out of the sample because they are unemployed are replaced by households which are employed or not in the labor force, then estimates of unemployment would be deflated. Paradata, including respondent contact history recorded by interviewers, may help understand those effects. For example, using propensity models to predict which of those who respond to the survey are similar to nonresponders, dropouts, and movers, then the different potential biases can be estimated.

Using Paradata to Understand Business Survey Reporting Patterns

Eric Fink (U.S. Census Bureau) and Joanna Fane Lineback (U.S. Census Bureau)

Paradata are increasingly used as a source of information about respondent behavior and survey outcomes. In this paper, we examine paradata from the 2010 and 2011 Annual Survey of Manufactures, including cost, follow-up and multi-unit electronic reporting instrument data.

Previous research was exploratory in nature, attempting to understand basic questions such as the median length of time from downloading to uploading the survey's software. Current research will be more targeted, whereby we use auxiliary data, including paradata, to model survey response. Additionally, we begin examining ASM costs as they related to the stages of the survey process.

What Can Paradata Tell Us About Business Reporting in the 2012 Economic Census?

Elizabeth Hoeffel (U.S. Census Bureau), Gianna Dusch (U.S. Census Bureau), and Joanna Fane Lineback (U.S. Census Bureau)

Survey paradata are increasingly used to understand respondent behavior and survey outcomes. We analyzed paradata to examine electronic reporting patterns of single-establishment businesses that used the web for the 2012 Economic Census. These paradata include the timing and frequency of user actions. In addition, unlike many surveys of individuals or households, we also have business characteristics from our business register. In this paper, we put together these data sources to examine web survey behaviors. We use our findings to explore ideas for using these paradata to influence data collection activities.

CONCURRENT SESSION D-4: *TECHNICAL DEMONSTRATIONS*

ADVANCES IN UTILIZING MAPS AND VISUALIZING DATA

A Demonstration Occupational Employment Statistics' Web-Based Chart and Map Visualization Tools

Tanner Beam (Bureau of Labor Statistics)

The Occupational Employment Statistics (OES) survey is a semi-annual survey of approximately 200,000 nonfarm business establishments which results in annual wage and employment estimates by occupation for approximately 1.2 million business establishments. In this presentation we illustrate the various tools OES has created for website users to visualize the large amount of data produced by the program.

OES produces wage and employment estimates for hundreds of occupations, industries, and geographic areas (all of which have different levels of aggregation, to complicate matters further) so developing tools that are both intuitive and informative has been a challenge. We discuss some of the strategies for dealing with these challenges, including how to best organize the data to illustrate patterns and maintaining continuity in the face of regularly changing occupational, industrial, and geographic classification systems.

The examples in the presentation will all utilize OES data, but the concepts will easily translate for anyone that needs to develop strategies for visualizing geographic datasets that contain a large number of variables and categories.

Visualizing Historical Agricultural Data: The Current State of the Art

Irwin Anolik (National Agricultural Statistics Service)

This paper reports on recent research conducted and methods implemented at the National Agricultural Statistics Service (NASS) to use a variety of visual displays of historical agricultural data which enhance data users' ability to view, analyze and interact with this data using any web browser or mobile device.

We discuss concepts and technologies implemented to enhance the ability of NASS's data customers to find historical data of interest, make comparisons, and see important trends across variables that are not possible with static tables or charts.

The visualization methods presented include 1) geographical animations which depict changes over time, 2) sparklines, invented by Edward Tufte and described as "word-sized graphics", employed in small multiples which allow the display of large amounts of historical data on a single web page or mobile device screen, and 3) interactive charts to enable our data users to drill down and focus on more detailed views of these data displays. Each of these methods facilitates the display of large volumes of data and allows data users to extract information from our statistics that is difficult or impossible to obtain from traditional static charts or tabular displays of data.

Hurricane Category Establishment Identification Project (HCat)

Peter Smith (Bureau of Labor Statistics)

This project identifies businesses along the Atlantic and Gulf of Mexico coasts that lie in areas at risk of flooding due to hurricane related storm surge. Using establishment data from the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW) program and hurricane storm surge zones from the National Hurricane Program (NHP), a joint initiative involving the United States Army Corp of Engineers, the Federal Emergency Management Agency and the National Weather Service, the project has produced a new subset of data which includes businesses, their industry, ownership, employment level and wage level data, as well as an indicator of the severity of hurricane that would put the business at risk for storm surge flooding.

The project uses ArcGIS software to integrate the QCEW data and the NHP hurricane zones. The resulting output can be designed into maps for a visual display of the data, which is helpful to emergency planners and media who prefer a quick and basic explanation of the data. There is also an excel spreadsheet produced for each county along the coast that includes the economic information for each business establishment within the county as well the code for what severity of hurricane could bring the potential of storm surge flooding. This data, which are confidential, are used by BLS sample frame users, who will then be able to identify potential survey non-respondents whose businesses are interrupted by hurricane-related flooding. They will also be able to factor these flooded establishments into their data estimates.

The aim of the project is to have a dynamic database of establishments at risk of hurricane-related flooding that is easily and quickly accessible. This will be helpful both for agencies in pre-storm planning as well as post storm evaluation processes.

Map Matching and Real World Integrated Sensor Data Warehousing

Evan Burton (National Renewable Energy Lab), Jeff Gonder (National Renewable Energy Lab), Adam Duran (National Renewable Energy Lab), and Eric Wood (National Renewable Energy Lab)

The inclusion of interlinked temporal and spatial elements within integrated sensor data enables a tremendous degree of flexibility when analyzing multi-component datasets. The proposed paper illustrates how to warehouse, process, and analyze high-resolution integrated sensor datasets to support complex system analysis at the entity and system levels. The example cases presented in the paper utilizes in-vehicle sensor system data to assess vehicle performance, while integrating a map matching algorithm to link vehicle data to roads to demonstrate the enhanced analysis possible via interlinking data elements. Furthermore, in addition to the flexibility provided, the examples presented in the paper illustrate concepts of maintaining proprietary operational information (Fleet DNA) and privacy of study participants (Transportation Secure Data Center) while producing widely distributed data products. Should real-time operational data be logged at high resolution across multiple infrastructure types, map matched to their associated infrastructure, and distributed employing a similar approach; dependencies between urban environment infrastructures components could be better understood. This understanding is especially crucial for the cities of the future where transportation will rely more on grid infrastructure to support its energy demands.

CONCURRENT SESSION E-1

WEIGHTING METHODS IN THE PRESENCE OF NONRESPONSE

Evaluation Study of Calibration Estimation for the Survey of Local Government Finance

Elizabeth Love (U.S. Census Bureau) and Bac Tran (U.S. Census Bureau)

Calibration is a weight-adjustment method and can be used to estimate finance population totals. Sample design weights are calibrated using known population totals from the most recent census year. Currently, the U.S. Census Bureau uses calibration for the Annual Survey of Local Government Finance. The estimates for these surveys are population totals aggregated at different levels. The nested table structure provides multiple population totals to use in calibration equations. The goal of this research is to determine which level of aggregation provides optimal estimates. Census data is available for 2002 and 2007. A probability-proportional-to-size (πPS) sample from the 2007 census data was selected without replacement. Estimates from the 2007 sample were based on calibration equations using 2002 population totals at varying levels of aggregation. These 2007 sample estimates were compared to 2007 census totals. Estimates from different levels of aggregation were compared using their mean squared error, the sum of sampling bias and variance. Since the generalized regression estimator is asymptotically equivalent to calibration when linear distance equations are used, other goodness-of-fit statistics such as R^2 were also used to compare results.

Calibration Methods in the Quarterly Summary of State and Local Tax Revenue

Justin Nguyen (U.S. Census Bureau) and Bac Tran (U.S. Census Bureau)

The Quarterly Summary of State and Local Tax Revenue (QTax) was designed to estimate quarterly property, sales, personal income, corporate income, and other taxes for state and local governments. The Bureau of Economic Analysis uses these estimates to develop estimates of the Gross Domestic Product (GDP). In this paper, we discuss the estimation approaches we developed and tested when the new survey sample for the local non-property component produced lower than expected response rates for several quarters. An adaptive design approach was used as we sampled the non-respondents and reconcile with the reported Annual Summary of State and Local Government Finances data. A calibration model and non-response adjustment with statistical analyses were used to produce quality estimations. In this paper, we describe our methodology and provide our results with an interpretation for the four QTax quarters estimations in a given year using big data and tried to model the non-response. A thorough simulated test was done using three methods: (1) Logistic Response Model and Generalized REGression estimator, (2) Calibrated without non-response follow-up data, and (3) Calibrated with non-response follow-up data. The results provided equivalent and superior ability estimates when calibrated with non-response follow-up and thus, this estimation method with item calibration was recommended for the QTax considering its lower response rates. Furthermore, the validation response propensity model and variance estimation are discussed (using calibration weighting to adjust for non-response and coverage errors). This model is limited or was invalid if the missingness is non-ignorable.

Performance of Generalized Regression Estimator and Raking Estimator in the Presence of Nonresponse

Daifeng Han (Westat and University of Maryland College Park), Richard Valliant (University of Maryland), Jill Montaquila (Westat), and Keith Rust (Westat)

Calibration weighting is widely used to decrease variance, reduce nonresponse bias, and improve the face validity of survey estimates. In the purely sampling context, Deville & Särndal (1992) demonstrate that many alternative forms of calibration weighting are asymptotically equivalent, so the generalized regression (GREG) estimator can be used to approximate some general calibration estimators with no closed-form solutions. It is unclear whether this conclusion holds when nonresponse exists and single-step calibration weighting is used to reduce nonresponse bias (i.e., calibration is applied to the basic sampling weights directly without a separate nonresponse adjustment step).

In practice, poststratification (as a special form of the GREG estimator) and raking (as an example of general calibration estimators) are commonly used calibration approaches, but decisions between these estimators are often made ad-hoc based on sample sizes and availability of external data. In this paper, we compare the performance of these estimators by examining their biases, variances, and effective coverage of the confidence intervals. The theoretical work and simulation study demonstrate the need to consider models for both the outcome variable and the response pattern. The model supporting the typical application of raking has main effects only while poststratification (and more general forms of GREG) can include interactions. A framework involving both design-based and model-based thinking is developed to simultaneously evaluate the impact of sampling, outcome variable structure, and nonresponse mechanism.

Since survey practitioners often lack the knowledge of the outcome variables and nonresponse mechanism in real-world surveys, we also develop a diagnostic that helps gauge the potential consequence of choosing an inappropriate calibration estimator. The results of this research will provide guidelines for choosing between the commonly used calibration estimators.

Patterns of Biomarker Participation in the L.A. Family and Neighborhood Survey

Bonnie Ghosh-Dastidar (The RAND Corporation) and Narayan Sastry (University of Michigan)

The objective assessment of respondents' health status in social science surveys has grown rapidly in the past several years, although relatively few studies to date have collected objective measures of children's health. The Los Angeles Family and Neighborhood Survey (L.A.FANS) is a panel survey of children, families, and neighborhoods in Los Angeles County (Sastry et al., 2006). It was designed to study the effects of family and contextual factors on health, development, and well-being of children and adults in Los Angeles. Wave 1 of L.A.FANS was fielded in 2000-2001 and Wave 2 in 2007-2008. L.A.FANS public use and restricted use data (available through ICPSR) have been widely used by researchers to study a large variety of topics (see http://www.lasurvey.rand.org/pubs/ for a comprehensive listing of publications).

All Wave 1 respondents were eligible for re-interview in Wave 2, with in-person interviews for respondents who lived in Los Angeles County and telephone interviews for those who moved away. In the second wave of the Los Angeles Family and Neighborhood Survey, we collected children's height, weight, and blood pressure measurements, spirometry, and blood and saliva samples. The study used three distinct modes of biomeasure collection. L.A.FANS also collected extensive information about dwelling characteristics, interviewer attributes, and neighborhood factors that support the detailed analysis of contact rates and response rates.

We examined rates of respondent participation in the biomeasure collection, differences in participation rates by mode of collection, and the effects of respondents' demographic, socioeconomic, and health characteristics on their participation. Data from the baseline survey and from the Wave 2 fieldwork provide a rich set of covariates that describe respondent and interviewer characteristics and contextual factors. We used the results to develop a set of nonresponse weights for analyses of the biomeasures and conducted a bias-variance assessment of these non-response weights.

CONCURRENT SESSION E-2

ADMINISTRATIVE RECORDS AND SURVEY RESEARCH

The Nature of the Bias When Studying Only Linkable Person Records: Evidence from the American Community Survey

J. David Brown (U.S. Census Bureau), Adela Luque (U.S. Census Bureau), and Amy O'Hara (U.S. Census Bureau)

Record linkage across survey and administrative record sources can greatly enrich data and improve their quality. The linkage can reduce respondent burden and nonresponse followup costs. This is particularly important in an era of declining survey response rates and tight budgets. Record linkage also creates statistical bias, however. The U.S. Census Bureau links person records through a person validation

process that assigns each record a Protected Identification Key (PIK). It is not possible to reliably assign a PIK to every record, either due to insufficient identifying information or because the information does not uniquely match any of the administrative records used in the person validation process. Non-random ability to assign a PIK can potentially inject bias into statistics using linked data. This paper studies the nature of this bias using the 2010 American Community Survey (ACS). The ACS is well-suited for this analysis, as it contains a rich set of person characteristics that can describe the bias. We estimate probit models for whether a record is assigned a PIK. Preliminary results suggest that young children, minorities, residents of group quarters, immigrants, recent movers, low-income individuals, and non-employed individuals are less likely to receive a PIK. We also estimate multinomial logit models that further divide the non-PIKed records into those with insufficient information to be sent through the PIK assignment process vs. sufficient but non-matching information. We find that young children, immigrants, and recent movers are particularly likely to have non-matching information.

Social Security Income Measurement in Two Surveys

Howard Iams (Social Security Administration) and Patrick Purcell (Social Security Administration)

Social Security is a major pillar of retirement income in the United States which influences economic well-being and poverty. Because of its importance to economic well-being, it is important that household surveys accurately measure Social Security income. Using Social Security Administration (SSA) records on payments, we examine the recording of Social Security income for 2009 in the Census Bureau's Current Population Survey (CPS) and Survey of Income and Program Participation (SIPP). A complication for measurement in surveys is the usual deduction of Medicare premiums from Social Security income before SSA pays the actual Social Security benefit to beneficiaries. The money deposited in the bank reflects the net amount of Social Security income after these Medicare deductions from the Social Security benefits. The Census Bureau states that the CPS and SIPP record the full Social Security income before deductions for Medicare Premiums. We find that the CPS recording of Social Security income closely approximates the payment records of the SSA for the CPS respondents. We find that the SIPP recording of Social Security income approximates the SSA record of the net payment rather than the full Social Security income payment. This SIPP underestimate approximates the impact of Medicare premiums deducted from Social Security income.

Changes in EITC Eligibility and Participation, 2005-2009

Maggie Jones, (U.S. Census Bureau)

The rate of participation in the Earned Income Tax Credit (EITC) has been widely studied, but changes over time in eligibility for the credit have received less attention. One question of importance to policy-makers is whether (or by how much) eligibility might increase during economic downturns. The EITC is fundamentally tied to work. During periods of high unemployment, eligibility may decrease due to a lower number of workers (especially low-skilled workers) filing for a given tax year. On the other hand, family structure and underemployment may lead to increases in eligibility. For example, earners may become eligible when a two-earner family loses one job or when an earner works part of the year or fewer hours.

Using IRS tax data linked with the Current Population Survey Annual Social and Economic Supplement (CPS ASEC), I examine changes in EITC eligibility and take-up between tax years 2005 and 2009, during which time the Great Recession began and ended. Employing fixed-effects models, I assess patterns of eligibility among demographic groups defined based on characteristics that also predict outcomes in the labor market. Results indicate that, in a period when overall EITC eligibility rates increased, the state unemployment rate had a significant positive effect on eligibility and a significant negative effect on take-up. Meanwhile, groups who especially felt the impact of the recession experienced increases in eligibility rates, although these same groups did not participate in the program at higher rates.

Profile of Supplemental Nutrition Assistance Program (SNAP) Usage at the State and County Levels: Evidence from Texas and New York SNAP Administrative Records and the American Community Survey

Benjamin Harris (U.S. Census Bureau), Erik Scherpf (Economic Research Service), Constance Newman (Economic Research Service), and Graton Gathright (U.S. Census Bureau)

Means tested programs can be administered more effectively by identifying the characteristics and locations of individuals who, despite being eligible, do not participate. Linking administrative records to individual survey responses permits novel research on the relationship between program eligibility and program participation. Household surveys provide most of the information needed to estimate program eligibility, but program participation in surveys is known to be underreported. Administrative records, on the other hand, provide accurate information on program participation, but contain little information on the characteristics of participants, and no information on non-participants. Coupling these two data sources is therefore key to providing an accurate and rich profile of likely-eligible non-participants. This paper uses administrative records from Texas and New York States' Supplemental Nutrition Assistance Program (SNAP) linked with survey responses from the American Community Survey (ACS) to generate detailed profiles of the demographic and economic characteristics of likely-eligible non-participants. These profiles are provided at both the state- and county-level for 2009 (Texas) and each year, 2008–2010 (New York). We find substantial differences across states and over time in participation rates by race, Hispanic origin, linguistic isolation, age, and employment status. SNAP administrators can use these findings to target program outreach to underserved populations more effectively.

CONCURRENT SESSION E-3

DATA DISCLOSURE ISSUES

A Disclosure Avoidance Research Agenda

Paul Massell, (U.S. Census Bureau)

One of the striking facts about disclosure avoidance (also known by names such as 'confidentiality protection' or 'statistical disclosure control (SDC)') is the diversity of the protection methods. They vary greatly with the type of data product being protected, e.g., (frequency) count tables, magnitude data tables, microdata from a survey or a census or a statistical model. Some of the older methods for protecting count tables, e.g., collapsing of categories by which rows and columns are defined, can be learned by someone with just a basic understanding of statistical tables. Data swapping requires a good understanding of (demographic) microdata. Cell suppression requires knowledge of optimization techniques (e.g., linear programming). Some new sophisticated methods such as synthetic data require knowledge of ideas from Bayesian statistics and experience with statistical modeling. Even with some of the simpler methods, knowing exactly in what situations it is appropriate to use the method and how to fine-tune its use, requires experience. Similarly experience is needed to determine which method is best to apply in situations in which a number of methods are possible choices. Lately hybrid methods have become popular. Trying to create a coherent overview of these methods is a useful project for an agency that often needs to extend them to different situations. A course in disclosure avoidance would be a nice side benefit of such an effort. Such a course could be taught to new researchers in this field as well as those subject matter researchers who are often involved in disclosure avoidance issues.

Evaluating Noise Infusion for Disclosure Protection for Two Time Periods

Jared Martin (U.S. Census Bureau)

The U.S. Census Bureau conducts the Economic Census of Island Areas quinquennially. For the 2007 Economic Census of Island Areas, the Census Bureau provided disclosure protection using noise infusion to comply with the confidentiality requirements in Title 13 of U.S. Code. Noise infusion consists of perturbing values for each contributing establishment prior to table creation by applying a random noise multiplier to the magnitude data (i.e., characteristics such as receipts, payroll, and number of employees). The use of noise factors introduces uncertainty into the published cell values that protects the

confidentiality of each respondent's data. When planning for disclosure protection for the 2012 Economic Census of Island Areas, Census Bureau staff considered the impact of noise infusion on estimates of change in published cell values from one census cycle to the next. In an attempt to minimize distortion in cycle-to-cycle changes while providing sufficient uncertainty to protect the confidentiality of the respondents, they considered several options, which include different levels of control in the magnitude and direction of the noise factors and the establishments to which they are applied. This paper presents the results of a simulation study used to compare and evaluate the options considered, for estimates of varying size and composition. The paper illustrates each method and shows the tradeoff between bias and low variability in an estimate due to noise.

Is an On-Line Microdata Tool That Uses American Community Survey Transportation Data Feasible?

Tom Krenzke (Westat), Penelope Weinberger (American Association of State Highway and Transportation Officials), Jianzhu Li (Westat), and Elaine Murakami (Federal Highway Administration)

This feasibility study investigates the value added of testing and developing an on-line microdata tool using the American Community Survey (ACS) for disseminating Census Transportation Planning Products (CTPP). The current approach for disseminating the CTPP data is through pre-specified tabulations. The tables are generated using key transportation variables, which includes means of transportation, commute time, and time leaving home for work. Several aspects of feasibility are explored. Issues relating to system implementation and hosting are addressed. The progress and timing of the developing systems are provided. Data user needs are gathered through a small survey of transportation planners to determine the appropriate geographic levels and topics in the potential system. Since confidentiality screening is a must in a real-time system that is processed from ACS restricted-use microdata, its trade-offs with user needs and data utility are discussed.

The Census Bureau's New Cell Suppression System

Philip Steel (U.S. Census Bureau)

Cell suppression is the one of oldest methodology for protecting establishment or company based data. The Census Bureau is in the process of replacing its network flow based system with a new, Linear Programming based system. We describe that system and compare its performance to the old system. We also look at the question of whether the LP system is as responsive to weighting as the old network flow based system.

Managing Confidentiality and Provenance Across Mixed Private and Publicly-Accessed Data and Metadata

Carl Lagoze (University of Michigan), William Block (Cornell Institute of Social and Economic Research), Jeremy Williams (Cornell Institute of Social and Economic Research), John Abowd (Cornell University), and Lars Villhuber (Cornell University)

Social science researchers are increasingly interested in making use of confidential micro-data that contains linkages to the identities of people, corporations, etc. The value of this linking lies in the potential to join these identifiable entities with external data such as genome data, geospatial information, and the like. Leveraging these linkages is an essential aspect of "big data" scholarship. However, the utility of these confidential data for scholarship is compromised by the complex nature of their management and curation. This makes it difficult to fulfill U.S. federal data management mandates and interferes with basic scholarly practices such as validation and reuse of existing results.

We describe in this paper our work on the CED^{2}AR prototype, a first step in providing researchers with a tool that spans the confidential/publicly-accessible divide, making it possible for researchers to identify, search, access, and cite those data. The particular points of interest in our work are the cloaking of metadata fields and the expression of provenance chains. For the former, we make use of existing fields in the DDI (Data Description Initiative) specification and suggest some minor changes to the specification.

For the latter problem, we investigate the integration of DDI with recent work by the W3C PROV working group that has developed a generalizable and extensible model for expressing data provenance.

CONCURRENT SESSION E-4

MEASURING, MONITORING, AND ASSESSING THE QUALITY OF SURVEY DATA

The Quality Assurance Reviews at Statistics Canada

Laurie Reedman (Statistics Canada) and Claude Julien (Statistics Canada)

The Quality Assurance Reviews are a low cost yet high benefit quality initiative. The Reviews are an innovative quality evaluation tool, focusing on quality assurance at the implementation and execution stages of a statistical program. The mandate of the Quality Assurance Reviews is to demonstrate how quality is being managed. The objectives are to identify risks to the statistical program delivering its regular product, to highlight good practices that can be shared with other programs, and to make recommendations for improvement. The recommendations address quality assurance practices or actions to mitigate risks, and have measureable impacts.

Regular survey programs, cost recovery programs and programs based on administrative data are in scope for the Quality Assurance Reviews. Also, horizontal topics such as coding methods and practices and certification standards have been the subject of Quality Assurance Reviews.

Reviewers are Statistics Canada employees at the middle management level. They are assigned to review programs outside of their current area of responsibility, so as to ensure independence and an unbiased assessment. They are guided and facilitated by a steering committee made up of higher level managers, who ensure that relevant and timely information is shared with the reviewers. The findings of the reviews are shared with Statistics Canada's highest level management committee.

Program managers are afforded a variety of benefits from having their programs reviewed. If a program is heading into a redesign phase, the recommendations from the review provide valuable direction. If on the other hand the review finds few unmitigated risks, the good practices already in place are highlighted and shared with other programs. The Quality Assurance Reviews complement other program evaluation tools in use at Statistics Canada.

Distribution of Quality Control to the Point of Data Collection in the Field: Impacts on Cost, User Experience, and Security

Andrea Johnson (U.S. Census Bureau), Sandy Dyer (U.S. Census Bureau), Michael Ratcliffe (U.S. Census Bureau), and Jonathan Krentel (Gunnison Consulting Group, Inc).

The effort to collect quality data from respondents in the field constitutes one of the single greatest cost drivers associated with any large survey. The quality control practices associated with field data collection are a significant portion of this cost. These quality control practices fall into two categories: (1) Practices to monitor and control the work of field representatives during field operations and (2) Quality-related processing of data after field operations are complete. In both of these areas, the distribution of business rules and quality control mechanisms to the point of data collection in the field can result in (a) substantial reductions in cost and (b) improved quality of field collection data.

In this presentation, we take the Census Bureau's 2009 Address Canvassing operation as a representative example of these dynamics. We focus on quality-related challenges pertaining to the detection of human error during the map-spotting process in the field.
Examples of these challenges include: "long strand" detection, mapspot "heaping", and ambiguous collection block determination. Clerical operations at the National Processing Center (NPC) were undertaken to resolve questions that arose in the data because of these types of challenges. Our argument is that in the future, these types of issues can be resolved at the point of data collection, thus eliminating the need (and associated cost) of such follow-up clerical operations.

TMO's new Listing and Mapping Application (LiMA) is presented as an example of how this works. Using LiMA as an example, we explore (1) the estimated cost benefits of this new paradigm, (2) the impacts it has on user experience in the field, and (3) associated security-related considerations. Finally, we look at current trends in mobile computing generally to see the extent to which they affirm or disconfirm the technical assumptions that support this approach.

An Approach to Measurement Error Assessment for a Household Expenditure Survey: A Review and Application

Brandon Kopp (Bureau of Labor Statistics), Roger Tourangeau (Westat), Lucilla Tan (Bureau of Labor Statistics), and Scott Fricker (Bureau of Labor Statistics)

In any survey that relies heavily on a respondent's memory, there is bound to be some degree of measurement error. The Consumer Expenditure Survey (CE) asks respondents detailed questions about the occurrence of expenditures for a wide range of household goods and services, which demands considerable cognitive effort from respondents. The respondent's comprehension of the meaning of the question, ability to retrieve the relevant information, and formulation of an answer are among the key steps in determining the accuracy of the response. Given the opportunities for error at each of these steps, measurement error is undoubtedly present in the CE.

An important first step in minimizing measurement error is to understand its magnitude, direction, and causes. The CE is currently conducting research on redesign options with the primary goal of improving data quality through a verifiable reduction in measurement error. To that end, the authors of this paper reviewed 37 research reports, conference proceedings, and journal articles spanning from 1964 to 2012 and identified four broad categories of methodologies that could be used to track the magnitude and direction of measurement error over time and help shed light on it's causes; (1) comparisons with administrative data on consumer expenditures, (2) comparisons with other household surveys with overlapping expenditure categories (e.g., American Community Survey, Medical Expenditure Panel Survey), (3) analysis of data within the CE (e.g., diary and interview comparisons, latent class analysis, and the use of paradata), and (4) respondent provided financial record validation studies. Each method has characteristic strengths and weaknesses that are discussed along with a strategy for using multiple methods as a means of tracking measurement error over time. Lessons learned and implications for the study of measurement error in surveys more broadly are also discussed.

A Model for Matching Error for Triple System Estimation

Richard Griffin, (U.S. Census Bureau)

Capture probability heterogeneity is known to produce bias in the dual system estimates (DSE) which have been used to estimate census coverage in U.S. Censuses since 1980. Triple system estimation using an administrative records list as a third source along with the census and post enumeration survey (PES) has the potential to produce estimates with less bias. Each of the three frames is incomplete and matching between the frames is required. Matching errors add to the bias of potential population estimates. Biemer (Survey Methodology, 1988) presents a model for the matching error associated with DSE. This paper presents a matching error model for triple system estimation. It is assumed that all N individuals in the population are exposed to possible inclusion in all three sources. In practice, sampling is necessary for the post enumeration survey. In addition, erroneous inclusions including within list duplicates have been removed from all lists. The multinomial model assumes that the Census list, the post enumeration survey list and the administrative list are created as a result of N mutually independent trials from one person to the next (all persons are captured independently of all other persons for a given list). An expression for the mean squared error (MSE) is derived using Taylor Linearization approximation methods. A simulation study is used to compare the MSE of DSE with MSEs of triple system estimators associated with log linear models assuming no second order interaction, conditional independence and joint independence.

CONCURRENT SESSION F-1

EVALUATING WEIGHTING AND VARIANCE ESTMATION METHODS

Effects of Poststratification and Raking Adjustments on Precision of MEPS Estimates

Sadeq R Chowdhury (Agency for Healthcare Research and Quality)

The Medical Expenditure Panel Survey (MEPS) weighting procedures include several rounds of nonresponse and poststratification/raking adjustments. Although poststratifications/rakings using external control totals are generally applied for reducing bias due to nonresponse and coverage, an added benefit of these adjustments is variance reduction. The survey variances of control totals used in adjustments become zero while the variances of other estimates that are correlated to these totals decrease depending on the extent of correlation. However, this reduction in variance due to poststratification/raking is not captured in computing the variances of MEPS estimates. The Taylor linearization method is generally used for estimating variances in MEPS which is unable to account for the effect of these adjustments. A shortcut replication approach is also used sometimes for computing variances of MEPS estimates. Based on a published BRR replication structure, the users can form BRR or Fay's BRR replicate weights from the final MEPS weight. Forming replicates from the final weight is a one-step approach which also does not capture the effect of poststratification/raking on variances. To assess the effects of these adjustments on the variances of MEPS estimates, a set of proper Fay's BRR replicate weights are computed starting with the base weight and independently applying all weighting adjustments to each of these replicates. Variances of selected MEPS estimates are then computed from these replicate weights and compared with those of the Taylor method. This talk presents the results of this comparison which shows that the variance estimates are generally lower under the BRR method implying that the overall weighting adjustments in MEPS generally reduces variance.

Model-Assisted Domain Estimation

Dan Liao (RTI International), and Phillip Kott (RTI International)

Suppose we had a probability sample and wished to compute a domain-level estimated mean or total. Using calibration weighting or probability-weighted linear prediction will produce the same results when there is a domain indicator among the calibration targets. But what if there isn't? Comparisons will be made between the validity (bias) and reliability (variance) of these two methods through a simulation study based on a national health survey. Although weighted prediction can produce a more efficient estimated mean (i.e. one with a smaller variance) than calibration weighting, the estimate may no longer be asymptotically design unbiased. Fortunately, we can test whether this bias is significantly different from zero. If it is not, and the model supporting this estimation strategy can be assumed to hold in the domain, then the variance of the domain estimate can be measured under the combination of the model and probability-sampling theory even when the sample design in not ignorable. This variance will be compared with the variance of the corresponding domain estimate derived using calibration weighting.

These rival methods are frequently available when there is a two-phase sample and the calibration targets for the final sample are computed from the first-phase sample. Assuming, as we initially do here, a single-phase framework allows us to investigate the issues in a more straightforward manner. Nevertheless, we will discuss the additional complications in standard-error estimation caused by the existence of two sampling phases.

Variance Estimation for Calibration to Estimated Totals

Siyu Qing (RTI International) and Michael Larsen (The George Washington University)

In survey weight calibration, the control totals that are used as targets of calibration are assumed known population values. Sometimes, however, the control totals are estimated from other sources, which increase the variation. Many researchers use naïve variance estimators as if the inflation of the variance associated with estimated control is insignificant and hence negligible. Simulation shows naïve estimators can seriously underestimate the true variance, producing substantial drops in confidence interval coverage. This paper presents a new Taylor linearization variance estimator for an estimator calibrated to

design-unbiased control totals. Comparisons are made to naïve and Jackknife estimators. The results of simulations indicate that this new estimator can mitigate the underestimation, thereby producing better confidence interval coverage. Applications to the NSF's Survey of Doctorate Recipients are presented.

Impact of Influential Observations on Enumeration and Variance Estimation in the National Crime Victimization Survey

Michael Planty (Bureau of Justice Statistics) and Lynn Langton (Bureau of Justice Statistics)

The National Crime Victimization Survey (NCVS) collects information from respondents about the number and type of victimizations they experience each year. A small percentage of respondents experience a high number of victimizations within the reference period, referred to as series incidents. These chronic victims are substantively important groups as they represent victims of violence from intimate partners and incidents that occur routinely in the workplace and at school. These victims, however, introduce the possibility of extreme weights in the NCVS data. These extreme weight outliers can be influential and significantly increase both point estimates and variance of estimates. One way to control for these impacts is to trim and smooth extreme weights. Trimming weights has the potential to increase bias, but the increase in bias can be offset by the decrease in variance. This paper describes various approaches to incorporating series incidents into the enumeration scheme through trimming while balancing concerns about point estimates and variance.

CONCURRENT SESSION F-2

INTERNATIONAL ECONOMIC STATISTICS

Estimating Foreign-Born Emigration from the United States Using Data from the American Community Survey

Mark Leach (U.S. Census Bureau)

The Population Estimates Program at the U.S. Census Bureau annually produces estimates of international migration. This paper focuses on a subcomponent of international migration, foreign-born emigration (FBE). The research explores potential impacts of using data from the American Community Survey (ACS) in place of Census 2000. In absence of comprehensive data on people who leave the United States to reside in another country, the Census Bureau uses a residual method to estimate FBE. In the most recent series of estimates, for example, the Census Bureau used data/information from Census 2000 to enumerate the foreign-born population and calculated an expected survived population in 2011 after accounting for mortality. The difference, or residual, between the expected population and the foreign-born population enumerated in the 2011 ACS, excluding immigrants who arrived after 2000, is assumed to be due to migrants leaving the country.

Research shows that recent arrivals, who are relatively less settled, are most likely to return to their home country. Continuing to rely on Census 2000 into the 2010s implies, however, that the Census Bureau's emigration estimates will be based on immigrants that have been in the country for at least 10 years. This may result in estimates that underestimate emigration and overestimate net international migration. The present research assesses the feasibility of using data from the ACS to calculate an expected survived population. Annual data from the ACS will ensure that the behavior of immigrants who arrived more recently is incorporated into the estimates. Specifically, the study assesses the implications of relatively small sample sizes and sampling error in the ACS for surviving a foreign-born population by age and sex. It also compares estimates based on residuals with different beginning and end years and length of survival time periods to understand the robustness of the estimates.

The Use of Theory and Model Averaging for Population Prediction: An Application to the U.S. Overseas Population

Sidney Carl Turner (Fors Marsh Group, LLC), Joseph Luchman (Fors Marsh Group, LLC), Andrew Therriault (Lightbox Analytics), Brian Griepentrog (Fors Marsh Group, LLC), Kinsey Gimbel (Fors Marsh Group, LLC), Fritz Scheuren (University of Chicago), Ali Mushtaq (Consultant), and Paul Drugan (Federal Voting Assistance Program)

Increasing international mobility poses a challenge to national governments, with respect to their ability to provide goods and services to their overseas citizens. Arguably, the most important service democratic governments provide to their overseas citizens is ensuring their ability to vote.

Addressing these policy challenges requires that national governments have information concerning the size and location of their overseas citizens. However, this is difficult due to the lack of comprehensive international surveys of in and out migrants. Methodological inconsistencies between national surveys can also lead to measurement errors and biases.

Although a number of global two-way migration databases are available (including those maintained by organizations such as the Organisation for Economic Co-operation and Development (OECD), United Nations, and World Bank), such available estimates are created using pooled samples of origin countries and consequently do not account for unobserved heterogeneity in origin countries with respect to outmigration tendencies. In addition, many country-level population estimates are based on rudimentary imputation methodologies, without thorough validation checks.

Our paper describes an effort to estimate the size and location of the overseas U.S. population using a combination of U.S. government immigration/emigration data, plus foreign government data on overseas U.S. citizens, and theoretically-derived predictors of international migration. The approach accounts for both methodological and substantive variation in foreign government estimates. Uncertainty in the model selections—both in terms of estimation method and variable inclusion—are addressed through the use of a modified *Ensemble Bayesian Model Averaging* methodology. With this technique, sets of predictions are derived from all subsets of variables and estimation methods and are averaged to arrive at a final prediction. Prediction weights are derived employing leave-one-out (i.e., jackknife) cross-validation. Parameters and fitness statistics are calculated using inverse probability weighting to account for potential bias.

Effects of the Eurozone Economy on Economic Indicators

Tracey Kyckelhahn (Bureau of Justice Statistics)

Global economic problems have impacted national statistics in a variety of ways. Many economic models rely on assumptions that may not apply during recessions and black market activities may increase, thus reducing the validity of available data. However, the continuing crisis in the Eurozone economies of the European Union in particular has spurred legislative action intended to make economic activity more transparent. Proposed and enacted laws have focused on limiting cash transactions and reducing tax evasion through tax havens. This paper will discuss how these changes coupled with technological advances that allow for the collection and manipulation of large amounts of data may alter economic indicators.

In addition to legislative activities that affect individuals and businesses, this paper will examine the greater enforcement of transnational rules regarding the production of official economic statistics due to the Eurozone crisis. Finally, the last section of the paper will discuss how some other actions taken by the EU and national governments during the economic downturn may make valid data more difficult to obtain.

Offshore Outsourcing, Portfolio of Intangible Capital, and Performance: Theory and Evidence in the IT and Pharmaceutical Industries

Wendy Li (Bureau of Economic Analysis)

The degree of offshore outsourcing has increased rapidly in past few decades and economists have argued that the role of intangibles in the U.S. economy has grown sharply during the same period time. This paper develops a forward-looking intangible investment model to estimate the depreciation rates of R&D assets and organizational capital for the IT and pharmaceutical industries. The data is from the Computstat dataset and covers the period of 1987 to 2010. The estimates are used to construct the stocks of these two types of intangible capitals for both industries. In general, R&D assets depreciate faster than organizational capital in both industries, which is consistent with the recent finding by Bloom and Reneen (2010) that the evolution of management practices is slow and new improvements take time to diffuse across firms. Since 2000, both industries have decreasing ratios of R&D capital to total intangible capital, a pattern implying that organizational capital have played an increasingly important role for firms to sustain competitiveness in R&D intensive industries. However, how each type of intangible capital relatively contributes to each industry's performance depends on the degree of offshore offshoring, the pace of technological progress, and the investment scale of intangibles.

Benchmarking Monthly Indicator Series with Quarterly Information Derived From Administrative Data

Pieter Vlag, Statistics Netherlands

Several countries in Europe use the Value Added Tax (VAT) for short-term statistics (STS). Two challenges exist when using VAT for short-term turnover estimates; 1) timeliness, part of the VAT is available too late, 2) periodicity, part of the VAT-data cover a different period than the required statistical output. Therefore VAT is complete when quarterly and monthly turnover estimates have to be produced. An European project has focused on this challenge.

Two situations can be distinguished when producing STS in a VAT based system with the largest enterprises being surveyed. Situation I is that VAT provides good coverage when the estimates have to be made. Situation II is that no or only few VAT is available.

Situation I generally exists for quarterly estimates. The main conclusion of the project is that good quality level and growth rate estimates can be produced with VAT under these circumstances. This is called the structural series.

Provided that structural series can be calculated for the quarters, the challenge is to find estimation methods for (early) month when no or few VAT is available. The project suggests that monthly estimates should be considered as an 'indicator', benchmarked with the structural series when VAT becomes available. A small survey under larger enterprises is sufficient, if benchmarking demonstrates over a longer period that the short-term trend is well measured by these enterprises. In this case, monthly estimates can be provided by either a) the growth rates of the surveyed enterprises only or b) the growth rates of the larger enterprises plus a correction for the difference in the long-term trend between surveyed larger and non-surveyed smaller enterprises. Results of a case study on the retail trade and commercial services enterprises in the Netherlands will be discussed during the presentation.

CONCURRENT SESSION F-3

ASSESSING THE EFFECTIVENESS OF STRATEGIES TO IMPROVE SURVEY RESPONSE RATES

Testing Contact and Response Strategies to Improve Response in the 2012 Economic Census

Erica Marquette (U.S. Census Bureau) and Michael Kornbau (U.S. Census Bureau)

While planning the 2012 Economic Census, the Census Bureau tested six various contact and response strategies to improve the overall response rate to the Economic Census while attempting to achieve earlier response and obtaining more internet responses. The following six tests were conducted during different phases of mailout to achieve this goal:

1. Sending a follow-up mailing containing only a letter to a sample of nonrespondent single-unit companies who initially received a questionnaire to determine industry classification.
2. Mailing a questionnaire to small multi-unit companies prior to the Economic Census to obtain better contact information.
3. Mailing an advanced letter prior to the initial mailout to single-unit companies in certain low-response industries.
4. Calling nonrespondent single-unit companies with an automated message reminding them to respond to their census questionnaire.
5. Sending only a letter instead of a questionnaire to single-unit companies during the first follow-up mailing.
6. Sending certified mailings to single-unit companies in certain low-response industries during the second and third follow-up mailings.

This paper will explain the objectives used for each test, how the tests were conducted, how the results were evaluated, and provide preliminary results.

Use of Smart Phones/Text Messaging to Increase Response Rates

Piper DuBray (ICF International), Naomi Freedner (ICF International), Kisha Bailly (ICF International), and Kristie Healey (ICF International)

Survey response rates have greatly declined in the past decade, causing researchers to seek new ways to increase participation. The Connecticut Department of Health (CT DPH) and ICF International conducted two pilot studies in 2012 using text messages to 1) increase response rates to the Behavioral Risk Factor Surveillance System (BRFSS) cell phone survey, and 2) increase participation in the BRFSS Non-Response Web Follow-up.

To evaluate the impact of an advance text message on survey response, the CT BRFSS cell phone sample was divided into 3 groups: Group 1 was sent a text asking the respondent to complete the telephone survey when called, also offering a $10 incentive. Group 2 received the text invitation with no incentive offer, and group 3, the control group, did not receive a text message.

The second pilot consisted of sending BRFSS telephone non-responders a text message invitation to complete the survey via Web. Non-responders were divided into 2 groups: Group 1 received 2 text messages inviting them to participate in the Web survey and offered a $10 incentive for participating. Group 2 was sent the text invitations, without an incentive.

Early results show that text invitations to the Web survey do not have a significant effect on response rates. Initial results of advance texts to cell phones show a 2% increase in CASRO over the control group, while advance text with incentives show a 3% increase in CASRO over the control group. We will conduct further analyses after all data has been collected, and determine whether this increase in response rate is significant.

Based on preliminary results, text messages as a tool to increase response had mixed results. Advance text messages increased participation in a telephone survey, but text messages to BRFSS non-responders were ineffective in increasing Web survey participation.

Do Names Matter? Experiments Comparing Different Branding and Levels of Personally Identifiable Information in a Mail Questionnaire

Sarah Hastedt Carroll (National Center for Education Statistics) and Andrew Zukerberg (National Center for Education Statistics)

Falling response rates on telephone administered surveys have led researchers to consider alternative modes for data collection including the use of mail surveys. As a result, renewed attention is being paid to the design of mail surveys. Previous research has shown that survey sponsorship and content can impact response rates to a survey. In this paper we will present the results of a study that experimentally varied the branding on survey materials between two different Federal agencies and the level of Personally Identifiable Information (PII) collected. These experiments were conducted as part of the 2012 National Household Education Survey (NHES). In the first experiment respondents were randomly assigned to receive all surveys materials (letters, envelopes, questionnaires) with the Census Bureau logo, letterhead and official signatures or the same materials with Department of Education branding. The NHES is a two phase survey conducted almost entirely by mail. In the first phase households are screened to determine if any children under the age of 20 are living if the household. If they are, the respondent is asked to provide basic information including the child's name, age, gender and grade in school. A child is sampled in each household from the returned questionnaires and a detailed topical questionnaire is sent to the household to obtain information about the selected child's educational experience. During Pilot testing of the NHES there was inconclusive evidence that asking child's name could lower response at the screener phase, but increase response at the second phase since the topical questionnaire could be personalized with the child's name. We randomly assigned cases to receive a screener that either collected the child's name or did not. This allowed us to see the interaction between branding and collection of child's name. This paper will report on the results of these experiments.

Time Use, Response Rates, and Data Quality by Time of Day

Rose Woods (Bureau of Labor Statistics) and Laura Wronski (Bureau of Labor Statistics)

The American Time Use Survey (ATUS) collects a detailed 24-hour time diary of all the activities respondents participated in the previous day. Interviews are conducted via CATI throughout the day, with most interviews being conducted between 9 a.m. to 9 p.m. (respondent time). Respondents may call in or set up an appointment to complete the survey outside of these hours. This may result in some interviews beginning as early as 6 a.m. and some ending as late as midnight. Given the level of detail of time-use diaries, we are interested in whether data quality suffers later in the day.

We examine survey response, data quality measures, and characteristics of people responding to the ATUS by the time of day when respondents were interviewed. These characteristics include respondents' demographic composition as well as their time use.

CONCURRENT SESSION F-4

ADAPTIVE APPROACHES TO SAMPLING AND SUBSAMPLING

Adaptive Design Features for Using Address-Based Sampling in a National CATI Survey of Households with Children

Charles DiSogra (Abt SRBI), David Finkelhor (University of New Hampshire), Heather Hammer (Abt SRBI), Stas Kolenikov (Abt SRBI), and Heather Turner (University of New Hampshire)

The primary purpose of Wave III of The National Survey of Children's Exposure to Violence (NatSCEV III) is to document changes in the incidence and prevalence of children's exposure to a broad array of

violence, crime and abuse experiences. NatSCEV III uses an address-based sample (ABS) design to construct a nationally representative sample of households with children and adolescents age 0-17. While 34% of US households have children under age 18, CATI surveys of these households generally require costly telephone screening. The availability of ancillary demographic information that can be matched by a sample vendor to some proportion of the addresses in an ABS sample suggests the feasibility of a more efficient approach that uses an optimal allocation among matched and unmatched strata based on the expected proportion of households with children. Although presence of children is one of several indicators that can be used to stratify matched addresses in an ABS design, the accuracy of the ancillary data is imperfect and variable across indicators. NatSCEV III incorporates an adaptive design where sample vendor-provided ancillary demographic data for the matched addresses and demographic data collected from the first 1,000 completed interviews will be used to assess the accuracy of the vendor-appended ancillary demographic data and to build a logistic regression model that predicts the prevalence of households with children in each of the matched and unmatched strata. These results will then be used to develop an optimal stratified allocation for the main survey effort. Findings and conclusions about the viability of this approach for NatSCEV III and other national surveys of households with children will be presented.

Adaptive Adjustment of the Multicriteria Optimal Allocation of a Hard to Reach Population Survey

Benjamin Phillips Abt SRBI) and Stas Kolenikov (Abt SRBI)

The effectiveness of a survey of a rare population hinges on the information available about the prevalence of this rare population within the general population and areas where this rare population may be concentrated. Matters are further complicated when overall population and subdomain estimates within the population of interest are required, as these estimation objectives are frequently in conflict. We describe our work on a survey of American Jews, in which the estimates of the total population, as well as its demographic, religious and social characteristics are all of interest; moreover, the survey objectives encompass both estimates of the population as a whole as well as of domains of interest (Orthodox and Russian Jews). To approach the sample design, data from previous sample surveys, geocoded establishment information (synagogues and religious education organizations), and sample vendor ethnic name prevalence estimates were all used in deriving county-level small area estimates of Jewish-by-religion incidence. These estimates were used to stratify the general population into several strata with incidence varying from under 1.5% to above 10%. Nonlinear programming running under readily available software (Excel solver) was used to obtain the optimal allocation that included multiple accuracy criteria as well as the required sample sizes. The availability of the data from the field allowed reviewing the accuracy of the small area estimates in retrospect. More importantly, as the field period progressed, the observed incidence across strata and landline/cell phone frames was used to update the sample allocation in the adaptive design pursuing the most efficient use of the interviewer time and samples acquired from the provider. We present the trajectories of the observed incidences throughout the field period, discuss our adaptive design decisions, and quantify the economic impact of the adaptive design on the resulting total survey cost.

Optimizing Unit Nonresponse Adjustment Procedures After Subsampling Nonrespondents in the Economic Census

Laura Bechtel (U.S. Census Bureau) and Katherine Jenny Thompson (U.S. Census Bureau)

With budgets throughout the federal government tightening, many agencies have begun to consider adaptive strategies for data collection. One such strategy is to select a probability sample of nonresponding units for follow-up, instead of attempting complete follow-up. Besides saving cost, this can improve the tabulation quality by replacing an essentially self-selected sample with a representative sample. In this framework, one can design adjustment procedures that that mitigate or even eliminate unit nonresponse bias.

The Economic Directorate of the U.S. Census Bureau is investigating nonresponse subsampling strategies for usage in the 2017 Economic Census. Although no unified methodology has been fully

established, the design will include a systematic sample of nonrespondents, sorted by a unit measure of size. Given this subsampling design, we examine how best to improve the quality of the adjusted tabulations via a simulation study using selected industries from seven of the trade areas covered by the Economic Census. We consider two separate adjustment strategies – composite imputation that creates a complete record (the currently used approach) and adjustment cell weighting – examining the statistical properties of each approach on three basic data items collected by all sectors of the Economic Census.

Our simulation study has several layers, as there are many different factors that need to be examined such as varying subsampling rates (of the systematic sample), varying unit response rates, and varying conversion rates (the rate at which the sampled nonrespondents respond). In addition to these sample size considerations, we examine the effect of the response mechanism on adjustment procedures, as the optimal procedures under a nonignorable response mechanism could differ from those under an ignorable response mechanism. Quality effects are assessed by examining the bias, mean absolute error, and mean squared error of each item over repeated samples.

Determining Allocation Requirements for Subsampling Nonrespondents From the Annual Survey of Manufactures

Daniel Whitehead (U.S. Census Bureau), Stephen Kaputa (U.S. Census Bureau), and Katherine Jenny Thompson (U.S. Census Bureau)

As survey costs increase while response rates decrease, many agencies have begun to consider adaptive strategies for data collection. One such strategy is to select a probability sample of nonresponding units for follow-up, instead of attempting complete follow-up on all nonrespondents. A potential benefit of subsampling, besides cost reduction, is that by focusing on a smaller and more targeted number of units for nonresponse follow-up, it is possible to reduce the mean squared error, if the non-response bias reduction offsets the variance increase due to subsampling. This report presents the results of a simulation study conducted to develop allocation strategies for selecting a subsample of nonrespondents for follow-up in the Annual Survey of Manufactures, given a fixed total cost. The simulation study accounts for the two-phase sampling, testing the sensitivity of the assumed response levels and response mechanism, by examining the increase in variance (or coefficient of variation) caused by subsampling while monitoring follow-up cost.

CONCURRENT SESSION G-1

ISSUES ASSOCIATED WITH USING ADMINISTRATIVE RECORDS AND LINKAGE

Assessment of Quality, Cost and Risk Factors in Statistical Work With Administrative Record Data: A Review

John Eltinge (Bureau of Labor Statistics)

In recent years, large-scale statistical organizations have expressed increased interest in the use of administrative records to supplement standard sample survey data. Practical decisions on use of administrative data involve a complex set of factors that affect the balance of quality, cost and risk in the statistical production process. Many of these factors are qualitative in nature, e.g., ones that involve legal, regulatory, contractual or operational constraints on use of administrative data. Other factors are quantitative, e.g., measures of missing-data rates, misclassification rates, or cost components.

This paper reviews and synthesizes previous literature on assessment of the abovementioned quality, cost and risk factors. Seven issues receive special attention:

(1) Conditions under which primary interest may center on, respectively, qualitative or quantitative assessment of these factors.
(2) Inferential goals for the proposed use of administrative data; and linkage of these goals with the information needs of primary stakeholders.
(3) Administrative-record extensions of concepts and methods developed primarily within the context of sample survey methodology. These include “total survey error” models; broader assessments of survey data quality; integration of multiple data sources; and adjustments for incomplete data, reporting errors, aggregation effects and definitional effects.
(4) Measurement and modeling of fixed and variable components of cost; and amortization of some cost components across time and across multiple production systems.
(5) Distinctions between aggregate risks (associated with the cumulative effects of a large number of independent events) and systemic risks (associated with a small number of high-impact events).
(6) The degree of control that the statistical organization exercises over specified dimensions of quality, cost and risk.
(7) The extent to which issues (1) through (6) may be affected by the architecture of a statistical production system.

Riddles Wrapped in Mysteries Inside Enigmas: Issues with Getting and Using Administrative Data for Impact Evaluations

Alisha Creel (ICF International) and Ronald Szoc (ICF International)

Evaluators turn to administrative data as an objective rather than perceptual source of information that can be used to help triangulate evaluation findings along with non-administrative sources. Administrative data are generally designed for purposes other than evaluation, however, may only partially meet evaluators’ needs, and bring their own challenges. The Office for Juvenile Justice and Delinquency Prevention funded nine programs in eight states to reduce drinking among underage Airmen at Air Force bases and their surrounding communities through its Enforcing Underage Drinking Laws (EUDL) discretionary grant program. Using an environmental approach, the program interventions included enforcement aimed at reducing the social availability of alcohol; compliance checks of local liquor establishments; impaired driving enforcement; local policy development; community-based awareness/media campaigns; and offering alternative activities that did not include drinking alcohol. In this paper, we report on our experiences using five administrative data sources to assess the impact of these programs. Grantees provided monthly data retrospectively for two years prior to program implementation and prospectively for up to two years after the beginning of program implementation. Some of the micro-level issues that we faced included data that were incomplete for some months, data stored in forms that prohibited the kinds of aggregation that we needed for the purposes of the evaluation, and transferring and summarizing data from repositories that were designed for other purposes. The macro-level issues included lack of

cooperation from data source gate-keepers, lack of access due to concerns with violating HIPAA rules, and the length of time required to collect the retrospective data. Additionally, reconciling data from different sources was occasionally problematic. Based on these experiences, we offer a number of lessons learned and solutions for dealing with these issues for future evaluations that use administrative data.

Administrative Data Initiatives at Statistics Canada

Julie Trépanier (Statistics Canada), Jean Pignal (Statistics Canada), and Don Royce (Statistics Canada)

Statistics Canada has long been using administrative data (i.e., any data collected externally by governmental or non-governmental organizations) in its statistical programs and is determined to increase the use of such data when it leads to a better outcome - that is a better balance between relevance, quality, costs and respondent burden. To achieve this objective, the organization created an Administrative Data Secretariat in the fall of 2012, with the mandate to develop and implement a corporate approach to increasing the use of administrative data.

The Secretariat has undertaken several initiatives as part of this mandate. They include a review of the legislative, policy and organizational frameworks for the statistical use of administrative data that exist in Canada and elsewhere, to identify approaches that might be adopted at Statistics Canada; the construction of a central inventory of administrative data sources currently received by Statistics Canada, to understand better what data the organization uses and how it could use it better; the development of an evaluation framework, to assess the quality of potential administrative data sources and their statistical usability; and the conduct of a Census Program research project, to study the feasibility of building a statistical database of all Canadians with their basic demographic information by using multiple administrative data sources.

This paper will provide an overview of these different initiatives, progress made so far and future plans.

CONCURRENT SESSION G-2

ASSESSING ALTERNATIVE WEIGHTING METHODOLOGIES

Performance of Weighted and Non-weighted Estimators in a Cell Phone Based Electoral Poll: An Academic Study of the 2012 Presidential Elections in Mexico

Olivia Carrillo-Gamboa (Tecnológico de Monterrey, Mexico), Rosa Isela Hernández-Zamora (Tecnológico de Monterrey, Mexico), and Jesús Cantú-Escalante (Tecnológico de Monterrey, Mexico).

Electoral polls and surveys applied to political purposes have being used widely in Mexico for the last two decades. Although most of the methodologies applied are based on multiple-steps cluster sampling procedures and face-to-face interviews, these procedures are becoming of difficult application due to safety related issues.

On the other hand, methodologies based on landline phone sampling frames are known to yield samples with demographic characteristics substantially different from those of the target population –and potentially biased results- due to low coverage of the sampling frame as well as to the presence of other non-sampling errors tied to the logistic of the methods applied.

An alternative telephone methodology was applied in a series of academic studies in the context of the past 2012 Mexico's presidential elections which were based on samples drawn from a cell-phone sampling frame. The cell-phones interviews yielded demographic characteristics such as gender, geographical distribution and urban/non-urban composition comparable to the target population. The only concern from the cell-phone sample representativeness could be raised by the age-group composition since it can be easily understood older citizens -60 years and older- might have lower cell-phone coverage and could become underrepresented by the sample and at the same time younger age groups would be overrepresented.

This paper focuses on the performance of non-weighted estimators and of estimators weighted by age group when compared against official electoral results of the 2012 federal election. Both estimators were found to have an excellent practical performance and, although the weighting procedure could be the best theoretically recommended from a statistical point of view, the non-weighted estimator was slightly closer to the official results. A plausible explanation for this finding is also provided in the context of the 2012 Mexico's presidential elections.

The Use of Signal Filtering for Hog Inventory Estimation

Stephen Busselberg (National Agricultural Statistics Service)

The National Agricultural Statistics Service (NASS) uses probability surveys of hog farm operations to estimate quarterly hog inventories in the United States at the state and national levels. NASS does not publish these survey results for two primary reasons: the survey results are biased and the survey results do not satisfy a set of industry specified constraints at the national level. These constraints are a function of inventory over time in addition to external nonproprietary hog industry data such as imports, exports, and quarterly slaughter counts. The official published hog inventory estimates are a consensus of an expert panel called the Agricultural Statistics Board (ASB). The ASB estimates of inventory are subjective, not reproducible, and provide no measure of statistical error. This paper proposes and illustrates estimation of U.S. hog inventories using signal filtering methodology. Allocation of the U.S. level inventory estimates to the state level is formulated using Restricted Least Squares theory. This methodology provides a consistent, reproducible, and statistically defensible solution for hog inventories which satisfies all specified constraints and provides measures of statistical error.

DepicitingTime-Dependent Changes in Environmental Data for Prospective Uses in Personalized Medicine

Turkan Gardenier (Pragmatica Corp)

Queries related to environmental exposure have arisen in recent applications of personalized medicine.. For example, one may wish to evaluate the levels of specific air pollutants in an area, or areas, where a patient lived over a period of time.

Data exist to answer the-environmental exposure query, yet tracing stability or change over time is not clear cut. The presentation will focus on an urban city and a rural are in Eastern U.S. where annual concentrations for carbon monoxide (CO) Nitrogen Dioxide (NO2), and Ozone (O3) were available for 20 years, 1990-2010. In analyzing the data, time-dependent effects were computed using various methods, and their relative effectiveness were compared as to when a change was signaled. Changes were depicted in a "step" or boundary-crossing framework during the time-period observed.. This is particularly relevant to personalized medicine, because the profile of changes in environmental pollution for a specific location may be associate with other data for individuals living there.

The methods used in the analysis included : (1) linear regression fit over the 20-year data; (2) three-category "bands" created by the 20-year-mean +/- one half standard deviation, 75th and 25th percentiles and 90th and 10th percentiles, and (3) visual inspection of the range in data in order to use "natural breaks," an option within Geographic Information Science (GIS). ARIMA, Auto-Regressive Integrated Moving Average fit to the data was also explored. Stability of annual observations were tested using the non-parametric sign test for successive observations.

Preliminary results show that the use of +/- one-half standard deviation in creating the monitoring bands was the most favorable method for identifying consistent shifts in pollutant levels over the 20 years, 1990-2010.

Prediction Performance of Single Index Principal Fitted Component Models

Jia-Ern Pai (National Highway Traffic Safety Administration) and Kofi Placid Adragni (University of Maryland, Baltimore County)

Scientists in many research fields are encountering regression problems where the number of predictors *p* is larger than the number of observations *n*. For example, in a rare traffic accident, the number of potential confounding variables that may cause such accident is much greater than the number of accident cases. Statistical analyses on data sets with *p* greater than *n* are often referred to as "small-*n*-large-*p*" problems. To make further statistical inference, such as predictions, a reduction of the dimensionality of the predictors is always desired.

When building regression models, selecting potential variables, and making predictions, the conditional distribution of $Y \mid \mathbf{X}$ is traditionally applied, where *Y* is the response variable and **X** is a *p*-vector of predictors, such that $\mathbf{X} = (X_1, ..., X_p)^T$, $X_i \in \mathbf{R}$, and $i = 1, ..., p$. A regression model for $Y \mid \mathbf{X}$ is called forward regression. Some forward regression methods, such as the partial least squares (PLS), LASSO regression, and principal component analysis are frequently used. When the predictors in **X** are fixed, the forward dimension reduction methods are naturally chosen for modeling $Y \mid \mathbf{X}$. However, when the predictors in **X** are random, modeling $\mathbf{X} \mid Y$ may be a viable approach to attempt a dimension reduction on **X**. A regression model for $\mathbf{X} \mid Y$ is called an inverse regression. An example of inverse regression is the sliced inverse regression.

When dealing with large *p* problems, a lofty goal is to reduce the dimension of the *p*-predictor vector **X** to a *d*-predictor vector $R(\mathbf{X})$, such that $d \le p$, and use $R(\mathbf{X})$ as a surrogate variable. Dimension reduction methods have been developed for that purpose. It has been showed that when $(Y, \mathbf{X})$ have a joint distribution, $Y \mid \mathbf{X}$ can be linked to $\mathbf{X} \mid Y$ through $R(\mathbf{X})$ that carries all of the regression information that **X** has about *Y*. In addition, the conditional distribution of $\mathbf{X} \mid Y$ provides more reductive information when encountering the large *p* problems. Inverse regression approaches have been proposed to dimension reduction in the regression context, which is called principal fitted components (PFC) models. They are likelihood-based approaches that model $\mathbf{X} \mid Y$.

PFC models are equipped to capture any type of associations between the predictors and the response variable through the use of a set of basis functions. In this research, we consider a special case of PFC models, where the predictors are linearly related to the response variable. We restrict our research scope to scenarios, where the dimension of the sufficient reduction is one. The obtained reduction $R(\mathbf{X})$ is plugged in the forward regression model as the following:

$$Y = \alpha_0 + \alpha_1 R(\mathbf{X}) + e .$$

The prediction performances of this model, referred to as principal fitted components regression (PFCR), is studied through the simulations. We compare the prediction performances of PFCR with other traditional forward regression methods, such as the PLS and LASSO regressions. We consider the following three scenarios:

(a) Large *n* case, where *n* is greater than *p*,
(b) Dense case, where *p* is larger than *n* and all the predictors are related to the response variable,
(c) Sparse case, where *p* is larger than *n*. However, only a few predictors are related to the response variable.

An Alternative Estimation Method for the Current Population Survey

Daniel Bonnéry (University of Maryland, College Park and U.S. Census Bureau), Partha Lahiri (University of Maryland, College Park), and Yang Cheng (U.S. Census Bureau)

The Current Population Survey (CPS), sponsored jointly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics (BLS), is the primary source of labor force statistics for the population of the United States. The CPS is a 4-8-4 rotating panel design, and a sample consists of eight rotation groups of housing units. A composite estimator, known as AK estimator (see Gurney, (1965)), is used to estimate number of employed, unemployed, and unemployment rates.

Estimates of other quantities are weighted estimators, where the weights are computed following a method, called composite weights estimation (see Fuller (1990)), those weights being computed so that they match AK composite estimators for unemployment rates. The computation of composite weight for the CPS is the last of many steps in the weight computation process, that consists in different weight adjustments.

In this presentation, we critically examine practical and theoretical aspects, in the context of the CPS, of an alternative estimation method, originally proposed by Fuller and Rao (see Fuller and Rao (2001)), called Regression Composite Estimation and used by the Statistics Canada. For this purpose, we propose different models for employment evolutions and, via simulations, we study and discuss the properties of both estimators. We also compare these different estimators when applied to CPS data. It appears that a practical advantage of CPS lies in the relative simplicity of the one-step weight computation procedure.

CONCURRENT SESSION G-3

HITTING THE TARGET IN HOSPITAL PROFILING: THE AHRQ QUALITY INDICATORS

Variation in Quality by Hospital Characteristics: True or False?

Eric Schone (Mathematica Policy Research), Alex Bohl (Mathematica Policy Research), David Jones (Mathematica Policy Research), Dmitriy Poznyak (Mathematica Policy Research), Jessica Ross (Mathematica Policy Research), Samantha Stalley (Mathematica Policy Research), Frank Yoon (Mathematica Policy Research), and Joe Zickafoose (Mathematica Policy Research)

Recent critiques of the AHRQ QIs in the academic literature and popular press have called into question the validity of the indicators in making comparisons of hospital quality. A key component of the critiques is the assertion that certain hospital types are fundamentally different in their mission, patient populations, and service delivery This raises the question whether QI results using only patient-level risk adjustment is sufficient to support hospital comparisons. These assertions are typically supported by limited evidence that the group means of the QIs (average risk- and reliability-adjusted point estimates) vary by hospital characteristics. We present a systematic review of the variation in the QIs by hospital characteristics that are suspected to explain a meaningful proportion of variability in QI results. We add to the evidence base by considering the estimated variance of the risk- and reliability-adjusted rates when testing the differences between hospital group means. We also compare the distributions of hospital performance categories reported on CMS's Hospital Compare website (an alternate way of comparing hospitals) to determine if there is variation by hospital groups. Finally, we explore sources of the observed differences in a multivariate framework, by estimating the degree to which various hospital characteristics drive the variation in hospital results.

The Role of Hospital Characteristics in Setting Appropriate Yardsticks for Quality Measurement

Frank Yoon (Mathematica Policy Research), Alex Bohl (Mathematica Policy Research), David Jones (Mathematica Policy Research), Dmitriy Poznyak (Mathematica Policy Research), Jessica Ross (Mathematica Policy Research), Eric Schone (Mathematica Policy Research), Samantha Stalley (Mathematica Policy Research), Sheng Wang (Mathematica Policy Research), and Joe Zickafoose (Mathematica Policy Research)

Hierarchical models in the AHRQ Quality Indicators (QI) adjust for patient-level risk factors but not for potential variation in quality by hospital attributes, such as teaching status or bed size. This variation is typically modeled by hospital-level random effects, for example, by setting prior means in a Normal distribution that depend on hospital attributes. Alternatively, we evaluate the performance of hierarchical models that assume more flexible random effect distributions than the Normal to better account for underlying variation in quality. Flexible distributions for the random effects might avoid the need to explicitly specify models with hospital attributes while addressing concerns about their use and interpretation in hospital profiling. We will discuss this advantage and demonstrate the application of these enhanced models in a nationally representative inpatient claims database.

Hospital Peer Groups, Reliability, and Stabilization: Shrinking to the Right Mean

Alex Bohl (Mathematica Policy Research), David Jones (Mathematica Policy Research), Dmitriy Poznyak (Mathematica Policy Research), Jessica Ross (Mathematica Policy Research), Eric Schone (Mathematica Policy Research), Samantha Stalley (Mathematica Policy Research), Frank Yoon (Mathematica Policy Research), and Joe Zickafoose (Mathematica Policy Research)

The AHRQ Quality Indicators (QIs) are reliability-adjusted or “smoothed” to the national mean to deal with the unstable QI estimates due to hospitals with small numbers of denominator cases and rare outcomes. Differences in hospital scope, size, and other characteristics allude to the possibility that smoothing to alternate means determined by hospital attributes, or “peer groups,” may reduce bias when comparing hospitals on their estimated QIs. Current research suggests that incorporating peer-group targets into the risk-adjustment model through random effects is not feasible due to high-dimensional parameters and computational limits of MCMC estimation. Two alternative approaches are to: (1) smooth to a peer-group’s prior mean in the current framework (using an empirical estimate of reliability using the signal-to-noise ratio); or (2) add fixed effects for hospital characteristics to the risk-adjustment model. This study aims to compare the performance of these alternative peer-group smoothing methods and discuss their conceptual implications. We will judge model performance based on changes to model efficiency and reduction in comparison bias.

CONCURRENT SESSION G-4

NEW APPROACHES TO SURVEY DESIGN, IMPLEMENTATION, AND USE

The Consumer Expenditure Survey’s New Design and Implementation Plans

Laura Paszkiewicz (Bureau of Labor Statistics), Jennifer Edgar (Bureau of Labor Statistics), Brett McBride (Bureau of Labor Statistics), Dawn Nelson (U.S. Census Bureau) and Adam Safir (Bureau of Labor Statistics)

The U.S. Bureau of Labor Statistics (BLS) began the Gemini Project in 2009 with a goal of redesigning the Consumer Expenditure Surveys (CE) as a response to increasing evidence in measurement error, declining response rates, the emergence of new data collection technologies, and need for flexibility in addressing changes in the interviewing environment. The primary mission of the Gemini Project is to improve data quality through a verifiable reduction in measurement error, with a particular focus on under-reporting. Earlier stages of the Gemini project focused on gathering information to inform redesign decisions. This included conducting and reviewing research on survey methodologies and prioritizing user needs. Additionally, the BLS reached out to experts from the Committee on National Statistics (CNSTAT) in 2010 to devise an expert panel charged with recommending different CE design options given the

project goals. The CNSTAT panel presented three alternate designs in September 2012.

The Gemini Project will meet a major milestone in July 2013 by releasing a new design proposal including input from the CNSTAT panel recommendations, other outside experts, and research conducted. The proposed design will include a single sample design using multi-mode collection methods and new technologies. This presentation will discuss the details of the new proposal and outline the roadmap for testing, evaluation, development, and implementation of the redesign. This presentation will also highlight how the CNSTAT's recommendations were addressed in the proposed redesign. The presentation will be of interest to survey methodologists, data producers, and data users.

Race and Ethnicity Measurement: Effects of Response Format

Randall Thomas (GfK Custom Research, LLC), Frances Barlas (ICF International), and Wendy Gross (GfK Custom Research, LLC)

In modern, multi-ethnic societies, a person's race and ethnicity are often viewed as determinants of his/her thoughts and behavior, yet measurement of race-ethnicity has varied over time. Though a 'Mexican' response was provided as a separate racial option only in the 1930 census, it was not until 1970 that a separate question on Hispanic origin was included on the long form. Following a 1997 OMB directive for federal data collection, the 2000 and 2010 U.S. Census forms used two separate questions to measure race-ethnicity, one for Hispanic origin and the other for race, which had five substantive race categories: white; black or African American; American Indian and Alaska Native; Asian; and Native Hawaiian or Pacific Islander. A sixth category, 'Some other race,' was the third most endorsed category, with 97% who endorsed it being Hispanic. As a result, the Census Bureau is considering changing race-ethnicity measurement by combining the Hispanic question with the race question, making it a single multiple response question. As prior work has shown, multiple response formats tend to yield lower item endorsement than yes-no questions requiring an answer for each response option (Smyth, Dillman, Christian, and Stern, 2006; Thomas and Klein, 2006). In our experiment, respondents were randomly assigned to one of three conditions: the 2010 Census version (2 items), a new single item version (multiple response), or a modified single item version presenting all options in a yes-no grid. Respondents were also asked how well they could read a newspaper or book in English and Spanish to examine correspondence with Hispanic identification. Over 3,600 U.S. respondents from a probability panel (GfK Knowledge Networks) completed this web-based experiment. We found significant differences between formats, with the yes-no grid format having the highest endorsement for individual options (e.g. Hispanic) as well as the highest multi-racial identification.

Backtranslation vs. Committee Approach: An Experiment Comparing How They Perform in Questionnaire Translation

Alisú Schoua-Glusberg (Research Support Services) and Ana Villar (City University, United Kingdom)

Twenty years ago, the only commonly used method for assessing questionnaire translations was backtranslation, a translation performed by a single translator, followed by a second translator translating it back into English. The original and backtranslated English versions are compared, supposedly allowing the monolingual researcher to assess the quality of the translation. If both versions differ in meaning, discrepancies are investigated. This process is expected to uncover problems in the original translation.

Janet Harkness began looking for alternative translation assessment approaches over fifteen years ago. She clearly identified the problems in using backtranslation exclusively, as it evaluates the translated instrument without actually looking at it. Backtranslation, she agreed with others, uncovers some problems in the original translation, some problems in the second or backtranslation, yet fails to uncover some problems in both steps. Harkness' developed the TRAPD (translation-review-adjudication-pretesting-documentation) model as an alternative method.

This paper will present results of a study comparing a translation obtained by backtranslation with one done by Committee Approach, a team approach that includes the steps in TRAPD and has gained acceptance in the industry since the mid 1990s, yet no experiments are published showing how it compares with backtranslation. Thus, many in the profession still believe backtranslation is the industry

standard, despite an increasing number of federal agencies and university researchers adoption of team approaches.

For this experiment, a survey scale of 66 items will be translated into Polish via Committee Approach. The original translation done by the translator with the strongest credentials will be sent to a backtranslator. A comparison of the backtranslated and the committee version will be performed to determine 1) how each process best identifies translation problems and 2) how each fares in producing an translation that native speakers in a focus group find most idiomatic.

Getting the Most out of a Limited Sample Size Field Test: Experiences From the National Survey on Drug Use and Health

Jonaki Bose (Substance Abuse and Mental Health Services Administration), Dicy Painter (Substance Abuse and Mental Health Services Administration), Doug Currivan (RTI International), Larry Kroutil (RTI International), and Kevin Wang (RTI International)

This conference presentation and paper will discuss how the 2012 National Survey on Drug Use and Health (NSDUH) Redesign Field Test was used to make decisions about the 2013 redesign dress rehearsal and the 2015 NSDUH redesigned survey implementation. This 2000-respondent field test was designed to evaluate the performance of revised sections of the questionnaire, new contact materials, new equipment and other changes. Based on an extensive analysis of the different facets of data collection including field performance and response rates; field staff opinions of the new equipment; field interview debriefing items on the new protocols; changes in skip patterns, missingness and imputation rates, and other specify usage; changes in actual estimates and other factors, decisions were made for the dress rehearsal and actual redesigned survey implementation. This paper draws on that experience and discusses factors to consider in the design and interpretation of field test results, especially when the sample sizes are not large enough to detect potential breaks in comparability and trends.

CONCURRENT SESSION H-1

CROWDSOURCING METHODOLOGIES THROUGHOUT THE SURVEY LIFECYCLE

Crowdsourcing in the Cognitive Interviewing Process

Joe Murphy (RTI International), Michael Keating (RTI International), and Jennifer Edgar (Bureau of Labor Statistics)

Crowdsourcing holds promise as an alternative or supplement to traditional cognitive interviewing methods. Faced with the challenge of assessing the extent to which a question or concept is understood before finalizing a survey instrument, researchers typically recruit subjects for an in-person administration of draft items using think-aloud methods to capture feedback. This can provide some deep understanding of individuals' reactions to a question or set of questions, but is limited in the number of interviews that can be completed in a short amount of time and at a low cost. Cognitive interview participants are typically not selected based on probability methods and beyond demographics, there is usually not an understanding of how those respondents may differ from the population of interest for the study.

Crowdsourcing allows for the option of recruiting many participants quickly and cheaply. Participants can be provided a question or set of questions to consider and to which they can react. Using an existing crowdsourcing platform, these interviews can be self-administered over the web to provide reactions to questions as quickly as in one afternoon.

We will discuss results of a pilot test of this method of testing survey items using crowdsourcing including advantages (speed, cost, and number of participants) and disadvantages (less control over who participates, loss of context of items in a larger survey) as compared to traditional cognitive interviewing methods.

Case Study Comparing Data Collected via Crowdsourcing vs. Trained Data Collectors

Annice Kim (RTI International), Allie Lieberman (RTI International), and Daniel Dench (RTI International)

Research Purpose: Researchers and state health agencies have traditionally deployed a small group of trained data collectors to hundreds of stores state-wide to collect information about how tobacco products are marketed in retail stores. The retail store is a complex and fast-changing environment and increasingly, more frequent, timely, and local-level data collection efforts are needed to monitor emerging tobacco products and to inform state/local government efforts to regulate tobacco advertising. Crowdsourcing may provide a faster and more cost-effective alternative to using trained data collectors but the quality of crowdsourced data is unknown. The purpose of this study is to compare the quality of data collected via crowdsourcing vs. trained data collectors.

Data Sources and Statistical Methods: As part of RTI's Retail Advertising Tobacco Study for the Florida Department of Health, trained surveyors conducted in-store audits of licensed tobacco retailers located in Miami and Tampa during a 3-week period (8/31/2012 to 9/19/2012). Using crowdsourcing application *Gigwalk,* we posted 170 jobs to residents in the Miami and Tampa area to collect a subset of measures at the same stores during the same time period. Each job ("gig") required the worker to visit a specific retail store and collect information and photos on tobacco product availability, advertising, and promotional offers. We computed percent agreement and kappa statistics to compare the quality of data collected via untrained crowdsourced data collectors vs. trained data collectors.

Conclusions: More than half of the gigs were completed within 24 hours of posting the job online. For most of the measures examined, there was high agreement (77.5% - 98.0%) between the untrained crowdsourced data collectors and trained data collectors with kappa statistics in the moderate to high range (0.4 – 0.7). These findings suggest that crowdsourcing may be a promising alternative method for collecting local-level observational data.

The Census Bureau Mail Return Rate Challenge: Crowdsourcing to Develop a Hard to Count Score

Chandra Erdman (U.S. Census Bureau) and Nancy Bates (U.S. Census Bureau)

In 2012, the Census Bureau posed a challenge under the *America Competes Act*, an act designed to improve the competitiveness of the United States by investing in innovation through research and development. The agency tapped Kaggle.com to host and manage a world-wide competition to develop the best statistical model to predict 2010 Census mail return rates. The agency provided competitors with a block-group level database consisting of housing, demographic, and socio-economic variables derived from Census 2010, five-year American Community Survey estimates, and Census 2010 operational data. The agency then challenged teams to use these data (and other publicly available data) to construct the models. One goal of the challenge was to leverage winning models as inputs to a new model-based hard-to-count (HTC) score – a metric to target and stratify geographic areas according to propensity to self-respond in surveys and censuses.

The winning models all used machine learning and random forest techniques to predict mail return but lacked any theoretical underpinnings to explain relationships between predictors and response. This made the models hard to interpret and impossible to directly translate to a HTC score. Nonetheless, the winning models contained insights toward building a new OLS regression-based score using variables from the database. This paper describes the original algorithm-based HTC score, insights gained from the Kaggle Challenge, and the model underlying the new, block-group-level HTC score.

A Methodological Framework for Crowdsourcing in Research

Michael Keating (RTI International) and Robert Furberg (RTI International)

The adaptation of crowdsourcing from commercial marketing in the private sector for use to support the research process is increasing, providing investigators a wealth of case studies from which to discern emerging best practices. The successes and failures of crowdsourced research have not yet been analyzed to provide guidance for those interested in these new methods, nor have these data been synthesized to yield a methodological framework. Drawing from relevant case studies for survey research, this presentation will provide an evidence-informed methodological framework that describes a set of concepts, assumptions, and practices to support investigators with an interest in conducting crowdsourced research.

Our case studies will cover the research lifecycle, beginning in the research design phase by examining open innovation challenges, moving to the implementation phase of supplemental data collection, and concluding with data analysis challenges. Successful implementations of crowdsourcing require the researcher to consider a number of dimensions, including clearly articulated research goals, determination of the type of crowdsourcing to use (e.g. challenge or microtask labor), identification of the target crowd, an assessment of the factors motivating the crowd, and a determination of where to apply crowdsourcing results in the overall research lifecycle. Without a guiding framework and process, investigators risk unsuccessful implementation of crowdsourced research activities. This presentation offers an initial step toward a more standardized crowdsourcing method.

CONCURRENT SESSION H-2

LINKED DATA FOR ANALYSIS AND SURVEY EVALUATION

A Comparison of Person-Reported Industry to Employer-Reported Industry in Survey and Administrative Data

Emily Isenberg (U.S. Census Bureau), Liana Christin Landivar (U.S. Census Bureau), and Esther Mezey (Axioma, Inc.)

We compare self-reported job industry in the American Community Survey (ACS) to employer reported industry from the Quarterly Census of Employment and Wages (QCEW) that is part of the Census Bureau's Longitudinal Employer-Household Dynamics (LEHD) Program. The QCEW firm level data, including industry, is linked to quarterly data on individual workers, primarily from unemployment insurance (UI) wage records. We link individual work histories from the LEHD Program to ACS data using Protected Identification Keys (PIKs) created by the Census Bureau. If an individual works in more than one LEHD job around the time of the ACS interview, we use ACS interview date and LEHD earnings to select a main LEHD job to use for industry comparisons. Using 2009 data, we find an overall industry sector match rate of 75 percent, and a match rate at the 4-digit Census Industry Code (CIC) level of 61 percent. Industry match rates vary by sector and by whether industry sector is defined using ACS or LEHD industry information. Two of the largest sectors, educational services and health care and social assistance, also have some of the highest industry sector match rates. The wholesale trade sector has relatively low match rates, especially for the LEHD sector, and jobs in this sector in the LEHD data are often categorized in the manufacturing or retail trade sectors in the ACS data, and vice versa.

The National Household Survey: Evaluating Data Quality in a Large Voluntary Survey

Sander Post (Statistics Canada)

Statistics Canada conducts a Census every 5 years. In recent Censuses, including the 2006 Census, a short questionnaire was given to 80% of dwellings, and a long questionnaire to the other 20%. Both forms were mandatory.

The collection approach was changed for 2011. In 2011, all dwellings received a mandatory short questionnaire, and a 33% sample also received a voluntary long questionnaire. The voluntary questionnaire is officially a separate survey called the National Household Survey (NHS).

It was unknown in advance what kind of response rate would be obtained for the NHS. Given an initial nonresponse rate of over 30%, later improved with subsampling and follow-up, there were many questions about the impact of nonresponse bias on data quality.

One tool used to evaluate this impact was a record-linkage of 2011 NHS respondents and 2011 NHS nonrespondents to 2006 long form respondents. NHS nonrespondents were linkable as we had their (mandatory) census responses, and the Census and NHS questionnaires for a household were linked. Common variables from 2011 and 2006 were then used for record linkage. As expected, we were able to evaluate and account for nonresponse bias. Less expected, we also learned about other data quality issues, such as proxy response bias.

As part of our commitment to data quality, we will advise users that due to the changes in methodology, NHS data is not strictly comparable to previous long form Census data. The first NHS releases are in May, June, and August 2013. It will be interesting to see what issues arise from data users who do not heed this advice, and what lessons there are to be learned for data users and data producers.

Enhancing the Medical Expenditure Panel Survey Through Data Linkages

Lisa Mirel (Agency for Healthcare Research and Quality) and Steven Machlin (Agency for Healthcare Research and Quality)

Linkages between household surveys and other data sources can enhance the analytic capabilities of a survey. In particular, these efforts can help improve data quality and provide the basis for methodological studies. This paper will describe selected data linkage enhancements to the Medical Expenditure Panel Survey Household Component (MEPS-HC). The MEPS-HC is a complex multi-stage nationally representative sample of the U. S. civilian noninstitutionalized population. Each year a new sample is drawn as a subsample of households that participated in the prior year's National Health Interview Survey (NHIS), which results in a natural link between these two surveys. Data are collected in the MEPS-HC through a series of five interviews that cover a variety of topics, including health conditions, use of medical care services, charges and payments, and access to care. This paper will focus on two types of linkages using the MEPS-HC. The first links a subset of MEPS sample persons that were selected and signed permission forms to contact medical providers for the MEPS Medical Provider Component (MEPS-MPC). The main focus of the MEPS-MPC is to collect data on charges and payments for care provided to sample persons. The MEPS-MPC data are linked to the MEPS-HC and used to supplement, replace or impute expenditure data and for methodological analyses. We will discuss the details of matching MEPS-HC to MEPS-MPC data, including variables used and the calculation of match weights. The second type of linkage that we will discuss is how drawing the MEPS-HC sample as a subsample from the NHIS offers the opportunity to link some MEPS sample persons to administrative record data sources (e.g., National Death Index and Medicare and Medicaid claims). We will describe the details of these linkages and potential research based on these types of linked files.

Characteristics of Medicare Beneficiaries with Part D Coverage: NHANES 2003-2004 Linked to 2006 Medicare Part D Data

Ryne Paulose (National Center for Health Statistics), Hannah Day (National Center for Health Statistics), and Steven Frenk (National Center for Health Statistics)

Background: Medicare Part D, also called the Medicare prescription drug benefit, went into effect on January 1, 2006. The National Health and Nutrition Examination Survey (NHANES) data has recently been linked to medication utilization and expenditure information from Medicare Part D data.

Objectives: Describe the population available from the linkage of NHANES data to Medicare Part D data and to compare demographic and health characteristics of Medicare beneficiaries aged 65 years and older with and without Medicare Part D coverage.

Methods: Data from NHANES 2003-2004 linked to the 2006 Medicare Part D data were analyzed. NHANES data were used to compare characteristics of Medicare beneficiaries aged 65 years and older at the NHANES interview by Medicare Part D enrollment status. Participants were classified as enrolled in Medicare Part D if he or she had at least 1 month of Part D enrollment in 2006.

Results: During NHANES 2003-2004, 1,494 participants were aged 65 years and older. Of these 86.7% (n=1,296) were linkage eligible and 13.2% were linkage ineligible (n=198). Linkage eligibility means that there was sufficient information provided by the NHANES respondent to allow their records to be linked to the CMS data. Of those linkage eligible, 87.8% (n=1,138) matched to the 2006 Medicare Part D Denominator file, 1.7% (n=22) did not match to Medicare in 2006, and 10.5% (n=136) died between the survey period and December 31, 2005. Among the matched sample, 59.9% (n=682) had Part D coverage at any time during 2006 and 40.1% (n=459) had no coverage.

Conclusion: Nearly 90% of linkage eligible NHANES Medicare beneficiaries were matched to the Medicare Part D records. However, not all beneficiaries had Part D coverage. Linked data, with questionnaire, exam and claim information may be used to further evaluate differences between beneficiaries enrolled in Part D and those not enrolled.

CONCURRENT SESSION H-3

HEALTH INSURANCE AND HOUSEHOLD INCOME IN THE CURRENT POPULATION SURVEY (CPS) REDESIGN AND PRODUCTION

The 2013 Annual Social Economic Supplement Income and Health Insurance Questionnaire Test Sample Design

David Hornick (U.S. Census Bureau)

The Annual Social Economic Supplement (ASEC) to the Current Population Survey (CPS) is collected every year in the months of February, March, and April. In the ASEC supplement, we collect data on family characteristics, household composition, marital status, educational attainment, health insurance coverage, county of origin, work experience, income, noncash benefits, poverty, program participation, and geographic mobility. During the 2013 data collection period, we tested a new questionnaire related to income and health insurance coverage. This paper sets the stage for the comparisons of the income and health insurance test questionnaire to the currently used income and health insurance questionnaire. We will discuss the ASEC sample design, the ASEC Test sample design, and the weighting of the data to make valid comparisons between the current and proposed income and health insurance questionnaire.

Non-response Bias in the March 2013 Content Test

Matthew Brault (U.S. Census Bureau)

In March 2013, the Census Bureau fielded a content test to evaluate new income and health insurance questions using a sample of 23,000 individuals drawn from participants of the 2011 and 2012 Current Population Survey Annual Social and Economic Supplement (CPS ASEC). The response rates in the test instrument were very low, which would potentially introduce bias in comparisons to estimates from the regular CPS ASEC. During the time between a sample case's last CPS ASEC interview and content test survey, individuals may have moved for reasons correlated with characteristics being studied. For instance, if low-income individuals disproportionally fail to respond, estimates of Medicaid coverage and income could be biased. The design of the content test is similar to that of the American Time Use Survey (ATUS), which has also suffered from low response, however, research on the sample design has shown that nonresponse bias did not greatly affect key ATUS estimates (Abraham, Maitland, and Bianchi 2006). Following a similar approach, I use logistic regression models to predict survey response, contact, and cooperation rates from sample characteristics obtained during their last CPS ASEC interview. From these models, I adjust the survey weights to account for the response propensities and compare them to unadjusted estimates to determine the effect of nonresponse bias on key estimates about sources of income and health insurance coverage.

Evaluating the Income and Health Insurance Questions 2013 CPS ASEC Content Test Using Timer Data

Aaron Cantu (U.S. Census Bureau) and Adam Bee (U.S. Census Bureau)

The Annual Social Economic Supplement (ASEC) to the Current Population Survey is one of the most widely used household surveys publishing national level estimates for income, poverty and health insurance. The Census Bureau is conducting a field test in March 2013 of a redesigned instrument. This paper will look into the functionality of the new instrument. The survey instrument automatically generates timer data. Timer data tells the analyst how long a section takes. The paper will use timer data, completion rates, and dropout rates to compare the ASEC and the redesign for the income and health insurance sections. The rest of the analysis is for the redesigned instrument only.

Evaluating the 2013 CPS ASEC Income Redesign Content Test

Jessica Semega (U.S. Census Bureau)

:

The Annual Social and Economic Supplement to the Current Population Survey (the CPS ASEC) is one of the most widely used economic surveys publishing national level medians and means on household income. The CPS ASEC is the official source of the US national poverty statistics. The ASEC asks each person detailed questions categorizing income into over 50 sources. As one of the nations' longest running surveys, it has been over 30 year since the last major redesign. In an effort to take better advantage of an automated environment and to update questions on retirement income and health insurance, the Census Bureau is conducting a field test in March 2013 of a redesigned instrument using a retired ASEC sample of 23,000 households. This paper will look into the outcomes of the field test and the specifics of the redesigned portion of the ASEC instrument, including the use of income brackets for respondent question refusals to income amounts (currently the instrument will move on to the next source of income) and a new "dual pass" approach that asks all income sources first, followed by income amount collection.

Health Insurance in the Current Population Survey: Redesign and Production

Carla Medalia (U.S. Census Bureau), Amy Steinweg (U.S. Census Bureau), Brett O'Hara (U.S. Census Bureau), David Lee (U.S. Census Bureau), Jessica Smith (U.S. Census Bureau), Joanne Pascale (U.S. Census Bureau), and Jonathan Rodean (U.S. Census Bureau)

The Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) generates widely used estimates on health insurance coverage and the uninsured. However, research has shown that the calendar-year measure is flawed and that estimates reflect a mixture of current and past year coverage. To address this concern, as well as to take advantage of technological advances in survey implementation and to add important new content to the instrument, the Census Bureau substantially redesigned the CPS ASEC health insurance module, which will be fielded in a large national test in March 2013. This paper compares the estimates of key health insurance estimates from the production ASEC instrument to the redesigned instrument. Specifically, we examine uninsured rates and rates of insurance by plan type, such as employer-provided plans, direct-purchase plans, Medicare, Medicaid, and VA coverage. Future research using the redesigned instrument, which may replace the CPS ASEC in 2014, will consider the implications of these changes when analyzing health insurance

CONCURRENT SESSION H-4

INNOVATIONS IN FRAME DEVELOPMENT

Sample Design Considerations for the Occupational Requirements Survey

Brad Rhein (Bureau of Labor Statistics), Chester Ponikowski (Bureau of Labor Statistics), and Erin McNulty (Bureau of Labor Statistics)

The Bureau of Labor Statistics (BLS) is working with the Social Security Administration (SSA) to carry out a series of tests to determine the feasibility of using the National Compensation Survey (NCS) platform to accurately and reliably capture data that are relevant to the SSA's disability program. The proposed new Occupational Requirements Survey (ORS) is envisioned to be an establishment survey that collects information on the vocational and physical requirements of occupations in the U.S. economy as well as the environmental conditions in which those occupations are performed. While NCS is also an establishment survey, sampled yearly from a national frame using probability proportionate to establishment employment size, it is unclear whether the NCS sample design will meet the goals of ORS. This paper discusses the advantages and disadvantages of integrating the sample design of ORS with the sample design of NCS, or whether an independent sample design for ORS would be more appropriate.

Exploratory Research on the Use of Google Earth to Create a Sampling Frame of Buildings

Katie Lewis (Energy Information Administration)

Since its inception, a majority of the sampling frame for the Commercial Buildings Energy Consumption Survey (CBECS), conducted on a quadrennial basis by the Energy Information Administration (EIA), has been created or updated using traditional area listing. The frame is expensive to maintain because field workers travel to selected area segments and list all eligible commercial buildings in those segments; for each building, information for locating the building at the interview phase and variables needed for sampling are recorded.

This presentation discusses our exploratory research on one potential less expensive alternative method for creating the frame: the use of Google Earth. In this project, we selected a sample of segments listed in the Fall of 2012, then a sample of listed buildings within those segments, and attempted to reproduce a subset of the variables the listers recorded for those buildings. We discuss our methods, comparisons of the frame variable values, the nature of the Google Earth images (e.g., their ages), the difficulties we encountered in this task, and recommendations for further research.

Surveying Community Stakeholders: Exploring the Viability of a National Sampling Frame

Barbara Robles (Board of Governors of the Federal Reserve System)

Community stakeholders are opinion makers, change agents and local leaders knowledgeable about current and emerging community economic conditions. Community economic development stakeholder surveys are often administered only at the local level by grant-funded foundation-university-public sector and non-profit organization partners. Some of these community surveys are called 'asset-mapping' or 'community needs assessments' and focus on limited respondent pools with limited questionnaires designed to capture a very narrow community sector issue or service.

One of the ways in which local economic conditions occurring at the community-level can be captured for a more aggregated perspective is to identify public domain data sources, at the local, municipal, state and regional levels. Once the universe of community stakeholders is minimally identified, one can employ public domain data sources aggregated into regional districts to determine the respondent pool. Using web tools, an internet-based survey designed to capture current and emerging community economic development conditions can be administered to varying sectors in the format of a stratified random sample. Community sectors would encompass: non-profit organizations, faith-based services, and public sector agencies that provide education services, health services, legal services, housing services, transportation services, food banks and social services. In addition, community banking services and small community-based businesses will be included in the sampling frame.

Preliminary findings comparing a national and a regional approach will be discussed. In addition, methodological approaches to triangulating survey findings using text and data analytics from separately administered regional community stakeholder surveys is explored. Future research questions addressing methodologies for implementing stratified clustering of the sampling frame will be raised.

Redesigning National School Surveys: Coverage and Stratification Improvement Using Multiple Datasets

William Robb (ICF International), Kate Flint (ICF International), and Ronaldo Iachan (ICF International)

This paper discusses the redesign of two national school surveys with a focus on sampling frame development. Historically, the sampling frames for these surveys have been developed from files acquired from a commercial vendor. Vendor files are used as they contain up-to-date contact information, facilitating recruitment efforts in support of high response rates. The vendor provided data set incorporates several sets of variables sourced from the NCES Common Core Data (CCD) and Private School Survey (PSS).

This paper explores the construction of an augmented frame, built by combining the vendor list with NCES files directly, using schools listed on the CCD for public schools and the PSS for non-public schools. A particular challenge is assembling an unduplicated frame for non-public schools in the absence of a unique identifier common across datasets. In addition, with multiple sources for both districts and schools, identifying a unique set of school-district relationships is difficult.

We provide an assessment of eligibility rates from each data source, and assess coverage improvements by grade, school type and geography.

CONCURRENT SESSION I-1

MOBILE DEVICES AND APPLICATIONS FOR SURVEY DATA COLLECTION

Bring Your Own Device and the 2020 Census Research and Testing

Evan Moffett (U.S. Census Bureau), Jay Occhiogrosso (U.S. Census Bureau), Scott Williams (U.S. Census Bureau), and Ryan King (U.S. Census Bureau)

The US Census Bureau is exploring the "bring your own device" (BYOD) concept for the 2020 Census. The Census is developing applications that run on commercial devices. To date, a Listing and Mapping application and an Enumeration application have been developed to run on Apple iOS, Google and Microsoft mobile Windows operating systems. This software was also developed to run on multiple form factors, including tablets and smartphones.

This paper will present the results of various tests conducted with the different devices and operating systems. The results of usability testing will be presented. The status of an enterprise information technology effort to deploy applications in a secure configuration simulating BYOD will also be presented. Finally, this paper will discuss preliminary survey results of the general public's willingness to use their own device in a work setting, such as a Census.

The goal of this work is to systematically establish a capability to develop and deploy secure applications in support of the 2020 Census field data collection operations.

Field Data Collection in Area Frame Surveys Utilizing iPads® - USDA's June Area Survey

Michael Gerling (National Agricultural Statistics Service), Eric Wilson (National Agricultural Statistics Service), Linda Lawson (National Agricultural Statistics Service), Alan Dotts (Iowa State University), and Andrew Vardeman (Iowa State University)

The United States Department of Agriculture National Agricultural Statistics Service's (USDA-NASS) June Area Survey is an area frame based survey conducted by field interviewers using paper questionnaires and an aerial photo of a sample segment of land. Interviewers visit land operators and draw out crop fields, pastures, woods, wasteland, etc. on the aerial photo. In addition, information on agricultural activity occurring within the designated area is collected on the questionnaire. NASS is currently developing a CAPI data collection instrument which eliminates the need for hard copies of the aerial photo and paper questionnaires. The field interviewer uses an iPad which displays the area of interest and provides the ability to draw off various agricultural fields and other land uses using a geographic information system (GIS) layer. Also, the field interviewer records survey responses on the land utilization within the iPad. The CAPI instrument is a hybrid of a thick/thin-client web application designed to collect data in both on-line and off-line situations. This presentation will show NASS's use of iPads for data collection in this survey and compare the results obtained from the GIS CAPI instrument to data reported by the farm operators for the same land.

The iPad® Computer-Assisted Personal Interview System - A Revolution for In-Person Data Capture?

Heather Driscoll (ICF International) and James Dayton (ICF International)

In-person interviewing has long utilized paper-and-pencil surveys as the data collection mode for observational studies. At a time of increased scrutiny from the public and rising costs, electronic data collection devices are dramatically changing the landscape of these types of studies. ICF has conducted several pilot studies using our iPad® Computer-Assisted Personal Interview system (iCAPI) since 2010 and found that it allowed for more efficient data collection, monitoring, cleaning, and analysis.

Most recently, ICF conducted a study of the economic impact of Pennsylvania's water trails on the state's economy. This was our first complete implementation of our newly developed iCAPI. Our interviewers surveyed visitors to water trails (rivers that have been designated as a recreational water trail because they are important corridors between specific locations) at hundreds of boat and kayak launch sites during the summer of 2012. Through record heat waves, intense thunderstorms, and unpredictable site conditions, our interviewers successfully collected expenditure data from roughly 400 water trail visitors, using the iCAPI.

Our most recent work in Pennsylvania confirmed and expanded on what we learned in our pilots, addressing questions, such as how easily can interviewers pick up the iCAPI system; how effective are the GPS and map capabilities; how do the iPads perform over weeks of data collection; and, is the iCAPI system greener, faster, better and cheaper? We were surprised by some in-field scenarios that were resolved with iCAPI; however, is the iCAPI system the perfect, sustainable intercept solution? Our paper will explore the advantages and limitations, as well as our ideas for refining the next iteration of applications for our iCAPI.

Innovative Retention Methods in Panel Research: Can SmartPhones Improve Long-term Panel Participation?

James Dayton (ICF International) and Andrew Dyer (ICF International)

Minimizing participant attrition is vital to the success of longitudinal panel research. One such example of longitudinal panel study conducted by ICF is the National Recreational Boating Survey (NRBS), sponsored by the U.S. Coast Guard to ensure that the public has safe, secure, and enjoyable recreational boating experiences. Specifically, the NRBS Program enables the Coast Guard to better identify safety priorities and coordinate and focus research efforts. The project features a several components, one such being the "Trip Panel". The Trip Panel is designed to capture actual exposure to recreational boating. This panel was recruited via dual-frame, dual-mode (Random Digit Dial telephone and mail) and has been in place for over a year. Respondent contact information includes e-mail address, mailing address, and telephone. In many cases, the provided contact number is a mobile device.

This presentation will explore ICF researchers' quest to improve panel retention though the introduction of a SmartPhone application that engages respondents in between survey waves by allowing them to communicate changes in contact information and even provide survey responses via SmartPhone rather than via the Web or traditional telephone. Active panelists who provide cell phone contact information will be randomly assigned to receive standard retention communications via mail, phone and e-mail (control) or alternate retention communications via a SmartPhone application and text message/SMS (treatment). The communications application for the treatment group includes study updates, various interactive communications, and mini-surveys.

ICF researchers will analyze the differences in control and treatment panel retention over a six-month period. We will also survey panelists' willingness to sign-on for another annual wave of the panel as well as their overall satisfaction with panel participation as an indicator of long-term continued participation.

CONCURRENT SESSION I-2

INTEGRATED APPROACHES TO DATA EDITING, IMPUTATION, AND DISCLOSURE CONTROL

Simultaneous Edit-imputation for Continuous Microdata

Hang Joon Kim (Duke University and National Institute of Statistical Science), Jerome Reiter (Duke University), Alan Karr (National Institute of Statistical Science), and Larry Cox (National Institute of Statistical Science)

We present a fully Bayesian, joint modeling approach to simultaneous editing and imputation of faulty and missing values for continuous microdata. In many surveys with continuous variables, especially economic surveys, the variables must satisfy systems of linear constraints, for example ratio edits. When records fail one or more edits, statistical agencies typically use a two-stage process of (i) localizing the particular values in error—most commonly using Fellegi-Holt routines—and (ii) imputing new values for those fields usually using a hot deck imputation. Our approach replaces the two step process with a single probability-based, data-driven approach in which we (i) specify a flexible joint probability model (a mixture of multivariate normal distributions) for the continuous variables, which can capture more complex associations than typical hot deck imputation schemes, (ii) use the model to identify values plausibly in error, which, unlike Fellegi-Holt routines, can reflect uncertainty over the unknown faulty values when making corrected data, and (iii) impute new values from the model in ways guaranteed to satisfy all linear constraints. In this talk, we describe this integrated edit-imputation approach. We illustrate it using manufacturing data from a survey in Columbia, South America. We compare the approach against Fellegi-Holt error localization methods, showing how the model-based approach can offer improved accuracy.

Simultaneous Edit-imputation for Categorical Microdata

Daniel Manrique-Vallier (Indiana University) and Jerome Reiter (Duke University)

We present a fully Bayesian, joint modeling approach to simultaneous editing and imputation of faulty and missing values for categorical microdata. Typically, certain combinations of categorical variables are *a priori* impossible, for example married three year old children. However, due to reporting or other errors, collected data often include such structural zeros. Many agencies use a Fellegi-Holt approach, or some approximation thereof, to select the categorical values to edit, followed by a hot deck imputation to make non-faulty records. We propose to use a latent mixture of multinomial distributions to model the underlying data, ensuring the distribution has no support over structural zeros. This model offers opportunities to capture complex dependencies that can be missed by hot decks and other imputation strategies (e.g., log-linear models and sequential regression imputation methods). We overlay a measurement error model that lets the data select values that are plausibly in error, and impute replacements using the mixture model that are guaranteed not to be structural zeros. We demonstrate performance of the approach using data from a public use microdata sample of the census of New York.

Statistical Disclosure Limitation and Edit-Imputation

Alan Karr (National Institute of Statistical Science), Hang Joon Kim (Duke University and National Statistical Institute of Science), Jerome Reiter (Duke University), and Larry Cox (National Institute of Statistical Science)

When releasing data to the public, statistical agencies typically have to deal with missing data, edit constraints, and disclosure limitation. While individually these topics have been heavily studied, there has not been much research on systems that integrate all three simultaneously. We present results of simulation studies illustrating the data quality of several such systems, including (i) edit-imputation first followed by disclosure limitation, (ii) disclosure limitation followed by edit-imputation, and (iii) simultaneous edit-imputation-disclosure control as a single process. We illustrate the first two approaches using common disclosure limitation strategies including swapping, noise addition, and microaggregation. We illustrate the third approach using multiple imputation methods for all three tasks (error localization,

imputation, and disclosure limitation). We illustrate the approaches using data from manufacturing surveys.

CONCURRENT SESSION I-3

EXAMINING THE IMPACTS OF SAMPLE DESIGN: COST, BIAS, VARIANCE AND OTHER SURVEY CONCERNS

Optimal Cutoff Sampling for the Annual Survey of Public Employment and Payroll

Brian Dumbacher (U.S. Census Bureau) and Carma Hogue (U.S. Census Bureau)

The goal of cutoff sampling is to save cost, reduce respondent burden, and maintain accuracy of estimates by reducing the number of small units in sample. For the Annual Survey of Public Employment and Payroll, the Governments Division of the U.S. Census Bureau uses a modified version of cutoff sampling in which a subsample of units below the cutoff are selected. In this paper, we examine a numerical method based on minimizing the sum of mean squared errors from linear regression models to find an optimal combination of cutoff and subsampling rate given a specified cost. Data from the 2002 and 2007 Censuses of Governments: Employment are used for this study.

Numerical Impact of a Simple Random Subsample on Consumer Spending for Children

Daniel Yang (Bureau of Labor Statistics) and Jeffrey Gonzalez (Agency for Healthcare Research and Quality)

The Bureau of Labor Statistics is in the process of redesigning the Consumer Expenditure Survey (CE). The primary goal of this effort is to reduce measurement error while neither increasing data collection costs nor imposing a greater burden on respondents. One redesign option that was initially considered was a split questionnaire design. The implementation of the split questionnaire we considered was a simple random subsample (SRSS). To evaluate the impact of this approach, we simulated data following a SRSS approach and examined to what extent, if any, it might affect data users and their respective economic analyses. One particular use of CE data is to characterize spending on children by different types of households. To do this, we compared estimates of children's expenditures derived from the full CE sample to those based on the simulated SRSS approach. Specifically, we examined results from descriptive statistics, logistic and linear regressions. We also explored the use of a Cragg's two stage model for estimating the marginal propensity to consume and income elasticities of expenditures on children. We found that compared to the full sample, the SRSS produces different estimates of expenditure means and generally higher standard errors as well as different estimates for regression coefficients. Finally, based on our findings, we provide recommendations for general design changes to the data collection procedures and statistical analysis methods.

Effects of Rotation Group Bias on Estimation of Smoking Prevalence

Younghwan Song (Union College)

The Tobacco Use Supplement to the Current Population Survey (TUS-CPS) has been administered as part of the Current Population Survey (CPS). The TUS-CPS has been a key source of national and state-level data on smoking in the U.S. household population because it is based on a large, nationally representative sample of the CPS. The CPS has eight rotation groups of households in each month that are repeatedly interviewed based on a sample rotation scheme. Previous research has found that, even though all eight rotation groups in the CPS are independent random samples of the population, some estimates, such as unemployment rates, tend to be significantly higher in the first rotation group than among other rotation groups. It has also been shown that the estimators using the full sample are biased unless these systematic biases across rotation groups cancel each other out.

This paper examines if smoking prevalence estimates based on the TUS-CPS suffers from rotation group bias because the TUS-CPS is based on the CPS. Furthermore, considering the finding in previous research that tobacco users are more likely to attrite in longitudinal surveys, this paper investigates if the

main reason for rotation group bias in the TUS-CPS is because smokers are less likely to be interviewed in the successive rotation groups in the CPS. If so, estimates using the full sample are likely to underestimate smoking prevalence. Using matched TUS-CPS and CPS data sets, this paper tests if smokers are less likely to stay in the subsequent TUS-CPS and CPS. This paper also compares estimates of smoking prevalence between the new/returning respondents and continuing respondents in the TUS-CPS. The results of this paper provide measures to improve the estimation of smoking prevalence using the TUS-CPS.

The Impact of Cellphone Sample Representation on Variance Estimates in a Dual-Frame Telephone Survey

A. Elizabeth Ormson (NORC at the University of Chicago), Kennon Copeland (NORC at the University of Chicago), Stephen Blumberg (National Center for Health Statistics), and Nadarajasundaram Ganesh (NORC at the University of Chicago)

Incorporating a sample of cellphone numbers ("cellphone sample") is a critical issue for any random-digit-dialed telephone survey. This is due to the dramatic increase in the percentage of households that have only cellphones, which has more than doubled in just four years, from 17.5% in early 2008 to 35.4% in early 2012 (Blumberg, Luke 2012). Optimum sample allocation of landline and cellphone sample, while the gold standard, can cause sticker shock due to the additional costs associated with fielding cellphone samples. Therefore, survey designers must determine how much cellphone sample to include in a telephone survey to achieve a reasonable level of precision, given specific cost constraints.

A common rule-of-thumb is to achieve a distribution of cellphones in the survey that is approximately half the population distribution percentage. For example, if cellphone-only (CPO) households constitute 36% of the population, then having 18% of the sample from the CPO domain, though not optimal, is assumed to allow direct weighting to this population with minimal impact on variance estimates.

This paper will review the impact of sample design and cellphone sample distributions on estimate variance given specific cost constraints. We will describe a simulation study that used data from the 2011-2012 National Survey of Children's Health, which is sponsored by the Maternal and Child Health Bureau and conducted by the National Center for Health Statistics. The simulation manipulated various levels of cellphone sample inclusion and two cellphone sample selection approaches (interviewing all cellphone respondents versus screening for the CPO population) and examined the impact on variance estimates.

CONCURRENT SESSION I-4

USE OF CAPTURE-RECAPTURE METHODS IN FEDERAL CENSUSES

Reducing Nonsampling Error in the Census Coverage Measurement

Tamara Adams (U.S. Census Bureau), Magdalena Ramos (U.S. Census Bureau), and Gia Donnalley (U.S. Census Bureau).

As part of the 2010 Census, the US Census Bureau conducted the Census Coverage Measurement, a capture-recapture study designed to measure the coverage error of housing units and persons. For the first time, we also published estimates of components of census coverage error. One of the primary goals of the 2010 Census Coverage Measurement was to reduce nonsampling error while supporting the goals of using dual-system estimation. During both the 1990 and 2000 censuses, nonsampling error issues caused the data from the coverage studies to need revision. We approached the problem in two ways: initial design changes to the Census Coverage Measurement and a sample cut to allow additional effort to be spent on reducing nonsampling error. In this paper, we discuss the measures taken and the effectiveness of those measures, where data are available.

Accounting for Missing Data in the Census Coverage Measurement Survey

Vincent Thomas Mule Jr. (U.S. Census Bureau) and Scott Konicki (U.S. Census Bureau)

The 2010 Census Coverage Measurement (CCM) program produced net coverage results showing undercounts or overcounts using dual system estimation. Additionally, the CCM program produced components of census coverage that included erroneous enumerations and omissions. This paper summarizes the methods and results to account for missing data. First, we will focus on the imputation of missing demographic characteristics and tenure. Second, we will focus on using logistic regression models to account for unresolved statuses for dual system estimation. Finally, we will focus on the imputation cell methods to account for the unresolved statuses for the estimates of the components of census coverage. We will highlight key results of these different methods.

Unresolved Matched Records in Capture-Recapture Methodology

Andrea Lamas (National Agricultural Statistics Service), Linda Young (National Agricultural Statistics Service), Denise Abreu (National Agricultural Statistics Service), Shu Wang (University of Florida), and Daniel Adrian (National Agricultural Statistics Service)

The National Agricultural Statistics Service (NASS) conducts a Census of Agriculture every 5 years, in years ending in 2 and 7. It is based on NASS's list of farm operations. The Census of Agriculture needs to be adjusted for non-response, under-coverage and misclassification. Capture-recapture is evaluated as the adjustment methodology. However, capture-recapture requires a second, independent survey. The second survey used in this methodology is NASS's June Area Survey (JAS) which is conducted annually in June. It is an area frame based survey with in-person interviews and is run independently from the Census of Agriculture.

In order to use the capture-recapture framework, a matched dataset of the Census of Agriculture and JAS is needed. This dataset is the foundation of modeling the probability of a JAS farm being captured by the Census of Agriculture. In the dataset, the farm status based on the JAS and the Census of Agriculture agree in most cases; however, in other cases, a record is a farm on either the JAS or Census and not a farm on the other.

In order to resolve the unresolved cases, a logistic model of the probability an operation is a farm based on the records with resolved farm status is developed. The final model is used to estimate the probability that each of the agricultural operations with unresolved farm status is a farm. These probabilities are used to adjust the weights of the unresolved records, which are then used in modeling the probability of capture and probability of misclassification in the capture-recapture adjustment.

Incorporating Misclassification into Capture-Recapture Methodology in the 2012 Census of Agriculture

Daniel Adrian (National Agricultural Statistics Service), Linda Young (National Agricultural Statistics Service), Denise Abreu (National Agricultural Statistics Service), Shu Wang (University of Florida), and Andrea Lamas (National Agricultural Statistics Service)

Capture-recapture methods, which originate from the study of fish and wildlife populations, have also been used to investigate undercount in the U.S. Census of Population for several decades. The National Agricultural Statistics Service (NASS), an agency within USDA, has further adapted this methodology to adjust the farm counts of the 2012 Census of Agriculture for undercoverage, nonresponse, and misclassification. This last issue, farm/non-farm misclassification, based on defining a farm as an agricultural operation with at least $1,000 in annual sales (or the potential to have such sales), requires modifications to traditional capture-recapture methods: for example, there is no analogous misclassification problem in the Census of Population (e.g. "a person is a person"). We emphasize such adaptations in our description of NASS's capture-recapture methodology.

CONCURRENT SESSION J-1

ANALYTIC CONSIDERATIONS OF DATA QUALITY IN RECORD LINKAGE

The Impact of Name Quality on Record Linkage

Mary Layne (U.S. Census Bureau), Deborah Wagner (U.S. Census Bureau), and Cynthia Rothhaas (U.S. Census Bureau)

The Census Bureau's Person Identification Validation System (PVS) assigns unique person identifiers to federal, commercial, and survey data to facilitate linkages across files. PVS uses probabilistic matching to assign a unique Census Bureau identifier for each person. The quality of input data affects the ability of a record to match. In this study, we examined the quality of the input name data for records that failed to link. This paper describes name data quality issues, as well as other person-level data issues affecting the ability of a record to link.

Analytic Challenges with National Survey Data Linked to a State-Level Cancer Registry

Eric Miller (National Center for Health Statistics), Dean Judson (National Center for Health Statistics), Yulei He (National Center for Health Statistics), Jennifer Parker (National Center for Health Statistics), and Hannah Day (National Center for Health Statistics)

Although data linkage provides an opportunity to conduct analyses that are not possible using each contributing data source alone, it also produces additional complexity. The CDC's National Center for Health Statistics (NCHS), in collaboration with the Florida Cancer Data System (FCDS) and University of Miami, conducted a pilot linkage between the 1986-2009 National Health Interview Survey (NHIS) and 1981-2010 FCDS to examine risk factors and characteristics of Floridians who were diagnosed with cancer compared to those without cancer. Due to residential mobility in and out of Florida, there are analytic challenges in constructing an analytic sample, calculating appropriate weights, and accounting for the longitudinal component of both datasets. We will describe these and other challenges encountered with the survey-registry linked data and discuss methods that could be used to address them. Although these examples are specific to the NHIS-FCDS linkage, this linkage will provides insight into the complexities of analyzing national data linked to state administrative data.

Clean-up and Statistical Analysis of Sets of National Files

William Winkler (U.S. Census Bureau)

This paper describes new methods for modeling/edit/imputation and record linkage that can be used in cleaning up individual files. These methods can be combined with new iterative models for adjusting statistical analysis for linkage error, statistical matching, recognition of frequency patterns of good data and not-so-good data, extensions of the basic modeling/edit/imputation model, and restraints from multi-variable optimization. Combining the methods with new computational methods that are 10-1000 times as fast as previously existing algorithms, a team of 10-20 skilled individuals might process a set of national files in 1-3 months.

Evaluating Hispanic and Asian Name Algorithms as a Tool for Improving Linkage to the National Death Index

Dean Judson (National Center for Health Statistics), Eric Miller (National Center for Health Statistics), Jennifer Parker (National Center for Health Statistics), and Hannah Day (National Center for Health Statistics)

The most recent linkage of individual records from National Center for Health Statistics (NCHS)' population health surveys to the National Death Index (NDI), with data up to 2006, was released in 2009. Auxiliary data from the Social Security Administration's Death Master File, data from previous linkages to administrative records from the Centers for Medicare and Medicaid Services, and other direct data collection efforts were used to improve the 2009 NCHS-NDI linkage. In the upcoming 2013 NCHS-NDI

linkage, Hispanic and Asian name and nickname algorithms were tested. The purpose of this paper is to specifically compare the additional records created by the Hispanic and Asian name and nickname algorithms with the original source records to determine if additional correct "hits" (true positives) were captured on the NDI. In addition, if "hits" were captured on the NDI later determined to not be correct (false positives), then an assessment would be performed to determine if the false positives did active harm, or merely created additional noisiness in the later post-processing and clerical review phases. An attempt to estimate false negatives (non-"hits" that were later determined by post-processing and or clerical review to be true "hits") will be included.

CONCURRENT SESSION J-2

UNDERSTANDING AND ADDRESSING SURVEY NONRESPONSE

On Checking Whether Response is Ignorable or Not

Michail Sverchkov (Bureau of Labor Statistics)

Most methods that deal with the estimation of response probabilities assume either explicitly or implicitly that the missing data are ‘missing at random’ (MAR). However, in many practical situations this assumption is not valid, since the probability to respond often depends on the outcome value. The case where the missing data are not MAR (NMAR) can be treated by postulating a parametric model for the distribution of the outcomes under full response and a model for the response probabilities. The two models define a parametric model for the joint distribution of the outcome and the response indicator, and the parameters of this model can be estimated in principle by maximization of the likelihood corresponding to this distribution. Modeling the distribution of the outcomes under full response can be problematic since no data are available from this distribution.

Sverchkov (2008, 2010) proposed a new approach that permits estimating the parameters of the model holding for the response probabilities without modeling the distribution of the outcomes under full response. The present paper shows how this approach can be used for testing whether the response is MAR or NMAR.

The Utility of the Integrated Design of the Medical Expenditure Panel Survey to Inform Trends in MEPS Nonresponse

Frances Chevarley (Agency for Healthcare and Research Quality) and Karen Davis (Agency for Healthcare and Research Quality)

This paper analyzes estimates from the Medical Expenditure Panel Survey (MEPS) matched with the National Health Interview Survey (NHIS) and uses practical tools to inform MEPS nonresponse estimates. MEPS is a nationally representative panel survey studying health care use, access, expenditures, source of payment, insurance coverage, and quality of care. Each year a new panel begins and each panel has 5 rounds of data collection over 2 ½ years that cover a two-year period.

The goal of this paper is to inform trends in MEPS nonresponse. Because MEPS uses the NHIS, conducted by the National Center for Health Statistics as its sampling frame, estimates are produced using variables from both the NHIS and MEPS across different categories of the MEPS response categories. Data used are from the 2006-2010 NHIS data matched with the 2007-2011 MEPS files along with additional paradata. Non-response rates are analyzed by standard demographic categories and MEPS response categories.

The Impact of Efforts to Increase Response Rates on Survey Estimates

Karen Wessels (ICF International) and Frances Barlas (ICF International)

Survey response rates have been declining substantially over the last several decades leading to concerns regarding the reliability of survey estimates. In an attempt to counteract declining response rates, researchers have invested in longer fielding periods, increased survey reminders, and offered more

than one survey mode. Even with these techniques in place, response rates have continued to fall, sparking research on the impact of declining response rates on survey findings. Previous research has found that, although more rigorous survey methodologies employing techniques, such as a longer fielding period and monetary incentives, result in higher response rates than standard survey methodologies, the majority of survey estimates do not differ significantly.

The current study compares estimates from a large survey of military community members for participants who responded 1) within the first four weeks of the survey administration, 2) within the first eight weeks, and 3) at any time during the fielding period. In addition, the study observes differences by survey mode (Web or paper) and compares regression models to determine whether findings change with increased response rates. Results indicate minimal changes in survey estimates between data collected through the eight week mark and the close of the survey; differences were significantly larger between data collected through the four week mark and the close of the survey. In addition, differences in estimates tended to be larger for data collected via paper than data collected via web, and for specific demographic groups in the paper group. Results provide recommendations for fielding length and number of survey reminders, and considerations of possible nonresponse bias for web and paper surveys.

Changing the Way We Look at Survey Nonresponse

Deborah Griffin (U.S. Census Bureau) and Dawn Nelson (U.S. Census Bureau)

For decades, survey sponsors, data users, and other stakeholders have judged the success of a survey by its response rate. The Office of Management and Budget and the Department of Commerce cite this measure of nonresponse as a key performance metric. The Census Bureau's Statistical Quality Standards require that a survey achieve certain levels of response to demonstrate acceptable quality. Because of the value placed on response rates, they have become a major measure used to evaluate data collection operations. We evaluate the performance of telephone interviewers, field representatives, and their managers based largely on the response rates they achieve. This focus on response rates has many consequences. One is a high burden placed on respondents when a survey repeatedly contacts sample households. Another is the pressure on interviewing staff to spend time and resources to increase their response rates and not to accept a noninterview. The reliance on response rates has overshadowed the value of alternative measures that may be equally, or even more, important from a survey management perspective. In the past few years, survey methodologists have questioned the utility of the response rate as a measure of survey quality. There is also growing concern about the costs and burden associated with obtaining high response rates.

This paper summarizes historical measures of survey nonresponse in the American Community Survey (ACS). It shows that the attention given to response rates in the ACS has resulted in annual weighted survey response rates between 97 and 98 percent with consistently high response rates achieved across all areas of the country. If controlling for survey nonresponse (as measured by consistently high subnational response rates) is the most important metric, the ACS is very successful. This paper looks more closely at ACS nonresponse, specifically on nonresponse associated with the final ACS data collection mode, Computer Assisted Personal Interviewing (CAPI). Here too we find exceptionally high response rates of about 95 percent with little subnational variation. The Census Bureau employs a contact history instrument during CAPI data collection. Paradata from this instrument allow us to look below the surface at what is required, in terms of both respondent burden and costs, to obtain these high response rates. We use these paradata to propose and demonstrate the utility of a set of alternative metrics to assess quality and productivity. We also suggest new measures of efficiency and quality that hold promise to tell us more than response rates ever could.

CONCURRENT SESSION J-3

INNOVATIVE APPROACHES FOR SAMPLING SPECIAL POPULATIONS

A New Source of Local Health Data: Facebook Likes

Steven Gittelman (Marketing, Inc.) and Elaine Trimarchi (Marketing, Inc.)

Diabetes, cancer, obesity and heart attacks have reached epidemic proportions. To deal with these health threats, we need strategies that deal with the behaviors that generate them. While morbidity and mortality data are generally available at the local level, behavioral measures are not.

Data from the BRFSS and other similar survey efforts is not available at the county level in sufficient quantities in over 90 percent of American counties. The median sample size by county is 66. Aggregating data over time does not supply sufficient granularity and buries trends. Other national data (i.e. countylevelrankings.org) are model based and/or drawn from already thin federal data.

We need new sources of local data, and the Big Data revolution may provide a partial answer. Social Networks, such as Facebook, have expanded to include over half of the US population, allowing for interesting data on respondent lifestyle in virtually every area of the country. These data are not explicitly health-related, but statistical analysis shows that when taken together, the 'network' of an individual's likes are predictive of many types of health behaviors and outcomes, regardless of the source.

Patterns of Facebook *Likes* on their own can predict health perceptions at a county level (as measured by the BRFSS) with an R^2 of .5, and can predict Life Expectancy at an R^2 of .75. In areas where traditional survey data are difficult, if not impossible to collect, the distribution of *Likes* may serve as a useful indicator.

We shall demonstrate how, on the basis of forty summary variables drawn from Facebook *Likes,* we create a predictive model and display not only predictability, but a categorization of health conditions that could allow for actionable county classifications that could inform community level health initiatives.

A New Online Sample Frame Concept: Passive Data Collection

Elaine Trimarchi (Marketing, Inc.), Steven Gittelman (Marketing, Inc.), and Philip Garland (Survey Monkey)

Online samples are biased by the non-coverage of those offline. Commercial online samples consist of double opt-in respondents who are often financially incented and thus have been dubbed "professional respondents". They participate in many online questionnaires and are thought to have behavioral biases even when demographically balanced.

In health research we seek the ability to collect data on a local basis. The costs associated with telephone research are prohibitive for small populations. In addition, telephone has its own sources of measurement error that appear to exist in sensitive health questions. The online panels do not have the feasibility nor do they properly represent the populations that we seek to study. Since the health of one's local community may have personal relevance to respondents, we suggest that community spirit may be a better driver of survey completion. Here we experiment with one method of conducting a survey through passive means with no financial incentives.

Survey Monkey provides a unique sampling environment. Researchers are often community based, conducting surveys with local populations who have an interest in the survey data. The local church, community center, school district or boy scout troop all reach out to their members in an effort to collect data of use to the group. For two weeks we collected data on BRFSS questions plus demographics from any respondent completing a study in Survey Monkey within the State of Georgia. We used no incentives and simply asked residents to help us collect health data on their community. We balanced the sample of over 10,000 respondents using post-stratification weights. A similar sample was collected from an online commercial sample, quota controlled and post-stratified.

The fit to telephone collected BRFSS reference points was quite good and equal or better than similar data collected from commercial panels.

The Use of Non-Probability Samples to Characterize Rare Conditions

John Boyle (ICF International)

Evidence-based medicine is a key element in current health policy. This is the use of best evidence in decisions about the care of individual patients "derived from high-quality research on population samples, to inform clinical decision-making in the diagnosis, investigation or management of individual patients."

Unfortunately, there is no high quality research using population samples for the vast majority of medical conditions outside of clinical settings. The vast majority of conditions are sufficiently uncommon that they are not represented in federal health surveys. Moreover, they also are sufficiently uncommon that the cost to conduct a survey for these individual conditions would be cost prohibitive using probability methods for interested parties, such as patient organizations.

This paper examines the potential for non-probability methods, specifically large-scale internet population panels, to describe the management and treatment of rare diseases in community settings. A national probability survey of 10,000 households was conducted for one of these rare or uncommon conditions. The probability sample provided good prevalence estimates for the condition, but yielded too small a subsample of patients to reliably assess management and treatment. To address this limitation, a large online panel was used to identify a nationally distributed, community based sample of 160 diagnostically eligible patients with this condition. The characteristics of the patient sample from the online panel matched those from the probability sample. The larger online sample, however, suggested that treatment and management of the condition was substantially different for patients seen in academic medical centers than those seen in the general community.

These findings suggests that community based assessments should be considered even for rare diseases, if we seek true evidence based medicine. They also suggest that non-probability samples, despite their many limitations, may offer a potentially valuable approach to this objective.

CONCURRENT SESSION J-4

NEW MEASUREMENT SOLUTIONS: AGREEMENT, INTRACLASS CORRELATION, AND OUTCOME PATTERNS

State Based Intraclass Correlation Values for Planning Group-Randomized Trials in Education: Within and Between District Estimates

Eric Hedberg (NORC at the University of Chicago) and Larry Hedges (Northwestern University)

Background

Cluster randomized experiments that assign intact groups such as schools or school districts to treatment conditions are increasingly common in educational research. Such experiments are inherently multilevel designs whose sensitivity (statistical power and precision of estimates) depends on the variance decomposition across levels. This variance decomposition is usually summarized by the intraclass correlation structure and, if covariates are used, the effectiveness of the covariates in explaining variation at each level of the design.

Objectives

This paper provides a compilation of school and district level intraclass correlation values of academic achievement and related covariate effectiveness based on state longitudinal data systems. These values are designed to be used for planning group-randomized experiments in education. The use of these values to compute statistical power and plan 2 and 3 level group randomized experiments is illustrated.

Research Design
We fit several hierarchical linear models to state data by grade and subject to estimate intraclass correlations and covariate effectiveness. We then compare our average of state estimates with the national work by Hedges and Hedberg (2007).

Agree or Disagree? A Demonstration of An Alternative Statistic to Cohen's Kappa for Measuring the Extent and Reliability of Agreement Between Observers

Qingshu Xie (MacroSys, LLC)

Cohen's Kappa has been widely used to measure the extent and reliability of agreement among observers. It is readily available in standard statistical software packages such as SAS and Stata, which further helps it to be conveniently used by researchers and practitioners. However, it has been recognized that Cohen's Kappa is inherently flawed in its assumption of complete randomness for chance-agreement correction and in its lack of accounting for bias and prevalence. There are often situations in which extremely low kappa coefficients are obtained when the level of agreement between observers are actually very high. Such obvious inconsistency leads to a search for a better alternative summary statistic for measuring the extent and reliability of agreement. One of the recently developed alternatives is Gwet's AC1. It is not yet readily available in standard software packages because it is relatively new. In this paper, we will compare Cohen's Kappa and Gwet's AC1, using a subset of the data from a recent study we performed to compute the two agreement coefficients. We will explore whether Gwet's AC1 is superior to Cohen's Kappa in measuring the extent and reliability of agreement among observers in the presence of bias and prevalence.

Statistical Tests of Agreement

Elizabeth Stanwyck (University of Maryland, Baltimore County), Bimal Sinha (University of Maryland, Baltimore County), and Barry Nussbaum (U.S. Environmental Protection Agency)

Assessing agreement between two methods of measurement plays a vital role in deciding if one of the methods (newer or cheaper) can be adopted in future experiments. Statistical tests assessing agreement between two comparable measurement methods are typically based on paired data (x, y) collected from two measurement methods under comparison from several units of the population. Assuming a bivariate normal distribution for the responses from the two methods, we derive a likelihood ratio test for a combined hypothesis of equality of means, equality of variances, and a known high value of pairwise correlation based on *unbalanced* paired data of the type (x, y1, y2, . . . ym). This situation arises when one observes multiple replications of one response for a specified single value of the other response from sampled units, corresponding to the situation when one measurement method (x) is much more expensive than the second method (having multiple observations). Our results provide a generalization of Yimprayoon, Tiensuwan, and Sinha (2006). We also explore a combined P-value approach to the problem, providing a sequential as well as simultaneous test. We illustrate our methods with an application from USEPA and simulations.

Measuring Health and Healthcare Disparities

James Scanlan (James P. Scanlan, Attorney at Law)

Researchers employ a number of methods to measure health (including healthcare) disparities. The most common of these are relative differences in either a favorable outcome or the corresponding adverse outcome, absolute differences between rates, and odds ratio. Research relying on these measure has been problematic, however, because of a failure to recognize the way each measure tends to be affected by the prevalence of an outcome. The rarer an outcome the greater tends to be the relative difference in experiencing it and the smaller tends to be the relative difference in avoiding it. Thus, as mortality generally declines relative differences in mortality tend to increase while relative differences in survival tend to decrease; as rates of appropriate healthcare generally increase, relative differences in rates of failing to receive appropriate care tend to increase while relative differences in receipt of appropriate care tend to decrease. Absolute differences and odds ratios tend also to be affected by the prevalence of an outcome though in a more complicated way than the two relative differences. Broadly, as uncommon

outcomes become more common absolute differences tend to increase; as already common outcome outcomes become even more common absolute differences tend to decrease. Further, as the prevalence of an outcome changes, absolute differences tend to change in the same direction as the smaller relative difference. Differences measured by odds ratios tend to change in the opposite direction of absolute differences. Although these patterns are apparent in a wide range of data, they are little known among persons attempting to determine whether health disparities are increasing or decreasing. This presentation will illustrate these patterns and describe a method of appraising differences in the circumstances of two group reflected by a pair of outcome rates that is unaffected by the prevalence of the outcome.

Author Index

FCSM RESEARCH CONFERENCE AUTHOR INDEX

Name	Session
John Abowd	C-3, E-3
Denise Abreu	I-4
Tamara Adams	I-4
Kofi Placid Adragni	G-2
Daniel Adrian	I-4
Luis Amador-Jiménez	B-2
Md Shohel Reza Amin	B-2
Irwin Anolik	D-4
Stephen Ansolabehere	B-4
Bettina Aten	B-3
Kisha Bailly	F-3
Marilyn Balogh	D-1
Frances Barlas	A-3, C-4, G-4, J-2
Robert Baskin	C-2
Nancy Bates	H-1
Danielle Battle	A-1
Kathy Batts	D-2
Mark Bauder	A-2
Michael Beaghen	D-1
Tanner Beam	D-4
Laura Bechtel	F-4
Adam Bee	H-3
Nikki Bellamy	A-3
Vladislav Beresovsky	A-2
Jennifer Romano Bergstrom	A-3
Marcus Berzofsky	D-1
Renuka Bhaskar	A-4
William Block	E-3
Stephen Blumberg	I-3
Alex Bohl	G-3
Daniel Bonnéry	G-2
Jonaki Bose	A-1, C-2, D-2, G-4
Michel Boudreaux	A-4
John Boyle	J-3
Matthew Brault	H-3
J. David Brown	E-2
Evan Burton	D-4
Stephen Busselberg	G-2
Kathleen Call	A-4
Aaron Cantu	H-3
Jesús Cantú-Escalante	G-2

Name	Session
Olivia Carrillo-Gamboa	G-2
Sarah Hastedt Carroll	F-3
Ricardo Carvalho	C-1
David Cella	B-4
Wan-Ying Chang	B-1
Sixia Chen	D-1
Yang Cheng	D-2, G-2
Frances Chevarley	J-2
Sadeq Chowdhury	F-1
Lisa Colpe	C-2
Kennon Copeland	I-3
Lance Couzens	D-1
Larry Cox	I-2
Carol Gotway Crawford	B-4
Alisha Creel	G-1
Arthur Cresce Jr.	D-2
Doug Currivan	G-4
John Czajka	C-2
Ashleigh Davis	A-3, C-4
Karen Davis	J-2
Hannah Day	H-2, J-1
James Dayton	I-1
Kevin Deardorff	B-1
Katie Dekker	C-1
Angela Delano	D-2
Daniel Dench	H-1
Satvinder Dhingra	B-4
Mamadou Diallo	A-2
Charles DiSogra	F-4
John Dixon	D-3
Sylvia Dohrmann	C-1
Gia Donnalley	I-4
Alan Dotts	I-1
Heather Driscoll	I-1
Paul Drugan	F-2
Piper DuBray	F-3
Brian Dumbacher	I-3
Adam Duran	D-4
Gianna Dusch	B-1, D-3
Andrew Dyer	I-1
Sandy Dyer	E-4

FCSM RESEARCH CONFERENCE AUTHOR INDEX

Name	Session
Jennifer Edgar	G-4, H-1
Michael Elliott	B-2
John Eltinge	G-1
Ned English	C-1
Chandra Erdman	B-1, H-1
Amy Falcone	A-3, C-4
Robert Fay	A-2
Leticia Fernandez	A-4
Eric Fink	D-3
David Finkelhor	F-4
Kate Flint	H-4
Naomi Freedner	F-3
Steven Frenk	H-2
Scott Fricker	E-4
Brett Fried	A-4
Robert Furberg	H-1
Nadarajasundaram Ganesh	I-3
Turkan Gardenier	G-2
Philip Garland	J-3
Graton Gathright	E-2
Michael Gerling	I-1
Joseph Gfroerer	C-2, D-2
Bonnie Ghosh-Dastidar	E-1
Kinsey Gimbel	F-2
Steven Gittelman	C-4, J-3
Elizabeth Goldberg	C-3
Jeff Gonder	D-4
Jeffrey Gonzalez	I-3
Brian Griepentrog	F-2
Christopher Grieves	D-2
Deborah Griffin	J-2
Richard Griffin	E-4
Wendy Gross	G-4
Heather Hammer	F-4
Daifeng Han	E-1
Benjamin Harris	E-2
Yulei He	J-1
Kristie Healey	F-3
Eric Hedberg	J-4
Sarra Hedden	C-2
Larry Hedges	J-4
Rosa Isela Hernández-Zamora	G-2
Pascal Heus	D-3
Elizabeth Hoeffel	D-3
Carma Hogue	I-3
David Hornick	H-3
Janey Hsiao	A-2
Arthur Hughes	A-1
Ronaldo Iachan	H-4
Howard Iams	E-2
Emily Isenberg	H-2
Jessica Jagger	A-3, C-4
James Johanson	C-3
Andrea Johnson	E-4
David Jones	G-3
Maggie Jones	E-2
Dean Judson	J-1
Claude Julien	E-4
Bob Kaplan	B-4
Stephen Kaputa	F-4
Alan Karr	I-2
Michael Keating	H-1
Annice Kim	H-1
Hang Joon Kim	I-2
Ryan King	I-1
Stas Kolenikov	F-4
Scott Konicki	I-4
Brandon Kopp	E-4
Michael Kornbau	F-3
Phillip Kott	C-2, F-1
Jonathan Krentel	E-4
Tom Krenzke	D-1, E-3
Larry Kroutil	G-4
Annemarie Kuhns	B-3
Tracey Kyckelhahn	F-2
Carl Lagoze	E-3
Partha Lahiri	G-2
Andrea Lamas	I-4
Liana Christin Landivar	H-2
Victor Lange	C-4
Lynn Langton	F-1
Michael Larsen	A-2, F-1

FCSM RESEARCH CONFERENCE AUTHOR INDEX

Name	Session
Linda Lawson	I-1
Mary Layne	J-1
Mark Leach	F-2
David Lee	H-3
Katie Lewis	H-4
Jianzhu Li	E-3
Ping Li	C-3
Wendy Li	F-2
Dan Liao	C-2, F-1
Allie Lieberman	H-1
Joanna Fane Lineback	D-3
Jeffrey Little	C-3
Elizabeth Love	E-1
Joseph Luchman	F-2
Donald Luery	A-2
Adela Luque	A-4, E-2
Steven Machlin	H-2
Amy Mack	A-3
Daniel Manrique-Vallier	I-2
Erica Marquette	F-3
Sean Marsh	C-1
Jared Martin	E-3
Troy Martin	B-3
Wendy Martinez	C-3
Paul Massell	E-3
Brett McBride	D-2, G-4
Jaki McCarthy	B-1
Joe McMichael	C-1
Erin McNulty	H-4
Cameron McPhee	A-1
Alex Measure	C-3
Carla Medalia	H-3
Esther Mezey	H-2
Lynn Milan	B-1
Eric Miller	J-1
Peter Miller	B-1
Lisa Mirel	H-2
Melissa Mitchell	B-1
Keith Evan Moffett	I-1
Jill Montaquila	E-1
Chris Moriarity	A-1
Katherine Morton	A-1
Tim Mulcahy	D-3
Vincent Thomas Mule Jr.	I-4
Elaine Murakami	E-3
Joe Murphy	H-1
Ali Mushtaq	F-2
Dawn Nelson	G-4, J-2
Constance Newman	E-2
Justin Nguyen	E-1
Amber Noel	A-1
James Noon	A-4
Kimberly Noonan	D-3
Barry Nussbaum	J-4
Eileen O'Brien	A-1
Amy O'Hara	A-4, E-2
Brett O'Hara	A-4, H-3
Colm O'Muircheartaigh	C-1
Jay Occhiogrosso	I-1
Catherine Okoro	B-4
Broderick Oliver	C-4
A. Elizabeth Ormson	I-3
Charlotte Oslund	B-3
Kathy Ott	B-1
Pamela Owens	C-2
Jia-Ern Pai	G-2
Dicy Painter	G-4
Timothy Park	B-3
Jennifer Parker	J-1
Joanne Pascale	H-3
Laura Paszkiewicz	G-4
Ryne Paulose	H-2
Michael Pemberton	D-2
Sandy Peterson	D-1
Jacqueline Pflieger	C-4
Benjamin Phillips	F-4
Jean Pignal	G-1
Michael Planty	A-2, F-1
Chester Ponikowski	H-4
Sander Post	H-2
Dmitriy Poznyak	G-3
Patrick Purcell	E-2

FCSM RESEARCH CONFERENCE AUTHOR INDEX

Name	Session
Siyu Qing	F-1
Susan Queen	C-2
Bhojinarine Rambharat	B-2
Magdalena Ramos	I-4
Dylan Rassier	B-3
Sonya Rastogi	A-4
Michael Ratcliffe	E-4
Laurie Reedman	E-4
Jerome Reiter	I-2
Trudi Renwick	B-3
Brad Rhein	H-4
Jamie Ridenhour	C-1
William Riley	B-4
Douglas Rivers	B-4
William Robb	H-4
Barbara Robles	H-4
Jonathan Rodean	H-3
Jessica Ross	G-3
Cynthia Rothhaas	J-1
Don Royce	G-1
Keith Rust	E-1
Adam Safir	G-4
Narayan Sastry	E-1
James Scanlan	J-4
Ashley Schaad	A-3, C-4
Erik Scherpf	E-2
Fritz Scheuren	F-2
Eric Schone	G-3
Alisú Schoua-Glusberg	G-4
Jessica Semega	H-3
Bonnie Shook-Sa	A-1, C-1, D-1
Richard Sigman	C-1
Avinash Singh	B-2
Bimal Sinha	J-4
Eric Slud	B-1
Jessica Smith	H-3
Peter Smith	D-4
Ted Socha	B-1
Rachel Soloveichik	D-2
Younghwan Song	I-3
Christopher Sroka	C-2
Samantha Stalley	G-3
Elizabeth Stanwyck	J-4
Philip Steel	E-3
Amy Steinweg	H-3
Robert Stillwell	C-3
Alex Strashny	D-2
Jonathan Strohl	A-3
Amang Sukasih	C-2
Paul Sullivan	D-2
Michail Sverchkov	J-2
Sam Szelepka	A-2
Ronald Szoc	C-4, G-1
Lucilla Tan	E-4
Andrew Therriault	F-2
Randall Thomas	C-4, G-4
Katherine Jenny Thompson	C-4, F-4
Nhien To	D-2
Roger Tourangeau	E-4
Bac Tran	E-1
Julie Trépanier	G-1
Elaine Trimarchi	C-4, J-3
Deborah Trunzo	D-2
Heather Turner	F-4
Joanna Turner	A-4
Sidney Carl Turner	F-2
Richard Valliant	E-1
Andrew Vardeman	I-1
Luciano Viera, Jr.	C-1
Lars Vilhuber	E-3
Ana Villar	G-4
Pieter Vlag	F-2
Richard Volpe	B-3
Gina Walejko	B-1
Deborah Wagner	J-1
Kevin Wang	G-4
Shu Wang	I-4
Penelope Weinberger	E-3
Karen Wessels	J-2
Daniel Whitehead	F-4
Adeline Wilcox	C-3

FCSM RESEARCH CONFERENCE AUTHOR INDEX

Name	Session
Jeremy Williams	E-3
Matthew Williams	C-2
Rick Williams	D-1
Scott Williams	I-1
Eric Wilson	I-1
William Winkler	J-1
Eric Wood	D-4
Rose Woods	F-3
Laura Wronski	F-3
Xi Xia	B-2
Qingshu Xie	J-4
Daniel Yang	I-3
Frank Yoon	G-3
Linda Young	I-4
Ellen Zapata	A-4
Alireza Zareie	B-2
Andrew Zbikowski	D-2
Joe Zickafoose	G-3
Andrew Zukerberg	F-3

Reports Available in the Statistical Policy Working Paper Series

1. Report on Statistics for Allocation of Funds, 1978 (NTIS PB86-211521/AS)
2. Report on Statistical Disclosure and Disclosure-Avoidance Techniques, 1978 (NTIS PB86-211539/AS)
3. An Error Profile: Employment as Measured by the Current Population Survey, 1978 (NTIS PB86-214269/AS)
4. Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistics, 1978 (NTIS PB86-211547/AS)
5. Report on Exact and Statistical Matching Techniques, 1980 (NTIS PB86-215829/AS)
6. Report on Statistical Uses of Administrative Records, 1980 (NTIS PB86-214285/AS)
7. An Interagency Review of Time-Series Revision Policies, 1982 (NTIS PB86-232451/AS)
8. Statistical Interagency Agreements, 1982 (NTIS PB86-230570/AS)
9. Contracting for Surveys, 1983 (NTIS PB83-233148)
10. Approaches to Developing Questionnaires, 1983 (NTIS PB84-105055)
11. A Review of Industry Coding Systems, 1984 (NTIS PB84-135276)
12. The Role of Telephone Data Collection in Federal Statistics, 1984 (NTIS PB85-105971)
13. Federal Longitudinal Surveys, 1986 (NTIS PB86-139730)
14. Workshop on Statistical Uses of Microcomputers in Federal Agencies, 1987 (NTIS PB87-166393)
15. Quality in Establishment Surveys, 1988 (NTIS PB88-232921)
16. A Comparative Study of Reporting Units in Selected Employer Data Systems, 1990 (NTIS PB90-205238)
17. Survey Coverage, 1990 (NTIS PB90-205246)
18. Data Editing in Federal Statistical Agencies, 1990 (NTIS PB90-205253)
19. Computer Assisted Survey Information Collection, 1990 (NTIS PB90-205261)
20. Seminar on Quality of Federal Data, 1991 (NTIS PB91-142414)
21. Indirect Estimators in Federal Programs, 1993 (NTIS PB93-209294)
22. Report on Statistical Disclosure Limitation Methodology, 1994 (NTIS PB94-165305)
23. Seminar on New Directions in Statistical Methodology, 1995 (NTIS PB95-182978)
24. Electronic Dissemination of Statistical Data, 1995 (NTIS PB96-121629)
25. Data Editing Workshop and Exposition, 1996 (NTIS PB97-104624)
26. Seminar on Statistical Methodology in the Public Service, 1997 (NTIS PB97-162580)
27. Training for the Future: Addressing Tomorrow=s Survey Tasks, 1998 (NTIS PB99-102576)
28. Seminar on Interagency Coordination and Cooperation, 1999 (NTIS PB99-132029)
29. Federal Committee on Statistical Methodology Research Conference (Conference Papers), 1999 (NTIS PB99-166795)
30. 1999 Federal Committee on Statistical Methodology Research Conference: Complete Proceedings, 2000 (NTIS PB2000-105886)
31. Measuring and Reporting Sources of Error in Surveys, 2001 (NTIS PB2001-104329)
32. Seminar on Integrating Federal Statistical Information and Processes, 2001 (NTIS PB2001-104626)
33. Seminar on the Funding Opportunity in Survey Research, 2001 (NTIS PB2001-108851)
34. Federal Committee on Statistical Methodology Research Conference (Conference Papers), 2001 (NTIS PB2002-100103)
35. Seminar on Challenges to the Federal Statistical System in Fostering Access to Statistics, 2002
36. Seminar on the Funding Opportunity in Survey Research, 2003
37. Federal Committee on Statistical Methodology Research Conference (Conference Papers), 2003
38. Summary Report of the FCSM-GSS Workshop on Web-Based Data Collection, 2004
39. Federal Committee on Statistical Methodology Research Conference (Conference Papers), 2005
40. Federal Committee on Statistical Methodology Research Conference (Conference Papers), 2007
41. Federal Committee on Statistical Methodology Research Conference (Conference Papers), 2009
42. Federal Committee on Statistical Methodology Research Conference (Conference Papers), 2012

Copies of these working papers may be ordered from NTIS Document Sales, 5285 Port Royal Road, Springfield, VA 22161; Phone: 1-800-553-6847. The Statistical Policy Working Paper series is also available electronically from FCSM's Web site www.fcsm.gov.

NOTES

NOTES

NOTES

The Federal Committee on Statistical Methodology and the Council of Professional Associations on Federal Statistics recognize and are abundantly grateful to the following organizations for sponsoring this year's FCSM Research Conference. Their commitment and support help to ensure the exchange of cutting edge statistical methods and findings among statisticians and other social scientists from the Federal government, academia, and the business community.

Special recognition to Nielsen for sponsoring a student to attend the Research Conference