FCSM Research Conference Planning Committee

Nancy Bates, Co-Chair, U.S. Census Bureau

Dawn Haines, Co-Chair, U.S. Census Bureau

Jeff Beaubier, Environmental Protection Agency

Benjamin Bridgman, Bureau of Economic Analysis

Jock Black, National Science Foundation

Thomas Broene, Energy Information Administration

Kevin Cecco, Statistics of Income Division, IRS

Amrut Champaneri, Bureau of Transportation Statistics

Chris Chapman, National Center for Education Statistics

Thesia Garner, Bureau of Labor Statistics

Charlie Hallahan, Economic Research Service, USDA

Paige Harrison, Bureau of Justice Statistics

Anna Holaus, U.S. Census Bureau

Kristen Hughes, Bureau of Justice Statistics

Howard Iams, Office of Research, Evaluation, and Statistics, SSA

Alesha Lewis, U.S. Census Bureau

Pamela McGovern, U.S. Census Bureau

Bill Mockovak, Bureau of Labor Statistics

Jennifer Parker, National Center for Health Statistics

Roberta Pense, National Agricultural Statistics Service

Michael Planty, National Center for Education Statistics

Edward Spar, Council of Professional Associations on Federal Statistics

The Federal Committee on Statistical Methodology Members

(January 2007)

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

Nancy Bates (Secretary), U.S. Census Bureau

Lynda Carlson, National Science Foundation

Steven B. Cohen, Agency for Healthcare Research and Quality

Lawrence H. Cox, National Center for Health Statistics

John Eltinge, Bureau of Labor Statistics

Robert E. Fay, U.S. Census Bureau

Dennis Fixler, Bureau of Economic Analysis

Larry Graubard, National Cancer Institute

William Iwig, National Agricultural Statistics Service

Arthur Kennickell, Federal Reserve Board

Nancy Kirkendall, Energy Information Administration

Jennifer Madans, National Center for Health Statistics

Renee Miller, Energy Information Administration

Susan Schechter, U.S. Census Bureau

Rolf Schmitt, Federal Highway Administration

Marilyn McMillen Seastrom, National Center for Education Statistics

Stephanie Shipp, National Institute of Standards and Technology

Monroe Sirken, National Center for Health Statistics

Nancy Spruill, Department of Defense

Philip Steel, U.S. Census Bureau

Clyde Tucker, Bureau of Labor Statistics

Katherine K. Wallman, (Champion) Office of Management and Budget

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

Additional Conference Support

Lourdes Hartman, U.S. Census Bureau

LoWanda Rivers, U.S. Census Bureau

Lee Ann Sklar, Council of Professional Associations on Federal Statistics

Note: Papers and discussant comments will be available in early 2008 on <www.fcsm.gov>.

Final Program

Federal Committee on Statistical Methodology Research Conference

Arlington, Virginia—November 5–7, 2007

Monday (11/5)

7:30 a.m.
Registration
(Ballroom Foyer)

Coffee
(Ballroom Foyer)

9:00–10:00 a.m.
Welcoming Remarks and PLENARY SESSION I
(Grand Ballroom)

10:00–10:30 a.m.
Break
(Ballroom Foyer)

10:30 a.m.–12 noon
CONCURRENT SESSION
II-A: (Ballroom A) **II-B:** (Ballroom B) **II-C:** (Ballroom C)

12 noon–1:15 p.m.
Open

1:30–3:00 p.m.
CONCURRENT SESSION
III-A: (Ballroom A) **III-B:** (Ballroom B) **III-C:** (Ballroom C)

3:00–3:30 p.m.
Break
(Ballroom Foyer)

3:30–5:00 p.m.
CONCURRENT SESSION
IV-A: (Ballroom A) **IV-B:** (Ballroom B) **IV-C:** (Ballroom C)

Tuesday (11/6)

7:30 a.m.
Registration
(Ballroom Foyer)

Coffee
(Ballroom Foyer)

9:00 a.m.–12:30 p.m.
Technical Demonstrations
(Lobby Atrium)

9:00–10:30 a.m.
CONCURRENT SESSION
V-A: (Ballroom A) **V-B:** (Ballroom B) **V-C:** (Ballroom C)

10:30–11:00 a.m.
Break
(Ballroom Foyer)

11:00 a.m.–12:30 p.m.
CONCURRENT SESSION
VI-A: (Ballroom A) **VI-B:** (Ballroom B) **VI-C:** (Ballroom C)

12:30–1:45 p.m.
Open

2:00–3:30 p.m.
CONCURRENT SESSION
VII-A: (Ballroom A) **VII-B:** (Ballroom B) **VII-C:** (Ballroom C)

3:30–4:00 p.m.
Break
(Ballroom Foyer)

4:00–5:30 p.m.
CONCURRENT SESSION
VIII-A: (Ballroom A) **VIII-B:** (Ballrooms B–C)

Wednesday (11/7)

7:30 a.m.
Registration
(Ballroom Foyer)

Coffee
(Ballroom Foyer)

9:00–10:30 a.m.
CONCURRENT SESSION
IX-A: (Ballroom A) **IX-B:** (Ballroom B) **IX-C** (Ballroom C)

10:30–11:00 a.m.
Break
(Ballroom Foyer)

11:00 a.m.–12:30 p.m.
CONCURRENT SESSION
X-A: (Ballroom A) **X-B:** (Ballroom B) **X-C:** (Ballroom C)

Meeting Rooms:

Ballroom A–C
(2nd Floor)

Grand Ballroom Foyer
(2nd Floor)

Lobby Atrium
(2nd Floor)

Final Program[1]

Monday, November 5

7:30 a.m.–5:00 p.m. **Ballroom Foyer**
Registration

7:30–9:00 a.m. **Ballroom Foyer**
Coffee

9:00–9:10 a.m. **Grand Ballroom**
Introduction and Welcoming Remarks

9:10–10:00 a.m.

PLENARY SESSION I

New Insights Into Questionnaire Design: How to Maximize the Validity of Your Measurements
Jon Krosnick (Stanford University, USA)

10:00–10:30 a.m **Ballroom Foyer**
Break

10:30 a.m.–12 noon **Ballroom A**

CONCURRENT SESSION II-A:
MONITORING, MEASURING, AND ADJUSTING FOR NONRESPONSE I

Chair: Nancy Bates (U.S. Census Bureau)

Monitoring Response to a Multi-Wave Medical Establishment Survey: How Different Are Responders From Nonresponders?
Jessica Graber (National Opinion Research Center, USA)

Factors Affecting Response to the Occupational Employment Statistics Survey
Polly Phipps (Bureau of Labor Statistics, USA)
Carrie Jones (Bureau of Labor Statistics, USA)
Clyde Tucker (Bureau of Labor Statistics, USA)

Review of the Weighting Methodology for the Canadian Community Health Survey
Cathlin Sarafin (Statistics Canada)
Steven Thomas (Statistics Canada)
Michelle Simard (Statistics Canada)

An Examination of Nonresponse Error and Measurement Error by Level of Effort Using Frame Information, Survey Reports, and Paradata
Sarah Dipko (Westat, USA)
Kerry Levin (Westat, USA)
Mary-Helen Risler (Internal Revenue Service, USA)

Session Organizer: Nancy Bates (U.S. Census Bureau)

[1] In the case of coauthors, the presenter is underlined.

10:30 a.m.–12 noon **Ballroom B**

CONCURRENT SESSION II-B:
TIME SERIES

Chair: Charlie Hallahan (Economic Research Service, USA)

Comparison of Methods for Computing Yearly Growth Rates From Weekly and Monthly Data, 1978 to 2005
Carol Blumberg (Energy Information Administration, USA)

The X-13A-S Seasonal Adjustment Program
Brian Monsell (U.S. Census Bureau)

Coherent Trends, Turning Points, and Forecasts for American Community Survey Data
Tucker McElroy (U.S. Census Bureau)

Empirical Evaluation of X-11 and Model-Based Seasonal Adjustment Methods
Richard Tiller (Bureau of Labor Statistics, USA)
Stuart Scott (Bureau of Labor Statistics, USA)
Dan Chow (Bureau of Labor Statistics, USA)

Session Organizer: Charlie Hallahan (Economic Research Service, USA)

10:30 a.m.–12 noon **Ballroom C**

CONCURRENT SESSION II-C:
FRAMES AND COVERAGE

Chair: Thomas Broene (Energy Information Administration, USA)

Practicability of Including Cell Phone Numbers in Random Digit Dialed Surveys: Pilot Study Results From the Behavioral Risk Factor Surveillance System
Michael Link (Nielsen Media Research, USA)
Michael Battaglia (Abt Associates, Inc., USA)
Larry Osborn (Abt Associates, Inc., USA)
Martin Frankel (Baruch College, CUNY and Abt Associates, Inc., USA)
Ali Mokdad (Centers for Disease Control and Prevention, USA)

Measuring and Adjusting for Frame Undercoverage of the State and Local Value Put-In-Place (VIP) Survey
Thuy Trang Nguyen (U.S. Census Bureau)
Shadana Myers (U.S. Census Bureau)

Comparing the Quality of the Master Address File and the Current Demographic Household Surveys' Multiple Frames
Xijian Liu (U.S. Census Bureau)

Impact of Preliminary Versus Final Economic Census Data on the Universe Extraction Process for Current Business Surveys
Kari Clark (U.S. Census Bureau)
Carol King (U.S. Census Bureau)

Session Organizer: Thomas Broene (Energy Information Administration, USA)

---

12 noon–1:15 p.m.
Open Lunch

1:30–3:00 p.m. **Ballroom A**

CONCURRENT SESSION III-A:
DISCLOSURE I

Chair: Michael Planty (National Center for Education Statistics, USA)

Comparative Evaluation of Four Different Sensitive Tabular Data Protection Methods Using a Real Life Table Structure of Complex Hierarchies and Links
Ramesh Dandekar (Energy Information Administration, USA)

Easy to Use Is Putting The Cart Before the Horse: Effective Techniques for Masking Numerical Data
Krish Muralidhar (University of Kentucky, USA)
Rathindra Sarathy (Oklahoma State University, USA)

Comparing Fully and Partially Synthetic Datasets for Statistical Disclosure Control in the German IAB Establishment Panel
Joerg Drechsler (Institute for Employment Research, Germany)
Agnes Dundler (Institute for Employment Research, Germany)
Stefan Bender (Institute for Employment Research, Germany)
Susanne Raessler (Institute for Employment Research, Germany)

Discussant: Marilyn McMillen Seastrom (National Center for Education Statistics, USA)

Session Organizer: Michael Planty (National Center for Education Statistics, USA)

1:30–3:00 p.m. **Ballroom B**

CONCURRENT SESSION III-B:
ADVANCES IN DATA EDITING

Chair: Dale Atkinson (National Agricultural Statistics Service, USA)

Investigation of Selective Editing Procedures for the Annual Survey of Government Finances
Loretta McKenzie (U.S. Census Bureau)
Terri Craig (U.S. Census Bureau)
Carma Hogue (U.S. Census Bureau)

Measuring Edit Efficiency in the Economic Directorate of the U.S. Census Bureau
Broderick Oliver (U.S. Census Bureau)
Katherine Jenny Thompson (U.S. Census Bureau)

Improving the Efficiency of Data Editing and Imputation for a Large-Scale British Annual Business Survey
Alaa Al-Hamad (Office for National Statistics, United Kingdom)
Gary Brown (Office for National Statistics, United Kingdom)
Pedro Silva (University of Southampton, United Kingdom)

An Empirical Investigation Into Macro Editing on Two Economic Surveys
Katherine Jenny Thompson (U.S. Census Bureau)
Laura Ozcoskun (U.S. Census Bureau)

Session Organizer: Roberta Pense (National Agricultural Statistics Service, USA)

1:30–3:00 p.m. **Ballroom C**

CONCURRENT SESSION III-C:
EXAMINING THE EFFECTS OF MODE AND RESPONDENT CHARACTERISTICS ON SURVEY PARTICIPATION

Chair: Rachel Caspar (RTI International, USA)

Incorporating a Multi-Modality Design Into a Random-Digit-Dialing Survey
Michael Battaglia (Abt Associates Inc., USA)
Larry Osborn (Abt Associates Inc., USA)
Martin Frankel (Baruch College, CUNY and Abt Associates Inc., USA)
Michael Link (Nielsen Media Research, USA)
Ali Mokdad (Centers for Disease Control and Prevention, USA)

Interviewer-Reported Reasons for Conducting Interviews by Telephone in the National Health Interview Survey, 2005
Barbara Stussman (National Center for Health Statistics, USA)
Catherine Simile (National Center for Health Statistics, USA)
James Dahlhamer (National Center for Health Statistics, USA)

Response Profile of the 2005 American Community Survey
Geoffrey Jackson (U.S. Census Bureau)

The Influence of Selected Factors on Student Survey Participation and Mode of Completion
Tracy Hunt-White (National Center for Education Statistics, USA)

Session Organizer: Pam McGovern (U.S. Census Bureau)

3:00–3:30 p.m. **Ballroom Foyer**
Break

3:30–5:00 p.m. **Ballroom A**

CONCURRENT SESSION IV-A:
EMPLOYMENT/LABOR STATISTICS

Chair: John Ruser (Bureau of Labor Statistics, USA)

Distorted Measures of Employment in Charitable Organizations: Some Remedies
Martin David (University of Wisconsin–Madison, USA)

A Proposed Model for Microintegration of Economic and Social Data
Paul De Winden (Statistics Netherlands)
Koos Arts (Statistics Netherlands)
Martin Luppes (Statistics Netherlands)

Methodologies for Estimating Mean Wages for Occupational Employment Statistics (OES) Data
Mallika Kasturirangan (Bureau of Labor Statistics, USA)
Shail Butan (Bureau of Labor Statistics, USA)
Tamara Zimmerman (Bureau of Labor Statistics, USA)

Estimating the Measurement Error in the Current Population Survey Labor Force - A Latent Class Analysis Approach With Sample Design
Bac Tran (U.S. Census Bureau)
Justin Nguyen (U.S. Census Bureau)

Session Organizer: Benjamin Bridgman (Bureau of Economic Analysis, USA)

3:30–5:00 p.m. **Ballroom B**

CONCURRENT SESSION IV-B:
CONFIDENTIALITY/PRIVACY

Chair: Jeff Beaubier (Environmental Protection Agency, USA)

Respondent Consent to Link Survey Data With Administrative Records: Results From a Split-Ballot Field Test With the 2007 National Health Interview Survey
James Dahlhamer (National Center for Health Statistics, USA)
Christine Cox (National Center for Health Statistics, USA)

Consumer Privacy
Howard Fienberg (Council for Marketing and Opinion Research, USA)

An Introduction to the National Inmate Survey
Rachel Caspar (RTI International, USA)
Chris Krebs (RTI International, USA)
Allen Beck (Bureau of Justice Statistics, USA)
Paige Harrison (Bureau of Justice Statistics, USA)

Discussant: Alvan Zarate (National Center for Health Statistics, USA)

Session Organizer: Jeff Beaubier (Environmental Protection Agency, USA)

3:00–5:00 p.m. **Ballroom C**

CONCURRENT SESSION IV-C:
TOPICS IN ESTIMATION AND MODELING FOR NATIONAL AND INTERNATIONAL SURVEYS

Chair: Tamara Rib (Statistics of Income Division, IRS, USA)

Estimating Unemployment for Small Areas in Navarra, Spain
Maria Ugarte (Public University of Navarra, Spain)
Ana Militino (Public University of Navarra, Spain)
Tomas Goicoa (Public University of Navarra, Spain)

Two-Step Versus Simultaneous Estimation of Survey-Non-Sampling Error and True Value Components of Small Area Sample Estimators
Swamy Paravastu (Bureau of Labor Statistics, USA)
Tamara Zimmerman (Bureau of Labor Statistics, USA)
Jatinder Mehta (Temple University, USA)

Weighting and Estimation Methodology and Results From the American Community Survey Family Equalization Project
Mark Asiala (U.S. Census Bureau)

Discussant: Michael Cohen (Statistical Consultant LLC, USA)

Session Organizer: Kevin Cecco (Statistics of Income Division, IRS, USA)

Tuesday, November 6

7:30 a.m.–5:30 p.m. **Ballroom Foyer**
Registration

7:30–9:00 a.m. **Ballroom Foyer**
Coffee

9:00 a.m.–12:30 p.m. **Lobby Atrium**

TECHNICAL DEMONSTRATIONS

Chair: William Mockovak (Bureau of Labor Statistics, USA)

Making Sense of Data Via the Web - A Case Study Using Agricultural Data
Irwin Anolik (National Agricultural Statistics Service, USA)

The National Health and Nutrition Examination Survey
Yinong Chong (Centers for Disease Control and Prevention, USA)
Rosemarie Hirsch (Centers for Disease Control and Prevention, USA)
Cheryl Fryar (Centers for Disease Control and Prevention, USA)
Jennifer Dostal (Centers for Disease Control and Prevention, USA)

Computer Audio-Recorded Interviewing (CARI)
Katherine Mason (RTI International, USA)

Using a Pen Based Windows XP Tablet PC for Data Collection: Development of a Mobile System for Health Care Settings
Sarah Kalsbeek (RTI International, USA)
Dick Paddock (RTI International, USA)
Reginald Pendergraph (RTI International, USA)
Helen Smith (RTI International, USA)
Vanessa Thornburg (RTI International, USA)

Demonstration of the Hand-Held Computer to Be Used for the 2010 Census
Karen Field (U.S. Census Bureau)

Use of Global Positioning Receivers at the National Agricultural Statistics Service
Michael Gerling (National Agricultural Statistics Service, USA)

Development and Evaluation of an Audio Computer-Assisted Self-Interviewing System for Handheld Computing Devices
Kevin Wilson, (RTI International, USA)
Stephen Litavecz (RTI International, USA)
Norman Goco (RTI International, USA)

Large Scale Applied Time Series Analysis With Program TSW (TRAMO-SEATS for Windows)
Agustin Maravall (Bank of Spain, Spain)

Technical Demonstration Organizer: William Mockovak (Bureau of Labor Statistics, USA)

9:00–10:30 a.m. **Ballroom A**

CONCURRENT SESSION V-A:
DISAGGREGATION OF ECONOMIC STATISTICS

Chair: Jeri Mulrow (National Science Foundation, USA)

Implementing a Reconciliation and Balancing Model in the U.S. Industry Accounts
Dylan Rassier (Bureau of Economic Analysis, USA)
Thomas Howells, III (Bureau of Economic Analysis, USA)
Edward Morgan (Bureau of Economic Analysis, USA)
Nicholas Empey (Bureau of Economic Analysis, USA)
Conrad Roesch (Bureau of Economic Analysis, USA)

Estimating State Price Levels Using the Consumer Price Index
Bettina Aten (Bureau of Economic Analysis, USA)

Converting Historical Industry Time Series Data From SIC to NAICS
Robert Yuskavage (Bureau of Economic Analysis, USA)

An Empirical Comparison of Methods for Temporal Disaggregation at the National Accounts
Baoline Chen (Bureau of Economic Analysis, USA)

Session Organizer: Jock Black (National Science Foundation, USA)

9:00–10:30 a.m. **Ballroom B**

CONCURRENT SESSION V-B:
VARIANCE ESTIMATION I

Chair: Charlie Hallahan (Economic Research Service, USA)

A New Application of Estimating Functions for Variance and Interval Estimation From Simple and Complex Surveys
Avinash Singh (Statistics Canada)

Weight Trimming Via Bayesian Variable Selection Method
Michael Elliott (University of Michigan, USA)

Using Markov Chain Monte Carlo for Modeling Correct Enumeration and Match Rate Variability
Andrew Keller (U.S. Census Bureau)

A Study of Basic Calibration Estimators and Their Variance Estimators in Presence of Nonresponse
Yves Thibaudeau (U.S. Census Bureau)
Jun Shao (University of Wisconsin–Madison, USA)
Jeri Mulrow (National Science Foundation, USA)

Session Organizer: Charlie Hallahan (Economic Research Service, USA)

9:00–10:30 a.m. **Ballroom C**

CONCURRENT SESSION V-C:
ATTRITION

Chair: Michael Rand (Bureau of Justice Statistics, USA)

Evaluation of Models for Longitudinal Attrition Nonresponse
Eric Slud (U.S. Census Bureau)
Leroy Bailey (U.S. Census Bureau)

The Effect of Attrition on the NLSY97
Alison Aughinbaugh (Bureau of Labor Statistics, USA)
Rosella Gardecki (The Ohio State University Center for Human Resource Research, USA)

First Cut Is the Deepest
James Halse (Department for Education and Skills, United Kingdom)
Iain Noble (Department for Education and Skills, United Kingdom)
Andrew Ledger (Department for Education and Skills, United Kingdom)

Attrition Bias in Panel Estimates of the Characteristics of Program Beneficiaries
John Czajka (Mathematica Policy Research, Inc., USA)
James Mabli (Mathematica Policy Research, Inc., USA)
Karen Cunnyngham (Mathematica Policy Research, Inc., USA)

Session Organizer: Paige Harrison (Bureau of Justice Statistics, USA)

10:30–11:00 a.m. **Ballroom Foyer**
Break

11:00 a.m.–12:30 p.m. **Ballroom A**

CONCURRENT SESSION VI-A:
AMERICAN COMMUNITY SURVEY

Chair: Wendy Hicks (Westat, USA)

Improving the Labor Force Questions in the American Community Survey: The Results of the 2006 ACS Content Test
David Raglin (U.S. Census Bureau)
Kelly Holder (U.S. Census Bureau)

Analysis of Changes to the Educational Attainment Question in the 2006 ACS Content Test
Alan Peterson (U.S. Census Bureau)
Sarah Crissey (U.S. Census Bureau)

A Comparison of Forced-Choice and Mark-All-That-Apply Formats for Gathering Information on Health Insurance in the 2006 American Community Survey Content Test
Leah Ericson (Carnegie Mellon University, USA)
Chuck Nelson (U.S. Census Bureau)

A Comparison of Closed- and Open-Ended Question Formats for Select Housing Characteristics in the 2006 American Community Survey Content Test
John Chesnut (U.S. Census Bureau)
Ellen Wilson (U.S. Census Bureau)
Jeanne Woodward (U.S. Census Bureau)

Session Organizer: Chris Chapman (National Center for Education Statistics, USA)

11:00 a.m.–12:30 p.m. **Ballroom B**

CONCURRENT SESSION VI-B:
MONITORING, MEASURING, AND ADJUSTING FOR NONRESPONSE II

Chair: Jennifer Parker (National Center for Health Statistics, USA)

Using the Multi-Level Integrated Database Approach
Tom W. Smith (National Opinion Research Center, USA)

Nonresponse Bias Patterns in the Current Population Survey
John Dixon (Bureau of Labor Statistics, USA)

Sample Maintenance: Internet Use by a Low-Income Population
Bryan Rhodes (RTI International, USA)
Ellen Marks (RTI International, USA)
Jun Liu (RTI International, USA)

Discussant: Steve Miller (Bureau of Labor Statistics, USA)

Session Organizer: Jennifer Parker (National Center for Health Statistics, USA)

11:00 a.m.–12:30 p.m. **Ballroom C**

CONCURRENT SESSION VI-C:
ADMINISTRATIVE RECORDS: APPLICATIONS OF DATA LINKAGES

Chair: Dawn Haines (U.S. Census Bureau)

RELAIS: Don't Get Lost in a Record Linkage Project
Tiziana Tuoto (Italian National Institute of Statistics (ISTAT), Italy)

Nicoletta Cibella (Italian National Institute of Statistics (ISTAT), Italy)
Marco Fortini (Italian National Institute of Statistics (ISTAT), Italy)
Monica Scannapieco (Italian National Institute of Statistics (ISTAT), Italy)
Laura Tosco (Italian National Institute of Statistics (ISTAT), Italy)

Allocated Values in Linked Files
Amy O'Hara (U.S. Census Bureau)

The Use of Free School Meal Status as a Proxy for Socio-Economic Status: Evidence From Matching the Longitudinal Study of Young People in England to the National Pupil Database
James Halse (Department for Education and Skills, United Kingdom)
Andrew Ledger (Department for Education and Skills, United Kingdom)

Discussant: John Eltinge (Bureau of Labor Statistics, USA)

Session Organizer: Dawn Haines (U.S. Census Bureau)

12:30–1:45 p.m.
Open Lunch

2:00–3:30 p.m. **Ballroom A**

CONCURRENT SESSION VII-A:
MAKING MUSIC OUT OF ALL THAT NOISE: USING ADMINISTRATIVE RECORDS AND SURVEY DATA IN HARMONY

Chair: Shelly Wilkie Martinez (Office of Management and Budget, USA)

How Do Surveys Differ in Reporting the Quality of Reported Medicaid Enrollment Data: CPS and State Surveys
Michael Davern (University of Minnesota, USA)
Kathleen Call (University of Minnesota, USA)

Differences in Estimates of Public Assistance Recipiency Between Surveys and Administrative Records
Victoria Lynch (U.S. Census Bureau)
Dean Resnick (U.S. Census Bureau)
Jane Staveley (Jacob France Institute, USA)
Cynthia Taeuber (Jacob France Institute, USA)

Developing the Chapin Hall Child Care Subsidy Eligibility Model
Dean Resnick (U.S. Census Bureau)

Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Survey and SSA Administrative Data
Martha Stinson (U.S. Census Bureau)

Session Organizer: Thesia Garner (Bureau of Labor Statistics, USA)

2:00–3:30 p.m. **Ballroom B**

CONCURRENT SESSION VII-B:
ESTIMATION ISSUES

Chair: Brian Meekins (Bureau of Labor Statistics, USA)

Imbedding Model-Assisted Estimation Into ACS: The Impact on Users
Robert Fay, III (U.S. Census Bureau)

A Larger Sample Size Is Not Always Better
Nagaraj Neerchal (University of Maryland, Baltimore County, USA)
Herbert Lacayo (Environmental Protection Agency, USA)
Barry Nussbaum (Environmental Protection Agency, USA)

Alternative Tests of Independence
Jai Choi (National Center for Health Statistics, USA)
Balgobin Nandram (Worcester Polytechnic Institute, USA)

Discussant: Mary Mulry (U.S. Census Bureau)

Session Organizer: William Mockovak (Bureau of Labor Statistics, USA)

2:00–3:30 p.m. **Ballroom C**

CONCURRENT SESSION VII-C:
STATISTICAL METHODS APPLIED TO RACE DESIGNATIONS AND POPULATION ESTIMATE

Chair: Kristen Hughes (Bureau of Justice Statistics, USA)

Bridging Estimates by Race for the Current Population Survey
William Davis (National Cancer Institute, USA)
Anne Hartman (National Cancer Institute, USA)
James Gibson (Information Management Services, Inc., USA)

Statistical Methods for Analyzing Multiple Race Response Data
Tommi Gaines (University of California, USA)

Genetic Analysis of Population Structure Relative to Self-Reported Race and Ethnicity in NHANES III
Christopher Sanders (National Center for Health Statistics, USA)
Ajay Yesupriya (National Center for Health Statistics, USA)
Lester Curtain (National Center for Health Statistics, USA)

Discussant: Marilyn McMillen Seastrom (National Center for Education Statistics, USA)

Session Organizer: Kristen Hughes (Bureau of Justice Statistics, USA)

3:30–4:00 p.m. **Ballroom Foyer**
Break

4:00–5:30 p.m. **Ballroom A**

CONCURRENT SESSION VIII-A:
BRIDGING THE PAST WITH THE FUTURE: INNOVATIONS AND SURVEY MEASUREMENT ISSUES

Chair: Cynthia Z.F. Clark (Retired, Office for National Statistics, United Kingdom)

Formulating the Laws of Studying Societal Change
Tom W. Smith (National Opinion Research Center, USA)

Analytical Comparison of the SIPP and CPS-ASEC Key Longitudinal Estimates
Smanchai Sae-Ung (U.S. Census Bureau)
C. Dennis Sissel (U.S. Census Bureau)
Tracy Mattingly (U.S. Census Bureau)

UK Household Surveys: Building on Survey Integration
Roeland Beerten (Office for National Statistics, United Kingdom)

Protocol Calibration in the National Resources Inventory
Cindy Yu (Center for Survey Statistics and Methodology, USA)
Jason Legg (Center for Survey Statistics and Methodology, USA)

Session Organizer: Michael Planty (National Center for Education Statistics, USA)

4:00–5:30 p.m. **Ballrooms B–C**

CONCURRENT SESSION VIII-B:
DATA QUALITY

Chair: Howard Iams (Office of Research, Evaluation, and Statistics, SSA, USA)

Quality Assessment of the Linkage Between the Canadian Community Health Survey and Hospital Data
Michelle Rotermann (Statistics Canada)

Do Teenagers Always Tell the Truth?
Janet Rosenbaum (Harvard University, USA)

The Accuracy of Reported Insurance Status in the MEPS
Steven Hill (Agency for Healthcare Research and Quality, USA)

Quality of Income Data in Household Surveys: Lessons From a Comparative Analysis
Gabrielle Denmead (Denmead Services, USA)
John Czajka (Mathematica Policy Research, Inc., USA)
Robert Weathers (Mathematica Policy Research, Inc., USA)
Joan Turek (U.S. Department of Health and Human Services, USA)

Session Organizer: Howard Iams (Office of Research, Evaluation, and Statistics, SSA, USA)

Wednesday, November 7

7:30 a.m.–12:30 p.m. **Ballroom Foyer**
Registration

7:30–9:00 a.m. **Ballroom Foyer**
Coffee

9:00–10:30 a.m. **Ballroom A**

CONCURRENT SESSION IX-A:
CHALLENGES AND STRATEGIES IN QUESTIONNAIRE DESIGN

Chair: Daniel Kasprzyk (Mathematica Policy Research, Inc., USA)

Effects of Language and Culture on Interpretation of Translated "Confidentiality" and "Mandatory" Survey Messages
Yuling Pan (U.S. Census Bureau)
Ashley Landreth (U.S. Census Bureau)
Marjorie Hinsdale-Shouse (RTI International, USA)
Hyunjoo Park (RTI International, USA)
Alisú Schoua-Glusberg (Research Support Services, USA)

Asking for Numbers and Quantities - The Design of Answer Boxes to Frequency Questions and Its Impact on Data Quality
Marek Fuchs (University of Kassel, Germany)

From Start to Pilot: A Multi-Method Approach to the Comprehensive Redesign of an Economic Survey Questionnaire
Alfred Tuttle (U.S. Census Bureau)
Rebecca Morrison (U.S. Census Bureau)
Diane Willimack (U.S. Census Bureau)

Background and Planning for Incorporating an Event History Calendar Into the Re-Engineered SIPP
Jason Fields (U.S. Census Bureau)
Mario Callegaro (Knowledge Networks, USA)

Session Organizer: Pam McGovern (U.S. Census Bureau)

9:00–10:30 a.m. **Ballroom B**

CONCURRENT SESSION IX-B:
DISCLOSURE II

Chair: Fan Zhang (National Science Foundation, USA)

Recent Developments in the Use of Noise for Protecting Magnitude Data Tables: Balancing to Improve Data Quality and Rounding that Preserves Protection
Paul Massell (U.S. Census Bureau)
Jeremy Funk (U.S. Census Bureau)

Model Based Disclosure Avoidance for Data on Veterans
Sam Hawala (U.S. Census Bureau)
Jeremy Funk (U.S. Census Bureau)

Microdata Risk Assessment in an NSI Context
Jane Longhurst (Office for National Statistics, United Kingdom)
Paul Vickers (Office for National Statistics, United Kingdom)

Discussant: Steve Cohen (National Science Foundation, USA)

Session Organizer: Jock Black (National Science Foundation, USA)

9:00–10:30 a.m. **Ballroom C**

CONCURRENT SESSION IX-C:
SAMPLE DESIGN

Chair: Chris Chapman (National Center for Education Statistics, USA)

Properties of Alternative Sample Designs and Estimation Methods for the Consumer Expenditure Surveys
John Eltinge (Bureau of Labor Statistics, USA)

The American Community Survey Sample Design: An Experimental Springboard
Megha Joshipura (U.S. Census Bureau)
Steven Hefter (U.S. Census Bureau)

An Adaptive Sample Allocation for a Multiple Objectives Survey of Business
Daniela Golinelli (RAND Corporation, USA)
Gregory Ridgeway (RAND Corporation, USA)
John Adams (RAND Corporation, USA)

Discussant: Jill Montaquila (Westat, USA)

Session Organizer: Chris Chapman (National Center for Education Statistics, USA)

10:30–11:00 a.m. **Ballroom Foyer**
Break

11:00 a.m.–12:30 p.m. **Ballroom A**

CONCURRENT SESSION X-A:
WEB APPLICATIONS

Chair: Paige Harrison (Bureau of Justice Statistics, USA)

Developments in Electronic Survey Design for Establishment Surveys
Grace O'Neill (U.S. Census Bureau)

Increasing Response Rates: Pre-Notification
Dominic Lusinchi (Far West Research, USA)

Enhancing Web-Based Data Collection Using Excel Spreadsheets
Dan Jackson (Bureau of Labor Statistics, USA)
Michele Eickman (Bureau of Labor Statistics, USA)

Discussant: Duane Cavanaugh (U.S. Census Bureau)

Session Organizer: Paige Harrison (Bureau of Justice Statistics, USA)

11:00 a.m.–12:30 p.m. **Ballroom B**

CONCURRENT SESSION X-B:
VARIANCE ESTIMATION II

Chair: John Bushery (U.S. Census Bureau)

On X11 Seasonal Adjustment and Estimation of Its Variance
Michail Sverchkov (Bureau of Labor Statistics, USA)
Stuart Scott (Bureau of Labor Statistics, USA)
Danny Pfeffermann (The Hebrew University of Jerusalem, Israel, and University of Southampton, United Kingdom)

Diagnostic Process to Assess the Effects of Truncating Extreme BRFSS Sampling Weights
Henry Roberts (Centers for Disease Control and Prevention, USA)
Elizabeth Hughes (Centers for Disease Control and Prevention, USA)
Ruth Jiles (Centers for Disease Control and Prevention, USA)
Robert Woldman (North Carolina Department of Health and Human Services, USA)

An Examination of Alternative Variance Estimators
Laura Ozcoskun (U.S. Census Bureau)
Samson Adeshiyan (U.S. Census Bureau)

Discussant: Amrut Champaneri (Department of Transportation, USA)

Session Organizer: Amrut Champaneri (Department of Transportation, USA)

11:00 a.m.–12:30 p.m. **Ballroom C**

CONCURRENT SESSION X-C:
IMPUTATION

Chair: Alan Jeeves (Bureau of Transportation Statistics, USA)

Multiple Imputation and Estimating Aggregate Productivity Growth in Manufacturing
Kirk White (U.S. Census Bureau)
Amil Petrin (University of Chicago, USA)
Jerome Reiter (Duke University, USA)

Multiple Imputation of Right-Censored Wages in the German IAB Employment Register Considering Heteroscedasticity
Thomas Buettner (Institute for Employment Research, Germany)
Susanne Raessler (Institute for Employment Research, Germany)

Imputing Missing Values in the Common Core of Data for Use in Computing the Averaged Freshman Graduation Rate
Jack Buckley (National Center for Education Statistics, USA)
Marilyn McMillen Seastrom (National Center for Education Statistics, USA)
Chris Chapman (National Center for Education Statistics, USA)

Discussant: Patrick Flanagan (U.S. Census Bureau)

Session Organizer: Amrut Champaneri (Bureau of Transportation Statistics, USA)

Abstract Booklet

This section represents abstracts received as of August 30, 2007.

The following abstracts have not been edited for content.

CONCURRENT SESSION II-A:
MONITORING, MEASURING, AND ADJUSTING FOR NONRESPONSE I

Monitoring Response to a Multi-Wave Medical Establishment Survey: How Different Are Responders From Nonresponders?
Jessica Graber (National Opinion Research Center, USA)

In an effort to improve chronic disease management and health outcomes among underserved populations, the Bureau of Primary Health Care (BPHC) now requires all federally-funded community health centers (HCs) to implement the Health Disparities Collaborative (HDC) model, a quality improvement program launched by the BPHC in 1998 that aims to enhance the treatment of chronic illness through evidence-based practices, clinical decision support systems, and patient self-management. NORC, in partnership with the University of Chicago and the MidWest Clinicians Network, evaluated the impact of the HDC at more than 150 HCs in the Midwest and West Central regions of the United States, collecting data in two waves to better assess the long term integration of key HDC administered questionnaires were mailed to 1,504 medical providers, administrators and other staff at HCs that had participated in the HDC for at least one year. Nonrespondents were subject to extensive follow-up; receiving additional questionnaires, letter of support by BPHC officials and telephone prompting producing an overall response rate of 68.1%. In 2005, HCs were re-contacted and 1,456 respondents were asked to complete a similar, but shorter questionnaire. Despite significantly less follow-up effort, the final response rate from the second wave was 68.7%. While the two cross-sectional samples were not identical, over half (53.7%) of Wave II respondents were also surveyed in the Wave I effort, with 45% completing surveys at both points in time. This overlap in sample affords us the opportunity to monitor survey response over time. In this paper we review individual and HC level characteristics of both responders and nonresponders, and develop strategies for maintaining or increasing survey participation across data collection waves.

Factors Affecting Response to the Occupational Employment Statistics Survey
Polly Phipps, Carrie Jones, and Clyde Tucker (Bureau of Labor Statistics, USA)

The Occupational Employment Statistics (OES) is a bi-annual establishment survey of wage and salary workers designed to produce data on occupational employment and wages by industry for the U.S., States and certain US territories, and Metropolitan Statistical Areas within States. This voluntary survey of establishments with one or more employees is conducted by State employment workforce agencies in cooperation with the Bureau of Labor Statistics. While the response rate achieved by OES is quite high, particularly when compared to other U.S. and BLS establishment size (Jones, 1999). Several studies have identified factors that influence the likelihood that establishments will respond to a survey request (Tomaskovic-Devey, Leiter, and Thompson, 1994; Willimack, Nichols, and Sudman, 2002). Tomaskovic-Devey and colleagues propose that organizational behavior and structure explains nonresponse, including factors related to authority, capacity and the motives of the organization and designated respondent.

Willimack and colleagues have established a conceptual framework for large organizations that includes factors affecting the external environment, the business, the respondent, and the survey design. We test the effect of a number of these conceptual factors on response to the 2006 OES survey at a state level, including state economic conditions (revenues, general fund balances, and others), establishment characteristics (multi-establishment firm status, industry, size, location), and survey design and administration factors (contact strategies, survey form types, nonresponse follow-up strategies, state staff composition, experience, turnover). We also test the effect of response burden measured through participation in other BLS surveys and respondents who report for multiple establishment units. Finally, we attempt to evaluate survey design and administration factors that OES could modify in order to improve state- and national-level response rates.

Review of the Weighting Methodology for the Canadian Community Health Survey
Cathlin Sarafin, Steven Thomas, and Michelle Simard (Statistics Canada)

The regional component of the Canadian Community Health Survey (CCHS) is a cross-sectional survey that is designed to collect general health-related data from a sample large enough to provide information for more than 100 health regions across Canada. To date three cycles (Cycle 1.1 - 2001, Cycle 2.1 - 2003, and Cycle 3.1 - 2005) of the survey have been released and collection for the most recent cycle of the survey, cycle 4.1, began in January 2007. Cycle 4.1 marks a turning point for the survey and the health-survey program, as it is the beginning of a continuous collection process. Instead of having data collected for a period of one year every second year, data will be collected over a period of two years with no break in collection between cycles. As part of this survey redesign, the methodology of the weighting process is being reviewed. This revision includes some improvements to the weighting process such as the simplification of the weighting strategy to reduce the number of adjustments, changes to the nonresponse methodology, as well as the use of paradata,

also known as data collection process information, when constructing the Response Homogeneity Groups (RHG) for the nonresponse adjustments. This presentation will discuss the issues that emerged during the review of the weighting process. An empirical study will also be presented that will compare two different methods of deriving RHGs: the segmentation method based on chi-square tests will be compared to a scoring method, where logistic regression is used to model the nonresponse mechanism and estimate individual response probabilities.

An Examination of Nonresponse Error and Measurement Error by Level of Effort Using Frame Information, Survey Reports, and Paradata

Sarah Dipko and Kerry Levin (Westat, USA), and Mary-Helen Risler (Internal Revenue Service, USA)

A recent telephone survey of low-income taxpayers affords the opportunity to estimate contributions of different types of nonresponse-to-nonresponse error. Sample frame characteristics are used to examine nonresponse bias for three estimates: percentage of males under age 30, percentage whofiled 2005 taxes, and percentage with earned income in 2005. Measurement error for interview reports of tax filing and earned income is assessed via comparison to frame information from 2005 tax records, considered more accurate than interview self-reports for this analysis. Conducted for the Internal Revenue Service during the summer of 2006, the survey examined effects of a new certification requirement on Earned Income Tax Credit (EITC) claimants. A 10-minute survey was administered to a sample of taxpayers required to fulfill the new certification requirement for tax year 2005 (Test subjects, randomly selected from 25,000 taxpayers that had been randomly assigned to the test program), and to a randomly selected sample of taxpayers subject to standard requirements (Control subjects). Nearly 60% of the telephone sample (n=9,912) was not locatable, and response rates ranged from 20-30% depending on sample type. Six categories of nonresponse are examined, including refusals, maximum calls, non-contact (no answer/answering machine), sample members not found, nonworking numbers, and other nonresponse. Operational interventions used to increase response; refusal conversion; re-releasing finalized cases for additional calls; and refielding final non-contacts for additional calls are considered to represent a higher level of effort. Nonresponse error is estimated for six categories of nonresponse at two points: prior to operational intervention, and pursuant to all interventions. Effect of these interventions on nonresponse error for three statistics is examined. Measurement error is examined to assess whether interviews obtained via intervention(s) were subject to greater response error than those requiring no intervention.

CONCURRENT SESSION II-B:

TIME SERIES

Comparison of Methods for Computing Yearly Growth Rates From Weekly and Monthly Data, 1978 to 2005

Carol Blumberg (Energy Information Administration, USA)

The Energy Information Administration (EIA) collects data on volumes of gasoline (all grades) and distillate fuel oil (mostly heating oil and diesel fuel for vehicles). Monthly data come from a census of appropriate companies. Weekly data are from samples of these companies. Estimates of total volume are then formed from these data. The approximate release times are: Monthly estimates based on weekly data (abbreviated MFW)—11 days Preliminary estimates (PE) from the monthly census—60 days Final estimates (FE) that include late submissions and resubmissions—June of the following year. The ideal ratio (IR) for computing yearly growth rates here is the FE for a particular month divided by FE for that month in the previous year. However, EIA customers, such as financial analysts and industry experts, want estimates of yearly growth quickly. So, the IR is not practical. This study focused on constructing practical alternatives to the IR. The questions investigated were: 1. If MFW and PE are used as the numerators, what are the best denominators to use? 2. If cumulative columns for 3-, 6-, or 9-months or an entire year based on MFW or PE are used: a. What are the best denominators? b. Is there seasonality in the growth rates? The criteria for deciding between the alternative ratios were the differences in means and standard deviations, mean square error, and correlations with IR.

Ratios were further compared on percentage of times they were within 1% and 2% or in the same direction (both positive or both negative) as the IR. Data from 1978 through 2005 were used. Although different methods did better on certain criteria, overwhelmingly the best denominator in all cases was the PE. No seasonality was found in the growth rates, even though there is seasonality in the monthly volumes for the products.

The X-13A-S Seasonal Adjustment Program

Brian Monsell (U.S. Census Bureau)

In collaboration with the current developers of the SEATS seasonal adjustment program, a beta release of a seasonal adjustment package that produces model based seasonal adjustments from SEATS as well as X-11 seasonal adjustments has recently been made available to users. This program allows users to generate X-11 and SEATS seasonal adjustments using the same interface, and compare these seasonal adjustments using a common set of diagnostics. This session will show new features for generating accessible output, metadata and XML output, demonstrate the use of new modeling and diagnostics integrated into the software, and discuss further directions for this work.

Coherent Trends, Turning Points, and Forecasts for American Community Survey Data

Tucker McElroy (U.S. Census Bureau)

The American Community Survey (ACS) provides one-year (1y), three-year (3y), and five-year (5y) estimates of various demographic and economic variables for each community, although for small communities the 1y and 3y may not be available. These survey estimates are not truly measuring the same quantities, since differing amounts of smoothing are utilized. We present a method for generating trends, turning points, and forecasts of ACS data at 1y, 3y, and 5y intervals, in such a way that the estimates are compatible, which allows for comparisons across communities. The filters utilized are non-model-based, require only a short span of data, and are designed to preserve the appropriate linear characteristics of the time series that are relevant for trends, turning points, and forecasts respectively. The basic method, which only requires polynomial algebra, is outlined and applied on ACS data. The resulting filters are analyzed in the frequency domain.

Empirical Evaluation of X-11 and Model-Based Seasonal Adjustment Methods

Richard Tiller, Stuart Scott, and Dan Chow (Bureau of Labor Statistics, USA)

X-11 and model-based seasonal adjustments are compared for 82 series from three U.S. Bureau of Labor Statistics programs, 1) establishment employment, hours, and earnings, 2) consumer prices, and 3) producer prices. Results are interpreted according to connections between X11 filter choices and ARIMA models following work of Depoutot and Planas and Bell, Chu, and Tiao. Both automatic andanalyst adjustments are analyzed. Weaknesses or shortcomings of automatic adjustments are pointed out, along with ways that the methods can complement each other in pointing to improvements via analyst-assisted adjustments. Some employment series exhibit special difficulties in modeling, while several price series require care in treating outliers and interventions. A variety of diagnostic statistics and graphs are used to support the findings.

CONCURRENT SESSION II-C:

FRAMES AND COVERAGE

Practicability of Including Cell Phone Numbers in Random Digit Dialed Surveys: Pilot Study Results From the Behavioral Risk Factor Surveillance System

Michael Link (Nielsen Media Research, USA), Michael Battaglia and Larry Osborn (Abt Associates, Inc., USA), Martin Frankel (Baruch College, CUNY and Abt Associates, Inc., USA), and Ali Mokdad (Centers for Disease Control and Prevention, USA)

Researchers are increasingly concerned about the rapid growth of cell phone-only households (i.e., households with no landline that are accessible only by cell phone) and the associated potential for bias in estimates obtained from telephone surveys, which do not sample from cell phone exchanges. A pilot study conducted in Georgia, Pennsylvania, and New Mexico as part of the Behavioral Risk Factor Surveillance System (BRFSS), the world's largest random digit-dialed (RDD) telephone survey, evaluated the effectiveness of conducting the BRFSS interview with a sample drawn from cell phone numbers. The BRFSS currently uses a list-assisted sample of landline telephone numbers, conducting interviews only over landlines. For the pilot, a sample of telephone numbers was drawn from dedicated cellular 1,000-blocks in each state. In screening for eligible sample members, all adults with only cellular telephone service were included and a sub-sample of adults with both cell phone and landline telephones was selected, resulting in approximately 600 interviews with each group. We report on response rates and demographic characteristics of respondents. Additionally, using a new approach to weighting cell phone samples, we illustrate how inclusion of cell phone respondents affects prevalence estimates of key health conditions and risk behaviors. The weighting procedure involved dividing the BRFSS

landline sample into adults with landline and cellular telephones and those with only landline telephones, and dividing the cell phone sample into adults with cell and landline telephone, and those with only cell phones. Using the joint distribution of cases with both landlines and cell phones as a starting point, the landline only and cell phone-only cases were incorporated into the overall weighted sample using external data sources for population totals. The methodology employed in this study and the lessons learned, including the costs of conducting surveys over cell phones, have wide application for other telephone surveys.

Measuring and Adjusting for Frame Undercoverage of the State and Local Value Put-In-Place (VIP) Survey

Thuy Trang Nguyen and Shadana Myers (U.S. Census Bureau)

The U.S. Census Bureau conducts monthly the State and Local (S&L) VIP survey to measure the value of construction put in place for building and non-building structures owned by S&L governments. We also collect fiscal year data on similar construction in the Annual Survey of State and Local Government Finances (ASGF). Conceptually, these estimates should be comparable on a fiscal basis; nevertheless, they have continued to differ during the past decades. The S&L VIP estimates are consistently lower than the ASGF estimates. The major difference is attributed to the undercoverage of the S&L VIP frame. This paper discusses the results of a study to measure the coverage of the S&L VIP sampling frame by determining the match rates of projects collected from an independent source to projects on the VIP frame. The frame undercoverage is then adjusted by applying undercoverage adjustment factors derived from the match rates. The study involved the following stages: (1) draw a sample of government agencies from the ASGF, (2) collect construction project information from the agencies, (3) sub-sample the projects for inclusion in the study, (4) match the sub-sampled projects from the agencies to projects in the S&L VIP frame to determine the coverage, (5) evaluate the match rates and estimate the undercoverage adjustment factors, (6) apply the undercoverage adjustment factors.

Comparing the Quality of the Master Address File and the Current Demographic Household Surveys' Multiple Frames

Xijian Liu (U.S. Census Bureau)

The current demographic household surveys conducted by the Census Bureau selected samples from a multiple frame system that obtained addresses from decennial Census, building permits, and area listings. The Census Bureau plans to redesign these surveys and will use the Master Address File (MAF) as the source of sample addresses. The Master Address File will initially be updated through various operations for the decennial census. The Census Bureau will continue to update the MAF using the U.S. Postal Service's delivery sequence file (DSF), which contains an updated list of their mail delivery points and using a field update operation designed to improve the MAF in targeted areas. To support this plan, the Census Bureau is conducting several evaluations. These evaluations compare the quality of the MAF and the current frames. Their focus is on the over-all quality at the national level as well as on two sub-universes: the new construction sub-universe and the sub-universe currently covered by an area frame that is primarily in rural area. This paper will present recent findings from these evaluations.

Impact of Preliminary Versus Final Economic Census Data on the Universe Extraction Process for Current Business Surveys

Kari Clark and Carol King (U.S. Census Bureau)

The U.S. Census Bureau selects a new sample for its current business surveys approximately once every five years. For use in constructing the sampling frame, the first step in the sample selection process is the extraction of establishment records from the Census Bureau's BusinessRegister used in the creation of an establishment list. As part of this process, census data will also be extracted for establishments that were active during the time of the latest Economic Census. The data from the Census and the Business Register are used to determine each establishments industry classification and the major kind of business for the sampling unit. Sampling units consist of aggregations of one or more establishments, based on the organization of the company. The industry classification is used to evaluate if the establishments could be considered in-scope to the current business surveys and is also used as a stratification variable in the sample design. Another stratification variable used in the sample design is a measure of size. This measure of size represents a full year of activity in terms of revenue at both the establishment and sampling unit level, and could also be calculated using Census data. When the extraction process was done for the current sample, only preliminary data from the 2002 Economic Census was available. Using the final census data would result in the extraction process being done a year later. Research was conducted to evaluate how using the preliminary rather than the final census data would have affected the industry classification and the measure of size for both the establishment and the sampling unit. This paper will detail the extraction process and the impact on the establishment list had the final census data been used.

CONCURRENT SESSION III-A:
DISCLOSURE I

Comparative Evaluation of Four Different Sensitive Tabular Data Protection Methods Using a Real Life Table Structure of Complex Hierarchies and Links
Ramesh Dandekar (Energy Information Administration, USA)

The practitioners of tabular data protection methods in federal statistical agencies have some familiarity with commonly used table structures and require some insight on how to evaluate appropriateness of various sensitive tabular data methods when applied to their own table structure. With that in mind, we use a real life typical table structure of moderate hierarchical and linked complexity and populate it with synthetic micro data to evaluate relative performance of four different tabular data protection methods. The methods selected for the evaluation are: 1) classical cell suppression 2) Lp-based CTA (Dandekar 2001), 3) USBC's network flow-based cell suppression and 4) USBC's micro data level noise addition method. The outcome from the comparative evaluation is available from http://mysite.verizon.net/vze7w8vk/.

Easy to Use Is Putting the Cart Before the Horse: Effective Techniques for Masking Numerical Data
Krish Muralidhar (University of Kentucky, USA) and Rathindra Sarathy (Oklahoma State University, USA)

In a recent paper, William Winkler of the Census Bureau observed the following: Statistical agencies have typically adopted masking methods because they are easy to implement. The easiest-to-implement methods seldom, if ever, have been justified in terms of preserving one or two analytical properties and preventing re-identification. In extreme situation, the crude application of masking methods may yield a file that cannot be used for analyses yet still allows some re-identification. In extreme situations, the crude application of masking methods may yield a file that cannot be used for analyses yet still allows some re-identification? (William Winkler, Modeling and Quality of Masked Microdata Research Report Series, Statistics 2006-01 http://www.census.gov/srd/papers/pdf/rrs2006-01.pdf). We could not agree more. We believe that recent advances in data masking techniques provide statistical agencies with sophisticated techniques that, while they may be little more difficult to implement compared to some other techniques, provide very low level of information loss and disclosure risk. The specific new techniques that we consider are the sufficiency based perturbation approach suggested by Burridge (2003) and the data shuffling (patent pending). We compare these techniques with the "easy to use" techniques such as noise added perturbation, micro-aggregation, and data swapping in terms of both disclosure risk and data utility. We will evaluate the risk of identity and value disclosure as well as the utility of these methods based on their ability to maintain marginal characteristics and multivariate relationships. In addition, we also propose to evaluate the ability of these techniques to maintain characteristics of sub-domains of the data, something that has not been evaluate previously. Obviously, give that an infinite number of possible sub-domains that can be formed, it is necessary to make some decisions that will allow us to make generalizations about the performance of these techniques.

Comparing Fully and Partially Synthetic Datasets for Statistical Disclosure Control in the German IAB Establishment Panel
Joerg Drechsler, Agnes Dundler, Stefan Bender, and Susanne Raessler (Institute for Employment Research, Germany)

For datasets considered for public release, statistical agencies have to face the dilemma of guaranteeing the confidentiality of survey respondents on the one hand and offering sufficiently detailed data for scientific use on the other hand. For that reason a variety of methods to guarantee disclosure control is discussed in the literature. In this paper we compare two approaches based on multiple imputation. The first, proposed by Rubin (1993), generates fully synthetic datasets while the second imputes values only for selected variables that bear a high risk of disclosure. We apply the two methods to a set of variables from the 1997 wave of the Ger-man IAB Establishment Panel and evaluate their quality by comparing results from an analysis by Thomas Zwick (2005) with the original data with results we achieve for the same analysis run on the dataset after the imputation procedure.

CONCURRENT SESSION III-B:
ADVANCES IN DATA EDITING

Investigation of Selective Editing Procedures for the Annual Survey of Government Finances
Loretta McKenzie, Terri Craig, and Carma Hogue (U.S. Census Bureau)

The U.S. Census Bureau re-engineered the edit procedures for the Annual Survey of Government Finances. Beginning with the 2004 survey, the Census Bureau investigated the use of selective editing. A score was assigned to each individual government unit based on the effect that a change in the government's reported data would have on the final estimates. Using predetermined critical score values, those government units with the largest potential impact on the estimates were pinpointed as candidates for manual review. Because the survey form covers a large number of variables, some of which are very volatile from one year to the next, determining which variables to use to develop a score for a governmental unit was problematic. We chose several score functions involving different combinations of these variables. For each score function we set edit referral rates and compared absolute pseudo-biases of resulting estimates to identify edit referral rates that would reduce edit burden and maintain quality. This paper reviews the selective editing technique, the edit research process, and the problems that we encountered in attempting to apply it to the Annual Survey of Government Finances. We used empirical research methods to conclude that other micro editing techniques should be used prior to selective editing. Data sources were the 2002 Census of Governments Finances and the 2004 Annual Finance Survey.

Measuring Edit Efficiency in the Economic Directorate of the U.S. Census Bureau
Broderick Oliver and Katherine Jenny Thompson (U.S. Census Bureau)

The correction of survey returns accounts for a significant portion of the cost and time of conducting sample surveys. Editing and imputation are carried out to detect and correct inconsistencies in the data resulting from sources such as the respondent, interviewer, and data capture instrument. In 2006-2007, the Economic Directorate of the United States Census Bureau conducted a series of studies to measure the editing efficiency of several of its Economic surveys and censuses. The first portion of these studies focused on a series of measures applied to the original reported data in comparison to the final tabulated data to assess the size of corrections to the data broken out by the source of the change; for example, (1) analyst corrections; (2) analyst imputes; (3) automated imputes. In this paper, we build upon the results of these initial studies by examining these same measures at different phases of the survey processing cycle, using data from the Annual Trade Survey, conducted by the U.S. Census Bureau. As we apply and interpret these measures, new ways of evaluating edit efficiency are revealed.

Improving the Efficiency of Data Editing and Imputation for a Large-Scale British Annual Business Survey
Alaa Al-Hamad and Gary Brown (Office for National Statistics, United Kingdom) and Pedro Silva (University of Southampton, United Kingdom)

This paper reports results from a project to evaluate and improve the editing and imputation approach adopted in the Annual Business Inquiry Part 2 (ABI/2). This is joint work carried out by the UK Office for National Statistics (ONS) and Southampton University's Statistical Sciences Research Institute. The ABI/2 is a large-scale annual survey covering most sectors of the British economy, with an annual sample of around 60,000 businesses. We examined detailed specifications for the current editing and imputation processes, and their connections to data collection instruments, data capture, and estimation methods. A variety of quality indicators, impact measures, and statistical editing and imputation techniques, were tested on three years of pre- and post-edited data. A number of alternative approaches to the overall editing and imputation process were investigated to maximize efficiency without impacting negatively on quality. Preliminary results suggest that these will yield increased benefits to the survey.

An Empirical Investigation Into Macro Editing on Two Economic Surveys
Katherine Jenny Thompson and Laura Ozcoskun (U.S. Census Bureau)

The identification of outliers in survey estimates prior to release is a widely accepted stage of data review. This identification process is used to determine whether outlying estimates are the results of uncorrected respondent or data capture errors or are in fact values that provide useful information (e.g., indicators of change in target estimates). Such identification is generally performed after completing micro-level review, during which the individual questionnaire returns are scrutinized and corrected on a flow basis. At the macro-level review phase, distributions of tabulated cell estimates are reviewed, within both the current collection period and in contrast to corresponding prior period estimates. Macro-editing techniques rely on distributional analyses.

Survey data estimates rarely have known parametric distributions. Moreover, quantitative economic data are often best assessed via ratio comparisons of totals (e.g., current to prior estimates, wage per employee). Consequently, macro-editing techniques that utilize survey data must employ non-parametric or robust methods. Moreover, since the original set of estimates will contain outliers, these methods should be resistant. Ratio comparisons are often quite effective at identifying outlying estimates, but can lead to redundant work since often the same estimation cells are repeatedly identified using different sets of estimates. A multivariate outlier detection method that simultaneously considers all key estimates to identify all (or most) outlying estimation cells could save considerable time. Thompson (2006) presents promising results with applications of the Hidiroglou-Berthelot edit and with a robust Mahalonobis distance measure to estimates collected from the U.S. Census Bureau's Annual Capital Expenditures Survey. In this paper, we apply these recommended techniques to data collected from two economic programs administered by U.S. Census Bureau with the goal of determining whether the recommended methods can be utilized with few modifications by other programs.

CONCURRENT SESSION III-C:

EXAMINING THE EFFECTS OF MODE AND RESPONDENT CHARACTERISTICS ON SURVEY PARTICIPATION

Incorporating a Multi-Modality Design Into a Random-Digit-Dialing Survey

Michael Battaglia and Larry Osborn (Abt Associates Inc., USA), Martin Frankel (Baruch College, CUNY and Abt Associates Inc., USA), Michael Link (Nielsen Media Research, USA), and Ali Mokdad (Centers for Disease Control and Prevention, USA)

The Behavioral Risk Factor Surveillance System (BRFSS) is a monthly state-based random-digit-dialing (RDD) survey that measures health risk factors and health conditions. Most RDD surveys are conducted using only the telephone survey mode of data collection. RDD survey response rates have been declining over the past ten years, making it important to examine alternatives to the single-mode approach. The paper describes a test of one alternative. In six states, a multi-modality design for the BRFSS was tested. A list-assisted RDD sample was selected and matched with a database of residential telephone numbers and addresses. For the sample telephone numbers with an address match, a mail survey with telephone follow-up was employed. The telephone follow-up involved contacting mail nonrespondent households and attempting to conduct the interview by telephone. For sample telephone numbers with no address match, the survey was conducted by telephone alone. After discussing the design and implementation of the six-state pilot survey, we focus on response rates by mode using the single-mode BRFSS survey in each state as a comparison sample. The paper will discuss weighting and estimation issues for multi-modality RDD designs and will also examine mode effects.

Interviewer-Reported Reasons for Conducting Interviews by Telephone in the National Health Interview Survey, 2005

Barbara Stussman, Catherine Simile, and James Dahlhamer (National Center for Health Statistics, USA)

In an effort to increase response rates and decrease costs, many survey operations have begun to use several modes of administration to collect relevant data. While the National Health Interview Survey (NHIS), a multipurpose household health survey conducted annually by the National Center for Health Statistics, Centers for Disease Control and Prevention, is primarily a face-to-face survey (e.g., 75% of interviews in 2005 were conducted entirely by personal visit), interviewers also rely on the telephone to complete some interview sections. Once a personal visit has occurred, interviewers may use telephone follow-up if a personal visit follow-up is not possible. Any one of the NHIS's four main sections (household composition, family, sample child, sample adult) may be conducted by a telephone follow-up. In 1997, 18% of all completed interviews included at least one main section that was conducted primarily by telephone. By 2005, the proportion of interviews in which at least one main section was conducted primarily by telephone had risen to almost 25%. The purpose of this study is to describe the field circumstances that give rise to interviewers use of the telephone instead of a personal visit in completing sections of the NHIS interview. Textual narratives detailing why the telephone was used are collected for every interview for which sections were administered primarily by telephone; in 2005, 10,461 such entries were collected. This study will summarize those data to describe the main reasons given for the use of the telephone. The results may be useful in adjusting field procedures and evaluating the impact on the quality of data collected.

Response Profile of the 2005 American Community Survey

Geoffrey Jackson (U.S. Census Bureau)

The objective of this paper is to use 2005 American Community Survey (ACS) data to compare and analyze the demographic, social, and economic characteristics of people who respond by mail, Computer Automated

Telephone Interview (CATI), and Computer Automated Personal Interview (CAPI). There is a high cost involved with ensuring a high response rate for the ACS. Potential respondents are first contacted through a mail questionnaire. Those that don't respond to the mail questionnaire are followed up with by CATI. Finally, a portion of CATI nonrespondents is followed up by CAPI. We created a profile of who responds to what mode. Survey nonresponse is typically due to three reasons: noncontact, resistance, and inability to complete the survey. We tested mail questionnaire nonresponse due to resistance and those who were unable to respond. A logistic regression model for nonresponse due to resistance and inability to respond was created.

The Influence of Selected Factors on Student Survey Participation and Mode of Completion
Tracy Hunt-White (National Center for Education Statistics, USA)

This paper will examine how selected factors pertaining to student characteristics, institutional characteristics, and survey design features are related to survey participation overall and by mode of completion (Web vs. telephone). The four factors to be examined are respondent characteristics, the institution's social environment (e.g., enrollment size) and technological environment (e.g., number of computers on campus), and survey design features (e.g., number of contact calls). The respondents for this paper come from the 2004 National Postsecondary Student Aid Study (NPSAS). The respondents consist of approximately 69,000 NPSAS study respondents who represent 18.3 million undergraduates enrolled in 2-year and 4-year postsecondary institutions in the U.S. 2003-04. Study respondents for NPSAS not only included those who completed a Web-based or telephone interview, it also included students who were not interviewed, but for whom key information could be obtained from school records or federal financial aid databases. To identify the variables that comprise the factors associated with survey participation, three sources of data are used. The 2004 National Postsecondary Student Aid Study (NPSAS) is used to group sample members and to supply variables that comprise the survey design and respondent characteristics factors. The Integrated Postsecondary Education Data System (IPEDS) is used mainly to identify institutional variables that comprise the social environment factor. Data from the College Board is used to form the technological environment factor. Researchers have watched their efforts to obtain high response rates grow more costly and time consuming. This paper will add to researchers' understanding of how student characteristics, institutional characteristics, and survey design features influence participation.

CONCURRENT SESSION IV-A:
EMPLOYMENT/LABOR STATISTICS

Distorted Measures of Employment in Charitable Organizations: Some Remedies
Martin David (University of Wisconsin—Madison, USA)

Public sector failures lead to a large understatement of employment in charitable organizations. Multiple forces lead to this understatement. Partitioning private business into charities, other exempt organizations, and for profit business has a low priority in Federal Statistical Agencies. Regulatory failures in IRS oversight of exempt organizations compromise available statistics; the count of active organizations, data on employment, coverage of available reports, and consistency in reporting. The incentive for IRS to regulate exempt entities is negative, as the activity does not generate net revenue. Because exempt organizations constitute a small part of private business, publication of estimates for establishments is limited by the imperative not to disclose proprietary information. Finally, regulation of burden in completing government forms leads to peculiar censoring of data within the population of exempt entities. This analysis demonstrates that existing published estimates of employment in charitable organizations is understated. We link IRS information returns to the BLS/QCEW. A substantial proportion of employers cannot be matched. Employment on IRS returns contains substantial nonresponse. Imputation of QCEW employment to matched organizations and augmenting the available census of IRS returns with employment in exempt organizations that are not covered produces aggregates that are substantially larger than the published Economic Census for 2002. A combination of more sophisticated imputation of information returns and matching of IRS information to records of payroll tax submissions, IRS/Form 941, can overcome understated employment.

A Proposed Model for Microintegration of Economic and Social Data
Paul De Winden, Koos Arts, and Martin Luppes (Statistics Netherlands)

Globalization affects all aspects of economic and social life. In order to study the effects of an open economy on employment and welfare, combined microdata from business surveys, social surveys and administrative registers are required to make causal inferences. Statistics Netherlands uses two sets of microdata to construct

a framework in order to analyze the complex relationships between the dynamics of enterprises and the outcome on employment and welfare. The combined dataset contains hierarchical data on four different statistical units: enterprise groups, enterprises, jobs and individual persons. The backbone of the system contains information from administrative (governmental) registers and therefore covers the total population of enterprises and employees. The required variables for analytical purposes are retrieved from business and social surveys. Depending on the type of surveys and cross sectional slices of units the researcher has to deal with the methodological challenge of sample reweighing and the use of multilevel models. However, these issues are outweighed by the benefits of combining microdata from different sources. In this study first results are presented on the relationship between job creation, job destruction and economic behavior of firms.

Methodologies for Estimating Mean Wages for Occupational Employment Statistics (OES) Data
Mallika Kasturirangan, Shail Butani, and Tamara Zimmerman (Bureau of Labor Statistics, USA)

The Occupational Employment Statistics (OES) is a joint Federal/State partnership program with a sample size of 1.2 million establishments over a 3-year period (six semi-annual panels each consisting of 200,000 establishments). The OES collects occupational employment and wage data for approximately 800 occupations at the MSA by 4-5 digit industrial (NAICS) level. Because of the burden on respondents, this survey is designed to collect wage data in intervals rather than exact wages for individual employees. In this talk, we will present the previous research work on the construction of lower and upper bounds of the intervals; alternative methods for estimating mean wages-arithmetic, geometric, and NCS mean wages; updating of wages from prior panels; and calculation of mean wages for the upper open-ended interval L (i.e., employees making $70 or more per hour in the years 2003–2005). This study further examines several methods for approximating mean wages for interval L for occupations that have significant employment (>5%) in interval L and validates the OES methodology on independent data sets from the Current Population Survey for years 2003, 2004, and 2005.

Estimating the Measurement Error in the Current Population Survey Labor Force - A Latent Class Analysis Approach With Sample Design
Bac Tran and Justin Nguyen (U.S. Census Bureau)

This paper describes the results of a Markov Latent Class (MLC) simulation study and its application to data from the Current Population Survey (CPS). Latent Class Analysis offers a new way of estimating response errors in longitudinal surveys. Latent Class Models (Clams) are mathematical models for characterizing the latent variables and their relationships with observed variables. The main purpose of the use of MLC models in the context of the CPS panel survey is to determine the classification error in the recorded employment status without the necessity of collecting reentries data. The simulation study determined how sensitive the MLC model estimates of classification errors are for violations of certain model assumptions. The results of a test performed with the CPS data (using Latent GOLD 5.0) give estimates of the sizes of the classification errors and of the latent classes when applying the relevant types of MLC models to CPS data while taking into account the complex sampling design.

CONCURRENT SESSION IV-B:
CONFIDENTIALITY/PRIVACY

Respondent Consent to Link Survey Data With Administrative Records: Results From a Split-Ballot Field Test With the 2007 National Health Interview Survey
James Dahlhamer and Christine Cox (National Center for Health Statistics, USA)

Data from the National Health Interview Survey (NHIS), a multipurpose household health survey conducted annually by the National Center for Health Statistics, Centers for Disease Control and Prevention, are routinely linked to other health-related administrative records to enhance their analytic potential. To improve the accuracy of person record matches, the NHIS attempts to collect unique identifiers such as Social Security Number (SSN) and Medicare number. However, as public concerns over identify theft have grown, the percentage of NHIS respondents providing this information has decreased dramatically. In recent years over 50% of adult NHIS respondents have refused to report their SSN, up from roughly 15% in 1993. In this paper we present preliminary results from the 2007 NHIS split-ballot field test of approaches designed to increase the number of respondents consenting to data linkage activities and to improve item response rates in the collection of Social Security and Medicare numbers. In ballot one, the respondent was first asked for permission to link. If

the respondent did not refuse, he/she was asked for the last four digits of his/her SSN. If applicable, a second question followed to obtain a similar portion of the Medicare number. In ballot two, the respondent was first asked to supply a portion of his/her SSN and Medicare number. If refused, the respondent was asked if he/she would consent to data linkage without use of the unique identification numbers. Descriptive, bivariate, and multivariate analyses are performed to assess the relative effectiveness of these sets of items for producing 1) data eligible for linkage, and 2) data eligible for linkage using unique identifying information. We discuss the implications of our findings for future NHIS data linkage activities.

Consumer Privacy
Howard Fienberg (Council for Marketing and Opinion Research, USA)

The goal of this presentation is to outline how consumer privacy laws, and professional codes, standards and best practices, impact survey and opinion research from small-scale surveys, to massive research efforts like the 2010 Census. It will broadly explain current laws and regulations, and introduce emerging legislative issues. In addition, the presentation will demonstrate how laws and trends that don't directly regulate researchers can impact and potentially harm their research, by discouraging respondent cooperation. The presentation will provide an overview of these issues for the research profession. Outline and Scope of Presentation 1. Contacting respondents a) The Telephone Consumer Protection Act (TCPA), the Telemarketing Sales Rule (TSR), and the Do Not Call Registry b) Contacting cell phones c) What time of day can (and should) you contact respondents? d) What disclosures should (or must) you make when talking with respondents? e) How can you use auto-dialers, faxes, email, and online cookies? 2. Data privacy a) Gramm-Leach-Bliley, HIPPA, and COPPA b) Fair information practices c) Privacy policies and practices d) Data security breach laws e) Informed consent 3. New Trends and Emerging Legislation Impacting Research 4. The Bad Guys Make Us All Look Bad—How practices and laws to which researchers are not subject can hurt respondent cooperation.

An Introduction to the National Inmate Survey
Rachel Caspar and Chris Krebs (RTI International, USA) and Allen Beck and Paige Harrison (U.S. Department of Justice)

The Prison Rape Elimination Act (PREA) of 2003 (P.L. 108-79) requires the Bureau of Justice Statistics (BJS) to develop new national data collections on the incidence and prevalence of sexual assault within correctional facilities. The Act requires BJS to survey each year not less than 10% of all federal, state, and local correctional facilities and to provide facility-level estimates of sexual assault for each sampled facility. To fully implement PREA, BJS has developed a multiple-measure, multiple-method data collection strategy. One of these data collection activities, the National Inmate Survey (NIS), involves obtaining information directly from adult inmates on their experiences with sexual assault. This presentation will provide an overview to the NIS, including a discussion of human subjects issues, the special considerations that must be made for conducting interviews in correctional facilities, design of the sample, development of the survey instrument, and decisions regarding the mode of data collection. A summary of the work completed to date will be provided. Results from a Pilot Study may be presented if they are available in time.

CONCURRENT SESSION IV-C:
TOPICS IN ESTIMATION AND MODELING FOR NATIONAL AND INTERNATIONAL SURVEYS

Estimating Unemployment for Small Areas in Navarra, Spain
Maria Ugarte, Ana Militino, and Tomas Goicoa (Public University of Navarra, Spain)

In the last few years, European countries have shown a deep interest in applying small area techniques to produce reliable estimates at the county level. With this as the goal, the EURAREA project <http://www.statistics.gov.uk/eurarea>, founded by the European Union between 2000 and 2003, has investigated the performance of various standard and innovative methods in several European countries. However, the specificity of every European country, the variety of auxiliary information as well as its accessibility, makes the use of the same methodology in the whole of Europe difficult. Navarra is a small autonomous community located at the north of Spain. It has 10.000 km^2 and only 600.000 inhabitants, irregularly distributed in seven subdivisions. Navarra Statistical Institute (IEN) has provided data to the Spanish Statistical Institute (INE) as a member of the EURAREA project. Nowadays, IEN is interested in providing precise estimates of the unemployment population in every of its subdivisions (called comarcas) in the context of the Spanish Labor Force Survey. In this work we review the current estimation procedure used to provide these estimates. In addition, we discuss the behavior of several design-based, model-assisted, and model-based estimators using different auxiliary information such as sex-

age groups, municipality sizes, education, and the Navarra unemployment population register. We also discuss the benefits of using these models, and provide several methods for estimating the prediction error. We comment on the results and the viability of its implementation. More specifically we comment on the difficulties of estimating in very small areas where the samples are both very scarce and unstable.

Two-Step Versus Simultaneous Estimation of Survey-Non-Sampling Error and True Value Components of Small Area Sample Estimators

Swamy Paravastu and Tamara Zimmerman (Bureau of Labor Statistics, USA) and Jatinder Mehta (Temple University, USA)

Any sample estimator for a small domain can be written as the sum of a true value and sampling and non-sampling errors. It can be based on survey data obtained from the domain, but the sample size within the domain is often either zero or too small to provide a reliable estimate. In this paper, the precisions of sample estimators for small areas for a period are simultaneously improved, using domain indirect model-dependent estimators that borrow strength from auxiliary data collected for those areas. What is new about this method is that it simultaneously estimates a true value and the sums of sampling and non-sampling errors in two or more sample estimators of the same true value for each small area. For state employment, the three sample estimators considered are the current population survey (CPS) composite estimator and the estimators given by data from the current employment statistics (CES) survey and by extrapolated data from the quarterly census of employment and wages (QCEW). In principle, the simultaneous estimation method developed in this paper is superior to a two-step method of estimating the true-value and sampling-error components of a sample estimator ignoring non-sampling errors. A cross-sectional model for state employment is used to explain the advantages of the simultaneous estimation method. The statistical method used to estimate this model is an iteratively re-scaled generalized least squares method. The method corrects for non-response and measurement-error biases and sampling error in the CPS employment estimator for each of 51 small areas consisting of the 50 United States plus the District of Columbia.

Weighting and Estimation Methodology and Results From the American Community Survey Family Equalization Project

Mark Asiala (U.S. Census Bureau)

Historically the American Community Survey (ACS) has produced inconsistent estimates of households and householders and inconsistent estimates of husbands and wives in married couple households even though logically these estimates should be equal. In the 2005 ACS, the size of these inconsistencies at the national level was approximately 3.7 million more householders than households and approximately 1.8 million more spouses than married-couple households. Likewise, for unmarried-partner households there are approximately 176,000 more unmarried-partners than unmarried-partner households. The cause of these data inconsistencies was rooted in the current person weighting methodology which was independent of the housing unit weighting and did not consider relationship to the householder. This paper describes the current weighting methodology and changes introduced to reduce these data inconsistencies while having a minimal impact on other estimates and on the variances of the estimates. A three-dimensional raking methodology is used where the marginal control totals are derived from the survey itself rather than an independent source for the first two dimensions that are related to equalizing spouses and householders. Changes in the estimation of housing unit characteristics are also discussed. Empirical results from the implementation of this new methodology are presented based on the 2004 and 2005 ACS data.

TECHNICAL DEMONSTRATIONS

Making Sense of Data Via the Web - A Case Study Using Agricultural Data

Irwin Anolik (National Agricultural Statistics Service, USA)

This demonstration illustrates results of research conducted at the National Agricultural Statistics Service (NASS) to determine ways to disseminate data on the World Wide Web so that data customers can freely and effectively view and analyze NASS data using any web browser connected to the Internet. Specifically, we demonstrate web-based solutions that enable viewing, analyzing, and dynamically interacting with summary data at the state and county level. The demonstration presents important concepts and technologies that we implement on the NASS web site to enhance the ability to find data of interest, and see structure and pattern inherent in the data. We display historical data through the use of animated maps and charts as well as sparklines as developed by Edward Tufte and described as small graphic charts that can be inserted within text on a page. We also display current data using dynamic maps and charts, and making use of small multiples, galleries, and multi-

variate column views; all meant to give more context and power to the data customer. While this demonstration focuses on Agricultural Data, the principles and methods can be applied to other sources of survey and census data.

The National Health and Nutrition Examination Survey
Yinong Chong, Rosemarie Hirsch, Cheryl Fryar, and Jennifer Dostal (Centers for Disease Control and Prevention, USA)

NHANES is the only national survey that collects extensive health information from both face-to-face interviews and medical examinations. NHANES datasets became publicly available on its website in 2000, and a growing number of analysts are using the NHANES data to address major public heath issues in the U.S. To meet the growing demands of a wide range of data users, a team of content experts, instructional designer, and web developer from NCHS and the National Cancer Institute jointly developed this web-based tutorial to promote broader and more proficient use of NHANES data. The tutorial simulates a real life experience of how to conduct an analytical project from beginning to end, with a step-by-step flow through typical analytic procedures. The data preparation component focuses on navigating through the data structure and vast amounts of information, and on identifying and retrieving NHANES data files and variables of interest. The analysis component focuses on conducting statistical analyses with appropriate attention paid to the nuances of NHANES data, given its complex sample design, weighting requirements, and data structure. In addition, annotated SAS and SUDAAN program code and analytical guidance are organized and integrated into different modules and tasks to facilitate this learning process. The web tutorial has gone through numerous rounds of formal and informal usability testing, which greatly enhanced the design features to accommodate different learning styles. The tutorial has also been accredited by CDCs online learning system as four stand alone courses, allowing physicians, nurses, and epidemiologists to receive continuous education units upon successful completion. This basic tutorial will be further expanded to include STAT programs, and modules specific to dietary data and historic NHANES datasets.

Computer Audio-Recorded Interviewing (CARI)
Katherine Mason (RTI International, USA)

What is it? Why is it needed? How do you do it? This software demonstration will show you how a survey instrument can be given CARI capability for collecting sound files during in-person interviews. The resulting sound files can be used for reviewing field staff performance, for confirming the authenticity of interviews or for evaluating how well a questionnaire item “works” with real subjects. The demonstration will include administration of a questionnaire, which captures audio files, followed by demonstration of a monitoring system for controlling, managing and evaluating the resulting sound files.

Using a Pen Based Windows XP Tablet PC for Data Collection: Development of a Mobile System for Health Care Settings
Sarah Kalsbeek, Dick Paddock, Reginald Pendergraph, Helen Smith, and Vanessa Thornburg (RTI International, USA)

This presentation covers the development and implementation of a data collection effort using Pen Based Windows XP Tablet PCs.

This study collects information on respondents receiving outpatient substance abuse services and health care providers delivering those services. The study is being conducted in six states and one tribal organization. The CAPI (computer assisted personal interview) instruments were especially designed for Tablet PCs using Blaise. The tablets selected for this study are Motion Computer’s Model LS800. These devices provide a unique mix a standard computing environment (Windows XP with Table Extensions) coupled with small size and modest weight (8.94 x 6.69 x 0.87 inches 2.2 pounds). This provides a standard development platform that can be used well in a data collection effort that requires the interviewer to be mobile.

Tablets were chosen over laptops due to their light weight, ergonomic design, and capabilities. Their design allows field staff members to hold them for long periods of time in mobile survey settings including emergency rooms and health clinics where desks might not be available.

Key topics of this presentation include the logistics of using the tablet PC, the unique capabilities of the tablet PC including signature capture for respondent informed consent (handwriting character recognition), security concerns, usability issues, and project modifications for field deployment.

As this is a continuing study, this presentation includes observations from field staff and an update on study progress.

Demonstration of the Hand-Held Computer to Be Used for the 2010 Census
Karen Field (U.S. Census Bureau)

Demonstration Focus: The US Census Bureau has contracted with the Harris Corporation to automate portions of the 2010 Census. The data for three of the main decennial census operations will be collected using a specially designed hand-held computer (HHC). We will demonstrate some of the key features of the HHC such as payroll, viewing/updating the address list, and report capabilities.

Use of Global Positioning Receivers at the National Agricultural Statistics Service
Michael Gerling (National Agricultural Statistics Service, USA)

The National Agricultural Statistics Service has explored using handheld Global Positioning System (GPS) receivers in data collection.

One way that NASS has utilized GPS receivers is to obtain latitude/longitude coordinates for a particular agricultural field of interest instead of the labor intensive method of using county/highway maps and an in-house, developed mapping software.

A second use of GPS receivers is to record a particular point in a field where a field enumerator is required to return each month to record crop counts. In the past, the enumerator would tie red ribbon around stakes to mark the area of interest. However, sometimes the ribbon comes loose, stakes are knocked down or the enumerator simply is unable to find the designated area.

The enumerator would then use the GPS receiver to assist in finding the location. The demonstration will show how both utilizations of the GPS receivers saved staff time and improved data quality. A short review of cost and benefits will also be presented.

Development and Evaluation of an Audio Computer-Assisted Self-Interviewing System for Handheld Computing Devices
Kevin Wilson, Stephen Litavecz, and Norman Goco (RTI International, USA)

RTI has developed an audio computer-assisted self-interviewing (ACASI) system for handheld computing devices, a technology that allows questionnaires to be self-administered. Respondents view questions and responses on the screen while listening as each are read through headphones to ensure the privacy of the respondent. The system uses prerecorded sound (WAV) files for the audio component. When a response is selected, it is highlighted and read again so that responses are recorded accurately. The respondent selects the answer by touching the response using a touch screen. ACASI technology is especially effective in collecting sensitive respondent data (e.g., drug use, sexual practices, etc.) and with low-literacy respondents.

The system offers the ability to administer numerous survey questionnaires in multiple languages and provides skip logic, numeric range checks, and date checking. Standard question types include informational, select-one, select-all, numeric, date and text entry. In addition, new methods of data capture are supported including the ability to record spoken responses and collection of participants signatures, allowing for verification of informed consent. The system is well-suited for global clinical trials due to its support for multiple languages, subject randomization and data blinding techniques.

The system will save project resources by requiring fewer interviewers to conduct group surveys, improving response rates by offering greater mobility and convenience for data collection, decreasing time required of staff and participants due to its ease of use and quick setup, and increasing response accuracy with systematic error checking during data entry. Handheld devices may be carried discreetly, which makes them a viable alternative where security is a concern.

An initial evaluation of usability and performance has strongly supported the use of handheld devices for ACASI-based data collection, concluding that the handheld system can provide similar functionality to current laptop-based systems.

Large Scale Applied Time Series Analysis with Program TSW (TRAMO-SEATS for Windows)
Agustín Maravall (Bank of Spain, Spain)

The demonstration will center on the application of program TSW to a large set of monthly time series. TSW is a Windows interface of updated versions of programs TRAMO (Time series Regression withArima noise, Missing values, and Outliers) and SEATS (Signal Extraction in ARIMA Time Series).

The program estimates a general regression-ARIMA model, and computes forecasts and interpolators for possibly nonstationary series, with any sequence of missing observations, and in the presence of outliers. The

program contains an option for automatic model identification and automatic detection and correction of several types of outliers, and for pretesting and estimation of Calendar-based effects. Several types of intervention variables can also be included.

Next, the program estimates and forecasts the seasonal, trend, transitory, and noise components in the series, using signal extraction techniques applied to ARIMA models. The program contains a part on diagnosis and on inference, and an analysis of the properties of the estimators and of the estimation and forecasting errors. The last part of the output is oriented towards its use in short term economic policy and monitoring.

TRAMO contains an extension (TERROR, or Tramo for ERRORs) to the problem of quality control of data in large data bases of time series; SEATS can be applied for estimation of long-term trends and (business) cycles.

The programs can efficiently and reliably handle, in an entirely automatic manner, applications to sets of many thousand series. They are already being used intensively in research, data producing agencies, policy making institutions, and business. (Perhaps the most widely used application is Seasonal Adjustment.) They are freely available, together with documentation, at the Bank of Spain web site <www.bde.es>.

CONCURRENT SESSION V-A:
DISAGGREGATION OF ECONOMIC STATISTICS

Implementing a Reconciliation and Balancing Model in the U.S. Industry Accounts
Dylan Rassier, Thomas Howells III, Edward Morgan, Nicholas Empey, and Conrad Roesch (Bureau of Economic Analysis, USA)

Stone et al. (1942) advocate a generalized least squares model to improve the statistical accuracy of independent estimates of national income and expenditures based on relative reliabilities of different source data used to construct the estimates. Researchers have subsequently revised the Stone approach to facilitate its implementation (Byron, 1978; van der Ploeg, 1982, 1984), but federal agencies responsible for producing national economic accounts have generally not implemented the approach due to lack of available technology that is typically required to solve the complex systems of equations faced by federal agencies.

As part of the U.S. Bureau of Economic Analysis' (BEA's) integration initiative (Yuskavage, 2000; Moyer et al., 2004a, 2004b; Lawson et al., 2006), the BEA is drawing upon the Stone method and Chen (2006) to reconcile the gross operating surplus component of value added from the 2002 expenditure-based benchmark input-output accounts and the 2002 income-based gross domestic product (GDP)-by-industry accounts. The objective of the reconciliation is to use information regarding the relative reliabilities of underlying data in both the benchmark input-output use table and the GDP-by-industry accounts in a balanced input-output framework in order to improve intermediate input estimates and gross operating surplus estimates in both accounts.

This paper presents work at the BEA to develop and implement the reconciliation and balancing model. The paper provides overviews of the benchmark use table and GDP-by-industry accounts, including features of external source data and adjustment methodologies that are relevant for the reconciliation. In addition, the paper presents the empirical model and briefly describes the technology used to solve the model. Preliminary work during development of the model shows that reconciling and balancing a large system with disaggregated data is computationally feasible and efficient in pursuit of an economically accurate and reliable benchmark use table and GDP-by-industry accounts.

Estimating State Price Levels Using the Consumer Price Index
Bettina Aten (Bureau of Economic Analysis, USA)

Differences in relative price levels for 38 metropolitan and urban areas of the United States have been recently estimated by Aten (2006) using the Consumer Price Index (CPI) of the Bureau of Labor Statistics. This paper builds on this work and develops exploratory estimates of state price level differentials using the underlying price and expenditure data of the CPI and supplementary data from the 2000 Census.

Converting Historical Industry Time Series Data From SIC to NAICS
Robert Yuskavage (Bureau of Economic Analysis, USA)

The introduction of the North American Industry Classification System (NAICS) in the late 1990s offered the promise of more relevant industry time series data. NAICS improves on the Standard Industrial Classification (SIC) system because it more consistently classifies establishments into industries based on similar production processes and provides greater detail for the fast-growing services sector. Statistical agencies faced major challenges implementing NAICS for their current statistical programs. Possibly an even greater challenge, however,

was the conversion of historical industry time series data from SIC to NAICS, given the limited availability of source data classified on a NAICS basis before 1997.

Consistent industry time series data are crucial for studying industry contributions to economic growth, structural change, and productivity. One of the key datasets for studying these issues are the annual GDP by Industry Accounts prepared by the U.S. Bureau of Economic Analysis (BEA). These accounts provide time series data on industry output and employment for about 65 industries going back to 1947. This paper describes the data sources and methodology that were used to convert the published industry estimates for 1947 through 1997 from SIC to NAICS.

In the absence of historical source data classified on the new basis, conversion procedures tend to rely heavily on concordances developed for a single year. Such static concordances are reliable for a limited number of years before the reference year, but they become increasingly unreliable over time as relationships change among the industries. Concordances that capture changes over time in the relative importance of new industries yield more reliable results. This paper describes the procedures that were used to develop dynamic time-varying concordances for several industry variables. It also describes the special challenges that were involved in converting both current-dollar and real (inflation-adjusted) estimates.

An Empirical Comparison of Methods for Temporal Disaggregation at the National Accounts
Baoline Chen (Bureau of Economic Analysis, USA)

This study evaluates five mathematical and five statistical methods for temporal disaggregation in an attempt to select the most suitable method(s) for routine compilation of sub-annual estimates through temporal distribution and interpolation in the national accounts at BEA. In addition, three software programs for benchmarking and temporal disaggregation are also evaluated in terms of their capability of producing high quality sub-annual estimates and the ease with which they can be operated in routine compilation of sub-annual estimates. The evaluation is conducted using 60 series of annual data from the National Economic Accounts. The final sub-annual estimates are evaluated according to five specific criteria to ensure high quality final estimates that are also in compliance with operational policy at the national accounts. The study covers the cases of temporal disaggregation when 1) both annual and sub-annual information is available; 2) only annual data are available; 3) sub-annual estimates have both temporal and contemporaneous constraints; and 4) annual data contain negative values. The estimation results show that on average the modified Denton proportional first difference method outperforms the other methods, though the Casey-Trager growth preservation model is a close competitor in certain cases. Lagrange polynomial interpolation procedure currently being used at the national accounts is inferior to all other methods. The results also show that software program BENCH consistently generates final estimates with better quality and is the most flexible to operate.

CONCURRENT SESSION V-B:
VARIANCE ESTIMATION I

A New Application of Estimating Functions for Variance and Interval Estimation From Simple and Complex Surveys
Avinash Singh (Statistics Canada)

It is known that the normal approximation to the standardized estimating function (EF) (serving as a pivot for testing and interval estimation) holds well even for moderate sample sizes unlike the normal approximation for the estimator obtained by the delta or Taylor method. In this paper, we exploit the above stable approximation to the distribution of EF to proposed a unified solution to improve variance and interval estimation for seemingly different problems whenever the delta method doesn't perform well due to moderate sample sizes or is not applicable due to nonsmooth EF. There do exist several special methods in the literature that tend to improve the interval estimation in such problem cases but they, however, do not address the problem of variance estimation. The proposed method based on replications is related to the recently developed EF-Bootstrap method of Hu and Kalbfleisch (2000, Can.Jour. Stat.). It is shown that while the proposed method with suitable choices of EF gives results identical to the existing improved interval estimates, it also gives rise to stable variance estimates. Examples include Wilson's interval for estimating proportions, Fieller's interval for estimating ratio of means, Woodruff's interval for estimating quantiles, and again Fieller's interval for estimating mean response and dose levels in bioassay. The proposed method can be quite versatile as it can produce new variance and interval estimates for existing problems by using alternate EFs. Through simulations, it is shown that the proposed method has very favorable finite sample properties in comparison to other available meth-

ods in terms of bias and variance of the variance estimator, and average length of coverage probability of the confidence interval.

Weight Trimming Via Bayesian Variable Selection Method
Michael Elliott (University of Michigan, USA)

Standard approaches to weight trimming; reducing the value of large case weights to reduce variance in population parameter estimators— typically do not use the data to make bias-variance tradeoffs. This presentation develops Bayesian variable selection methods for weight trimming to supplement standard, ad-hoc design-based methods in disproportional probability-of-inclusion designs where variances due to sample weights exceed bias correction. Strata are formed according to the probability of the various models that separate or pooling juxtaposed inclusion probability strata is determined. The posterior predictive distribution of the population parameter of interest can then be obtained by averaging over the posterior probabilities that a given model is "correct." These methods are used to estimate linear and generalized linear regression model population parameters.

Using Markov Chain Monte Carlo for Modeling Correct Enumeration and Match Rate Variability
Andrew Keller (U.S. Census Bureau)

The Census Bureau conducted the Accuracy and Coverage Evaluation Revision II (A.C. E. Revision II) with the goal of producing improved estimates of the net coverage error of Census 2000. A.C.E. Revision II used dual system methodology to estimate the net coverage error. Dual system estimates were created for population subgroups called post-strata. A problem in model-based evaluation of coverage with respect to smaller geographies is that the variance between blocks and within those geographies needs to be specified in order to estimate a coverage correction factor. To improve coverage estimates, effort has been made towards advancing models involving smaller geographies. This paper offers Markov chain Monte Carol (MCMC) methods as a computer intensive method to generate these estimates. Specifically, random effects models of the correct enumeration and match rates at the block level are developed to specify this variance with an accompanying statement of precision.

A Study of Basic Calibration Estimators and Their Variance Estimators in Presence of Nonresponse
Yves Thibaudeau (U.S. Census Bureau), Jun Shao (University of Wisconsin–Madison, USA), and Jeri Mulrow (National Science Foundation, USA)

Sarndal et al. (1992) discuss calibration estimators for complex surveys. These estimators rely on information carried by survey covariates to reduce mean squared error. Sarndal and Lundstrom (2005) extend the concept of calibration estimators to cover survey-estimation situations in presence of non-response. We discuss one of the calibration estimators proposed by Sarndal, the one-way classification estimator, and its use in the context of the Research & Development Survey sponsored by the National Science Foundation. We recall the formula proposed by Sarndal et al. to estimate the variance of the one-way estimator, accounting for nonresponse. We compare Sarndal's variance estimation procedure and jackknife variance estimation. This customized jackknife espouses the concepts presented by Sarndal, but it does not involve a parametric interface to estimate the variance. We discuss the advantages and downsides of these two types of variance estimators in the context of calibration estimation in presence of nonresponse for the R&D survey.

CONCURRENT SESSION V-C:
ATTRITION

Evaluation of Models for Longitudinal Attrition Nonresponse
Eric Slud and Leroy Bailey (U.S. Census Bureau)

There is not much published methodological work on how to measure the biases of nonresponse adjustment from the internal evidence of a longitudinal survey. In connection with SIPP, Leroy Bailey has produced a series of Census reports on the topic culminating in an ASA paper of 2006, and Slud and Bailey (2006) studied the estimates and standard errors of biases between Wave 1 totals of various SIPP96 cross-sectional survey items and their estimates (called adjustment bias) based on nonresponse-adjusted totals of the same Wave 1 items using only response data from a later Wave (4 or 12). They concluded that the relative and standardized magnitudes of the estimated biases varied considerably and unpredictably from one adjustment model to another. The present research defines a composite metric of quality of a model used for nonresponse adjustment of a longitudinal survey. This metric combines the magnitudes of estimated between-wave adjustment biases based on subsets of the sample, relative to the estimated total, for various survey items. The metric is based on the

largest of the adjustment biases for estimated totals calculated from the first K sample units, as K ranges from 1 to the size of the entire (wave-1) sample. This maximal absolute bias is averaged over a number of random re-orderings either of the whole sample or of the units within specified cells (which are then also randomly re-ordered), and the average maximal adjustment bias is divided by the estimated wave-1 item total to give the metric value. Theoretical (weak-convergence) results are used to estimate confidence bands for the estimated metric. The metric is then applied to judge the quality of a collection of adjustment-cell and logistic-regression models for nonresponse adjustment in SIPP96.

The Effect of Attrition on the NLSY97

Alison Aughinbaugh (Bureau of Labor Statistics, USA) and Rosella Gardecki (The Ohio State University Center for Human Resource Research, USA)

The National Longitudinal Survey of Youth 1997 (NLSY97) is a national sample of about 9,000 youth who were ages twelve to sixteen on December 31, 1996, and living in the U.S. Starting in 1997, interviews have been conducted annually. Currently, data through Round 8 are available, at that time respondents are ages 20 to 24. Though the focus of the NLSY97 is employment, the data set covers a broad array of topics including schooling, training, marriage, fertility, and income, thus allowing one to examine how different areas of life are interrelated.

The usefulness of any data set depends on whether the data are representative of the population of interest. Maintaining the representativeness of a longitudinal data set is crucial because these data gain value as information is collected from the same units over time, giving researchers the ability to examine whether and how early events are tied to later outcomes.

Although unit non-response has always been a concern for longitudinal surveys, over the past fifteen years the level of attrition has increased for longitudinal data sets. Atrostic et al. (2001) study attrition in six U.S. government household longitudinal surveys and find that the rate of unit non-response increased in all six over the 1990s. This pattern emerges with the National Longitudinal Surveys as well. Whereas the response rate in the NLSY97 was 89.9 percent in round 4, in its predecessor survey, the NLSY79, the response rate did not fall to such a level until round 12.

Recent examinations of longstanding U.S. panel data sets find that attritors differ significantly from non-attritors, but that the effects of this non-random attrition do not impact the pictures of the labor market outcomes presented by the data. In this study, we will examine the effects of attrition on the NLSY97 using the methods that MaCurdy, Mroz,

First Cut Is the Deepest

James Halse, Iain Noble, and Andrew Ledger (Department for Education and Skills, United Kingdom)

Response bias poses special problems for survey researchers. Significant efforts are put into detecting non-ignorable biases and devising weighting schema to counteract them. One advantage for longitudinal surveys is that data from earlier waves can be used in this. In later waves this often shows the biasing effects of attrition as relatively small.

One question that remains, however, is about differences between non-response at the first Wave and subsequent non-response and the possible inability of reweighing using only survey data to compensate. Non-response, especially refusal, can be substantially greater at first waves than subsequently. Several hypotheses exist about the possible effect of this, among them: that there is a quantitative difference between early and later non response, with larger, undetected, biases in the earlier, this also means that biases occurring at early Waves may not recur at subsequent ones. One result of either of these being true is: corrective weighting based solely on measurements made after Wave 1 may not correct for all non-ignorable biases. A second: general conclusions drawn about survey non-response using evidence from Wave 2 and subsequently of panel studies, may be invalid.

Until recently such hypotheses would be difficult to test as most samples have relatively restricted information pre-interview about those drawn. The increasing availability of individual level administrative data, and the use of such data sets as sampling frames, changes this. Such data sources can provide rich background information on sample members drawn from them.

This paper uses data from UK administrative sources and further data from linkages to test these hypotheses on results from the first three waves of two large scale studies of young people. It then assesses the relative value of sample data against survey data in detecting and compensating for bias in these two studies.

Attrition Bias in Panel Estimates of the Characteristics of Program Beneficiaries

John Czajka, James Mabli, and Karen Cunnyngham (Mathematica Policy Research, Inc., USA)

Attrition is a serious and growing problem in panel surveys, and it raises concerns about the continuing representativeness of panel samples after the initial interview. In the nine-wave 2001 panel of the Survey of Income and Program Participation (SIPP), only 65 percent of the original sample (excluding those dropped because of a sample reduction) qualified for a full panel weight, meaning that they would be available for a longitudinal analysis spanning the full length of the panel (or any portion thereof). A similar nine-wave calculation for the 1996 panel (which ran 12 waves) yields nearly the same figure. These estimates compare to 76 percent for the 1993 panel, which included the same number of waves as the 2001 panel.

This paper uses administrative records linked to survey records from the 1996 and 2001 panels of the SIPP to evaluate the extent of attrition bias among a sample of Social Security and Supplemental Security Income (SSI) beneficiaries. Estimates of bias are presented for both cross-sectional and longitudinal measures, building on earlier studies that have shown greater bias in estimates of change than in characteristics measured near the end of the panel. Social Security beneficiaries (who include retired and disabled workers as well as survivors and dependent children) have lower attrition rates than the general survey population, whereas SSI recipients compare to the broader population. In our investigation of attrition bias we look for evidence of life events that may have contributed to attrition while also contributing to changes in characteristics, creating differences between attriters and non-attriters. In presenting our conclusions we propose ways in which the survey data might be adjusted to compensate for attrition bias and improve the representativeness of the data for both cross-sectional and longitudinal analysis.

CONCURRENT SESSION VI-A:

AMERICAN COMMUNITY SURVEY

Improving the Labor Force Questions in the American Community Survey: The Results of the 2006 ACS Content Test

David Raglin and Kelly Holder (U.S. Census Bureau)

The unemployment rate from the American Community Survey (ACS) has tended to be higher than the unemployment rate from the Current Population Survey (CPS), the official source of the unemployment rate. We tested a modified version of the employment series as part of the 2006 ACS Content Test, which tested new and revised content for possible inclusion into the ACS. The paper examines the changes that were made to the labor force battery of questions and compares the modified version of the questions to the existing ones. Data from the current and revised questions were also compared to data collected in a content follow-up from survey respondents using the CPS questions.

Analysis of Changes to the Educational Attainment Question in the 2006 ACS Content Test

Alan Peterson and Sarah Crissey (U.S. Census Bureau)

In 2006 the American Community Survey (ACS) conducted the first test of new and modified content since the ACS reached full implementation levels of data collection. The results of that testing, the 2006 ACS Content Test, helped determine the content for the 2008 ACS. This paper examines the alternative question tested for educational attainment intended to improve reporting of specific levels of attainment as well as provide more detail for the secondary education grades. The alternative question (test version) included changes such as a use of headers to group schooling levels, a write-in category for grades 1 through 11, and a change in wording for the post-high school categories. Discussion will focus on the impact each of these changes had on response, including how mode of implementation may influence the effect.

A Comparison of Forced-Choice and Mark-All-That-Apply Formats for Gathering Information on Health Insurance in the 2006 American Community Survey Content Test

Leah Ericson (Carnegie Mellon University, USA) and Chuck Nelson (U.S. Census Bureau)

The 2006 ACS Content Test included two different versions of a new health insurance question for testing. The health insurance item intends to estimate health insurance coverage as well as private versus public health care coverage. Version 1 of the question used a forced-choice format. The forced-choice format requires respondents to indicate yes or no for each of seven health care types. Version 2 of the question used a mark-all-that-apply format. In this format, respondents read or hear a list of seven types of health care coverage, selecting as many as apply to them. The results suggest that the forced-choice version of the question produced more dual coverage estimates than the mark-all-that-apply version. In addition, when comparing the responses from the two formats to responses given in a more detailed follow-up interview the forced-choice version resulted in less

systematic error than the mark-all-that-apply version. These empirical findings will be discussed in terms of the methodological differences between each version.

A Comparison of Closed- and Open-Ended Question Formats for Select Housing Characteristics in the 2006 American Community Survey Content Test

John Chesnut, Ellen Wilson, and Jeanne Woodward (U.S. Census Bureau)

In January through March 2006, the U.S. Census Bureau conducted the first test of new and modified content for the American Community Survey since the survey reached full implementation levels of data collection. The results of that testing will help determine the content for the 2008 American Community Survey. One of the research objectives of this test was to conduct an experimental study of the impact of using open-ended question formats compared to closed-ended question formats for three different housing questions B property value, number of vehicles kept by members of the household, and number of rooms and bedrooms in the household. The three questions require varying amounts of knowledge to respond and thus may differ in how respondents use an open- versus closed-ended response format. For each of the three items, this paper examines the differences in data quality resulting from the two different response formats in terms of item nonresponse, reliability, and systematic response error.

CONCURRENT SESSION VI-B:

MONITORING, MEASURING, AND ADJUSTING FOR NON-RESPONSE II

Using the Multi-Level Integrated Database Approach

Tom W. Smith (National Opinion Research Center, USA)

We live in an information age in which considerable data exist in many forms and at many levels (e.g., individuals, households, neighborhoods, communities, etc.). These data sources can be tapped to aid in the collection, adjustment, and analysis of surveys. As Stoop has observed, by linking sampled cases to a large number of other registers and administrative records, a large among of data is available on nonrespondents and respondents with which nonresponse can be analyzed. The multi-level, integrated database approach (MIDA) can help to deal with the problems of both low response rates and the detection and adjustment for non-response bias. In addition, MIDA will also further substantive analysis by providing aggregate-level information for contextual analysis. Thus, MIDA will both advance survey-research methodology as well as enhance substantive research. MIDA uses databases to collect as much information as practical about the target sample at both the case-level and at various aggregate levels during the initial sampling stage. The first step in MIDA is to extract all relevant, public information at both the case-level and aggregate levels from the sampling from which the sample addresses are drawn. The second step is to augment the sampling frame by linking all cases in the sample to other databases at both case and aggregate levels. The third step in MIDA is to take information gained from the initial case-level linkages to secure additional information. The final stepis to record, process, clean, and maintain a large amount of paradata for each case.

Nonresponse Bias Patterns in the Current Population Survey

John Dixon (Bureau of Labor Statistics, USA)

Nonresponse rates have been used as a proxy for survey quality since they indicate the relative potential for nonresponse bias. The patterns of nonresponse rates (e.g., seasonal, time in sample) can provide insight into those rates. The current study uses different measures of nonresponse bias to see if there are similar patterns for bias which might be different than the patterns for rates.

Sample Maintenance: Internet Use by a Low-Income Population

Bryan Rhodes, Ellen Marks, and Jun Liu (RTI International, USA)

The ability to locate and re-interview baseline survey participants at the time of the follow-up survey is critical in longitudinal surveys. Tracking participants between rounds of data collection minimizes the number of respondents lost to follow-up and decreases the potential for non-response bias. Researchers have employed interim tracking activities that allow respondents to update locating information in a variety of ways including through the mail, over the telephone, and, increasingly, via the Internet. Researchers must balance the costs of each method with the expected outcome. Since lower-income populations are less likely to have internet access,

the use of the web for interim tracking may be seen as less cost effective for studies of these populations. The work presented here examines the interim tracking of a survey of lower-income parents concerning savings attitudes and behaviors. The study allows respondents to update locating information (address and telephone number) though the mail, over the telephone, or via a website. During the two years since the baseline survey, four rounds of sample maintenance have been conducted. Fourteen percent of the survey respondents have used the website to update their contact information. This paper uses baseline telephone survey data to determine characteristics of respondents who have used the web option. It examines differences between web responders and those who have only used mail or telephone methods based on a variety of demographic factors including race, ethnicity, age, family composition, education, and employment status. The analysis will model changes in web responders across rounds of sample maintenance. Preliminary analysis shows significant differences between web responders and mail or telephone responders based on several factors including education and income. These differences show the importance of a web option for retaining a representative sample at follow-up.

CONCURRENT SESSION VI-C:

ADMINISTRATIVE RECORDS: APPLICATIONS OF DATA LINKAGES

RELAIS: Don't Get Lost in a Record Linkage Project

Tiziana Tuoto, Nicoletta Cibella, Marco Fortini, Monica Scannapieco, and Laura Tosco (Italian National Institute of Statistics (ISTAT), Italy)

The combined use of statistical survey and administrative data is largely widespread: joint analyses of statistical and administrative sources allow to save time and money, reducing survey costs, response burden, etc. However, data sources are sometimes hard to combine since errors or lacking information in the record identifiers may complicate the integration. The record linkage techniques are a multidisciplinary set of methods and practices whose purpose is to identify the same real word entity, which can be differently represented in data sources. Record linkage has been subject of research for several decades: a huge number of record linkage solutions have been proposed, based on probabilistic and empirical paradigms. Being record linkage a complex process, it can be decomposed in separate phases, each requiring a specific technique and depending on given application and data requirements. Due to such complexity and application dependency, we propose a toolkit for record linkage, called RELAIS (Record Linkage At IStat). The toolkit is based on the idea of choosing the most appropriate technique for each phase, and of dynamically combining such techniques in order to build a record linkage workflow, on the basis of application constraints and data features provided as input. The modular structure of the workflow allows to omit and/or iterate some phases of the record linkage process. Moreover, the RELAIS toolkit is configured as an open source project, which allows to the users community to: gather the most appropriate techniques from the efforts of several independent experts. A real case study, based on the Post-Enumeration Survey of the XIV Italian Population Census, validates the RELAIS idea and provides a methodological pattern for driving the design of a record linkage workflow on the basis of the requirements of a real application.

Allocated Values in Linked Files

Amy O'Hara (U.S. Census Bureau)

This paper discusses the incidence and handling of allocated values when using linked data sets. Administrative records from other agencies are combined with U.S. Census Bureau survey data using protected identification keys. The linked files are useful for modeling and evaluation. This paper discusses a merged Current Population Survey Annual Social and Economic Supplement (CPS ASEC) and Internal Revenue Service Individual Master File (IRS IMF) data set. This file is analyzed at the individual income tax return level to evaluate modeled filing status and adjusted gross income, both important inputs to the tax calculator used to generate liability and credit amounts. The evaluation is complicated when the survey responses are incomplete. Demographic characteristics and income fields may be missing. CPS ASEC processing employs a hot deck imputation to fill these gaps. This paper introduces the issue of how to use these cases in evaluations. Income is not imputed to withstand scrutiny at the person level, conflicting with the primary benefit of administrative record linkage. Given the extent of income item non-response in the CPS ASEC, omitting all allocated records diminishes the usefulness of the linked file and renders person weights useless. By comparing reported income on the survey to reported income from the administrative record, a tolerance can be determined that can be applied to the allocated amounts. In this fashion, a proportion of the records with imputed income may be used in evaluations. Separate tolerances are designed to account for differences in filing status, worker class, and income types present. Cases with imputed amounts that meet the tolerance criteria are used in model evaluation; these results are compared with an evaluation where records with imputed income are omitted. The impact on disclosure and weighting when using either method is discussed.

The Use of Free School Meal Status as a Proxy for Socio-Economic Status: Evidence From Matching the Longitudinal Study of Young People in England to the National Pupil Database

James Halse and Andrew Ledger (Department for Education and Skills, United Kingdom)

Socio-economic status (SES) is seen as a key determinant of children's educational achievement, yet in the UK we know relatively little about the SES of pupils. Children living in households below a certain income threshold are eligible for Free School Meals (FSM). FSM eligibility is recorded on administrative databases and is used in much educational research as a proxy for SES. Overall, children who are eligible for FSM do less well academically at school, yet we do not really know who the FSM children are, and what their families are like.

Data from a major new face to face survey of young people and their parents/guardians, the Longitudinal Study of Young People in England (LSYPE), has been linked to the National Pupil Database (NPD). For the first time, we can now understand what the FSM pupils and their families are really like. We investigate how closely FSM status is associated with true measures of SES collected through LSYPE such as parental social class, income and education, as well as looking at the attitudes and aspirations of children and their parents by FSM status.

Finally, we evaluate the use of FSM in the Department for Education and Skills (DfES) contextual value added models. DfES use multiple regression to predict what level a pupil is expected to attain at age 16 given a certain level of prior attainment at age 11, along with contextual information such as ethnicity, gender and FSM status. Schools are evaluated according to their ability to add value to their pupils' expected progress. By replacing FSM in the model with a true measure of SES, we demonstrate that other coefficients in the model are biased, in some cases severely, by the use of FSM as a proxy for SES.

CONCURRENT SESSION VII-A:

MAKING MUSIC OUT OF ALL THAT NOISE: USING ADMINISTRATIVE RECORDS AND SURVEY DATA IN HARMONY

How Do Surveys Differ in Reporting the Quality of Reported Medicaid Enrollment Data: CPS and State Surveys

Michael Davern and Kathleen Call (University of Minnesota, USA)

Category administrative records session Description Survey measurement error can contribute to the tendency for surveys to show lower counts of Medicaid enrollment than administrative data (i.e., Medicaid Undercount). This happens two ways: first, respondents can misreport that they are uninsured when they really have Medicaid, biasing survey estimates of the uninsured and those on Medicaid. Second, respondents can answer they have coverage other than Medicaid, biasing the Medicaid count and not the uninsured count. The first form of error is more detrimental to policy research than the second. We compare the prevalence of both error scenarios using two types of surveys in which we know the Medicaid enrollment status of respondents: four state surveys, the Current Population Survey, and several state surveys. Preliminary analysis shows that rates of the first type of error are much higher in the CPS than other surveys.

Differences in Estimates of Public Assistance Recipiency Between Surveys and Administrative Records

Victoria Lynch and Dean Resnick (U.S. Census Bureau) and Jane Staveley and Cynthia Taeuber (Jacob France Institute, USA)

This study results from a collaboration between the Jacob France Institute, the Maryland Department of Human Resources, and the U.S. Census Bureau. In the first phase of the study, we examine discrepancies between the published estimate of public assistance households in Maryland from the U.S. Census Bureau's American Community Survey/Supplementary Survey for 2001 (SS01), and the count of public assistance households from the Client Automated Resource and Eligibility System (CARES), the administrative database for Maryland's public assistance programs. Records from these two sources are matched at the individual level and the results are used to generate match results at the household level to answer the following questions. What is the discrepancy between the two data sources when the administrative count is adjusted to cover the same universe as the ACS/SS01 published estimate? What is the discrepancy attributable to survey error? What is the discrepancy attributable to error from respondents giving false-negative reports?

In the second phase of the study we test hypotheses about the relationship between household characteristics and the probability that respondents give false-negative reports about participation in the TANF program, the main type of public assistance. Using the subset of TANF households from the research file created in phase one, we address the following questions. What household characteristics relate to false-negative reporting? What is the relative risk of false-negative reporting associated with particular household characteristics?

Developing the Chapin Hall Child Care Subsidy Eligibility Model

Dean Resnick (U.S. Census Bureau)

For program policy analysis, it is often useful to have data not only on program participants but also on persons, households, or families that are eligible under program rules. This information is useful for analyzing take-up rates and decisions, as well as making inferences as to the effects of the program. Unfortunately, program administrative records data usually covers only enrollees or at most enrollees and applicants. To evaluate the Child Care Subsidy program we sought to develop a model that imputed eligibility status onto households captured in the ACS/SS01. We built this model by statistically modeling family structure and family income and comparing output from these models as well as additional family status information to known eligibility rules. This eligibility model is now being used to understand CCS take-up and understand how this program impacts participants.

Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Survey and SSA Administrative Data

Martha Stinson (U.S. Census Bureau)

We quantify sources of variation in annual job earnings data collected by the Survey of Income and Program Participation (SIPP) to determine how much of the variation is the result of measurement error. Jobs reported in the SIPP are linked to jobs reported in a new administrative database, the Detailed Earnings Records (DER) drawn from the Social Security Administration's Master Earnings File, a universe file of all earnings reported on W-2 tax forms. As a result of the match, each job potentially has two earnings observations per year: survey and administrative. Unlike previous validation studies, both of these earnings measures are viewed as noisy measures of some underlying true amount of annual earnings. While the existence of survey error resulting from respondent mistakes or misinterpretation is widely accepted, the idea that administrative data is also error-prone is new. Possible sources of employer reporting error, employee under-reporting of compensation such as tips, and general differences between how earnings may be reported on tax forms and in surveys, necessitates the discarding of the assumption that administrative data is a "true" measure of the quantity collected by the survey. In addition, errors in matching SIPP and DER jobs, a necessary task in any use of administrative data, also contribute to measurement error in both earnings variables. Exploiting the presence of individuals with multiple jobs and shared employers over time, we estimate an econometric model that includes random person and firm effects as well as a common error component shared by SIPP and SSA earnings. We do not impose ancillary orthogonal design assumptions. Hence, our model is more general than conventional fixed effect estimators for this problem. The estimation equation includes two independent error components that represent the variation unique to each earnings measure. All fixed effects, random person and firm effects estimated jointly from the two data sources, and the shared residual are interpreted as components of "true" variation that represent differences in earnings across people, firms and time periods due to underlying economic reasons. The independent error components are interpreted as measurement error. Using these estimates we calculate several different measures of "true earnings" that take account of our measurement error estimates. We compare SIPP, DER, and true earnings for different demographic and education groups of SIPP respondents. We also calculate different measures of changes in earnings for welfare respondents moving on and off public assistance and changes in earnings for individuals switching jobs. Finally we estimate a standard earnings equation model using SIPP, DER, and true earnings and compare the resulting coefficients. While there are some interesting differences across the three types of earnings measures, in general the results are more similar than not. We conclude that at least for individuals whose survey-reported job information is of high enough quality to enable a match to administrative records, the earnings measures in the SIPP are as reliable as those in the administrative data.

CONCURRENT SESSION VII-B:

ESTIMATION ISSUES

Imbedding Model-Assisted Estimation Into ACS: The Impact on Users

Robert Fay, III (U.S. Census Bureau)

The American Community Survey (ACS) began full implementation in 2005 as a replacement for the decennial census long form. In 2010, ACS estimates will be released for the 5-year period 2005–2009; this release will be the first to offer ACS results at the geographic detail previously provided by Census 2000. A model-assisted approach has been proposed to reduce the variances of estimates for geographic units below the county level for both 3- and 5-year period estimates. The approach has been investigated as part of the Multi-Year Estimates Study, based on an ACS test in 34 counties during 1999–2005. The anticipated variance reductions were empirically confirmed, although substantial reductions occurred for a few key variables and far more modest ones

for others. Given the variance impact, the primary focus of this paper will be to address some of the remaining concerns potential ACS users may have about the model-assisted methods, such as their impact on bias and whether their introduction complicates analysis.

A Larger Sample Size Is Not Always Better

Nagaraj Neerchal (University of Maryland, Baltimore County, USA) and Herbert Lacayo and Barry Nussbaum (Environmental Protection Agency, USA)

In a previous paper Neerchal, Lacayo and Nussbaum (2007) explored the behavior of the well known problem of finding the optimal sample size for obtaining a confidence interval of a pre-assigned precision (or length) for the proportion parameter of a finite or infinite binary population. We illustrated some special problems that arise due to the discreteness of the population and precision being measured by the length of the interval rather than by the variance. Specifically, the confidence level of an interval of fixed length does not necessarily increase as you increase the sample size. However, if the confidence levels are computed using normal approximation, the non-monotonic behavior is not observed. In this paper we consider the corresponding problem under the Poisson approximation and show that the non-monotonicity persists and one should be beware of this peculiar behavior in recommending sample sizes for studies involving estimation of p-proportions.

Alternative Tests of Independence

Jai Choi (National Center for Health Statistics, USA) and Balgobin Nandram (Worcester Polytechnic Institute, USA)

The chi-squared test for independence in two-way categorical tables depends on the assumptions that the data follow the multinomial distribution. Thus, we suggest alternatives when the assumptions of multinomial distribution do not hold. First, we consider the Bayes factor which is used for hypothesis testing in Bayesian statistics. Unfortunately, this has the problem that it is sensitive to the choice of prior distributions. We note here that the intrinsic Bayes factor is not appropriate because the prior distributions under consideration are all proper. Thus, we propose using Bayesian estimation which is generally not as sensitive to prior specifications as the Bayes factor. Our approach is to construct a 95% simultaneous credible region (i.e., a hyper-rectangle) for the interactions. A test that all interactions are zero is equivalent to a test of independence in two-way categorical tables. Thus, a 95% simultaneous credible region of the interactions provides a test of independence by inversion.

CONCURRENT SESSION VII-C:

STATISTICAL METHODS APPLIED TO RACE DESIGNATIONS AND POPULATION ESTIMATES

Bridging Estimates by Race for the Current Population Survey

William Davis and Anne Hartman (National Cancer Institute, USA) and James Gibson (Information Management Services, Inc., USA)

Due to a directive from the Office of Management and Budget (OMB), the Current Population Survey (CPS) changed its race/ethnicity questions in January 2003. Respondents may now select more than one race when answering the survey. This paper provides a method to construct single race estimates using data from the post-2003 CPS Tobacco Use Supplement (TUS-CPS). The method is useful when trends over time are being examined for single race groups using both pre-2003 and post-2003. The method uses the post-2003 race/ethnicity responses to multiply impute the (unknown) pre-2003 race/ethnicity response. The method is particularly useful for racial groups whose respondents often report multiple races. This is the case for two races that are often underserved; the American Indian or Alaska Native (AIAN) and the Native Hawaiian or Other Pacific Islander (NHOPI). The method is used to show the recent trend in current smoking by race/ethnicity and gender. Computer code is included to allow the user the ability to apply the method to determine other smoking trends using TUS-CPS.

Statistical Methods for Analyzing Multiple Race Response Data

Tommi Gaines (University of California, USA)

Collection of racial data is ubiquitous throughout research as an important measure of the demographic characteristics of the study population. However, the validity of racial data has been a concern, prompting several agencies to modify their measurements by allowing individuals to identify with multiple racial categories. These

revisions pose analytic dilemmas for the researcher by inhibiting compatibility between different data collection systems, presenting difficulty with monitoring trends over time, and possibly leading to insufficient sample sizes to generate statistically reliable estimates. This research aims to add to the current methodology for analyzing data collected from multiracial individuals by comparing different statistical methods for analyzing multiple race responses as well as single race categories for data generated from the California Health Interview Survey (CHIS). Three distinct methods are explored for analyzing outcomes that indicate whether individual health behaviors are consistent with goals of the Healthy People 2010 program. One approach uses supplementary data from the Census Bureau and the California Department of Finance to rake multiple-race respondents into single-race categories consistent with the 1977 OMB standards. The second method, following Schenker and Parker (2003), imputes a single race category for multiple race respondents to produce population health estimates. The third method, which we call multiple covariate adjustment, simultaneously controls for indicators of all self-identified race categories (using one group as the referent) in a regression analysis. The sensitivity and robustness of the three methods are checked by fitting models to simulated populations, developed from the CHIS data set, that allow the proportion of multiple-race responders in the population to vary. Attention will be focused on inference for the proportion of individuals who meet Healthy People 2010 goals, which has a common interpretation across methods, as well as inference about racial disparities in achieving those goals.

Genetic Analysis of Population Structure Relative to Self-Reported Race and Ethnicity in NHANES III
Christopher Sanders, Ajay Yesupriya, and Lester Curtain (National Center for Health Statistics, USA)

Race and ethnicity variables in the Third National Health and Nutrition Examination Survey (NHANES III) are primarily self-reported with a small number from interviewer observations. The NHANES III survey, conducted prior to the OMB directive on collecting multiple race information, only allowed the participant to report one race/ethnicity. This could lead to an oversimplified representation of a participant's race/ethnicity. Another method to examine race/ethnicity is population structure analysis using genetic variations to infer the probabilities that an individual is of one or more given race/ethnicities. Since many statistical analyses use genetic data stratified by race/ethnicity, it would be valuable to assess the correlation of self-reported race/ethnicity relative to inferred population structure. We used meticulous genetic analysis to infer population structure and compared it to self-reported and observed NHANES III race and ethnicity variables. Also, we analyzed the extent of race/ethnicity admixture in the NHANES III population. We performed these analyses using the program Structure 2.1 on genetic data from more than 7000 individuals. Population structure results were analyzed for each participant relative to NHANES self-reported and observed race/ethnicity variables. These analyses showed a high level of correlation between self-reported race/ethnicity and the participant's predicted population structure. Population structure analysis of admixture can also be used to estimate the proportion in the sample that might have reported multiple race if that category had been available. We also used population structure results to predict which race/ethnicity category would be most appropriate for participants that had been placed in the Other category of NHANES III. Population structure analysis is a useful tool for estimating population ancestry, assessing of mixed race admixture and in some cases allowing the categorization of individuals with missing race/ethnicity data.

CONCURRENT SESSION VIII-A:
BRIDGING THE PAST WITH THE FUTURE: INNOVATIONS AND SURVEY MEASUREMENT ISSUES

Formulating the Laws of Studying Societal Change
Tom W. Smith (National Opinion Research Center, USA)

First Law: The way to measure change is not to change the measure. Second Law: When constant measures produce non-constant measurement, change the measure to measure change. Using extensive examples from the Census, General Social Survey, and other time series, this paper discusses the above laws of studying societal change. It considers 1) why it is more difficult to follow these simple laws than first appears, 2) when the second law should be followed rather than the first law, and the 3) serious consequences of failing to follow these laws.

Analytical Comparison of the SIPP and CPS-ASEC Key Longitudinal Estimates
Smanchai Sae-Ung, C. Dennis Sissel, and Tracy Mattingly (U.S. Census Bureau)

The longitudinal survey, Survey of Income and Program Participation (SIPP) and the cross-sectional survey, the Current Population Survey Annual Social and Economic Supplement (CPS-ASEC) have identical survey universe and sampling frame, and are similar in demographic and socioeconomic contests, sample and questionnaire designs, and data collection and processing systems. However, the SIPP collects monthly data but the CPS-

ASEC collects annual data. The SIPP sample households are generally interviewed for nine to twelve times over three to four years but the CPS-ASEC ones are only interviewed for two consecutive Marches. As a longitudinal survey, the SIPP follows the movers and classifies survey universe leavers (e.g., dead); but, as a cross-sectional survey, the CPS-ASEC does not. Inspired by similarity between these two surveys, the low non-response and attrition rates in and establishment stature of the CPS-ASEC, we develop an approach to analytically compare the SIPP's key longitudinal estimates with those of the COP-ASEC as a means to assess the longitudinal data quality of the SIPP. In this approach, we construct a quasi-longitudinal dataset from two consecutive CPS-ASEC cross-sectional datasets, and develop a probabilistic model to simulatively categorize the sample people in the first of the two consecutive CPS-ASEC's who became a mover or a universe leaver in the second CPS-ASEC. The simulative categorization of the mixed cohort of movers and survey universe leavers synthetically transforms the quasi-longitudinal dataset to be a complete longitudinal dataset. This coupled with longitudinally weighting the CPS March ASEC quasi-longitudinal dataset using the SIPP longitudinal weighting procedure enables us to produce the longitudinal annual estimates for analytical comparison with those produced by the SIPP. We apply this approach to the comparable 2002 and 2003 CPS-ASECs and 2001 SIPP Panel. This approach is also directly applicable to the upcoming reengineered SIPP, the Dynamics of Economic Well-being System (DEWS).

UK Household Surveys: Building on Survey Integration

Roeland Beerten (Office for National Statistics, United Kingdom)

In recent years the UK has witnessed an increase in evidence-based policy making, which has led to an increased need for survey information in a number of policy areas. Sometimes these needs can be met through minor modifications to the design of existing surveys, or new surveys can be set up in order to collect the necessary information. However, both approaches are generally expensive and sometimes inefficient. In trying to overcome these problems the UK Office for National Statistics (ONS) has adopted a more pro-active approach to survey taking in which flexibility and integration are key principles. This presentation will outline how the household survey strategy in ONS has taken on board a number of new survey requirements based on the principle of integration rather than designing new surveys in isolation. It will give an overview of the new surveys we have set up in the past few years and how ONS has used principles of survey integration in order to realize statistical and efficiency benefits. One of the key developments in this survey strategy is the Integrated Household Survey, which aims to fully implement the integration principle by offering a flexible, modular survey system.

Protocol Calibration in the National Resources Inventory

Cindy Yu and Jason Legg (Center for Survey Statistics and Methodology, USA)

The National Resources Inventory (NRI) is a large-scale longitudinal survey conducted by the National Resource Conservation Service in cooperation with the Center for Survey Statistics and Methodology (CSSM) at Iowa State University. The primary sampling units are areas of land called segments. About 300,000 segments were observed in 1997. Subsamples of the 300,000 segments of sizes 40,000 to 70,000 segments have been observed yearly since 2000. A key NRI estimate is year-to-year change in acres of developed land, where developed land includes roads and urban areas. Prior to 2004, acres of developed land in a segment were delineated by data gatherers using a transparent overlay and an aerial photograph. The data gatherers were also responsible for measuring the delineated area. In 2004, a fully digital procedure was introduced. Certain types of developed land such as industrial areas, roads, and large structures are delineated much as before, but boundaries are entered digitally. In the remaining area, the data gatherer marks residences as points. A computer program has been developed by CSSM to convert the points to area of developed land in each segment. Two independent determinations of the new procedure plus the determination by the old procedure were made on each segment for a sample of 2700 segments. A measurement error model is postulated for the relationship between old and new measures, where the duplicate measurements are used to estimate the error variances. Preliminary analyses suggest that the measurement error variances are a function of the true values. Possible nonlinearities between the two measures are studied by binning the data and by fitting measurement error models to subsets of the data. The results are used to modify the new protocol so that estimates are due to real change and not protocol change.

CONCURRENT SESSION VIII-B:
DATA QUALITY

Quality Assessment of the Linkage Between the Canadian Community Health Survey and Hospital Data
Michelle Rotermann (Statistics Canada)

Background/ Rationale: The Longitudinal Health and Administrative Data (LHAD) Initiative is a project between the federal and provincial Ministries of Health and Statistics Canada. One objective of the Initiative is to address information gaps through research using hospital in-patient data (HPOI) linked to population health survey data. An assessment of the quality of the linkage is an important preliminary step. Methods: To assess the quality of the linkage between cycle 1.1 of the Canadian Community Heath Survey (CCHS) and HPOI, the number of individuals admitted to hospital if their health insurance number (HIN) matched a HIN from at least on HPOI record; survey weights were applied to these records to calculate the total number of individuals hospitalized. Because HPOI is a virtual census of all hospital admissions, this count was regarded as the standard. The difference in counts between the two data sources indicated the extent of linkage failure. Quebec residents were excluded because the HINs in HPOI are scrambled. Results: According to HPOI for September 2000 - November 2001, 1,627,576 Canadians aged 12+ were hospitalized. The weighted estimate from the CCHS was 27% lower (1,192,920). The rate of linkage was comparable for men and women; lower among persons living in Manitoba (51%) or Nunavut (35%); better among individuals aged 12-74 (77%) than among those aged 75+ (61%). Conclusion: The number of Canadians (outside Quebec) aged 12+ who were hospitalized, based on hospital records, was 27% higher than the estimated number based on CCHS records linked to hospital records. The discrepancy varied by some socio-demographic characteristics; thus use of the linked file could give rise to biased estimates depending on the sub-population studied.

Do Teenagers Always Tell the Truth?
Janet Rosenbaum (Harvard University, USA)

Adolescents do not always answer survey questions truthfully. Researchers cannot detect lies directly, but they can detect logical contradiction, such as when an adolescent gives different answers to the same question in a short time frame. This study identifies question properties associated with higher inconsistency and estimates the increased error due to inconsistency. The data are from a reliability study of the Youth Risk Behavior Survey (YRBS), in which a diverse convenience sample of 4628 high school students answered the same 72 questions at an interval of two weeks. In the two-week interval, the prevalence of 41 of the 72 risk behaviors changed significantly. Questions about substance use are more consistent than questions about other topics, possibly because substance use questions are unambiguous. Longer questions and questions at the end of the 90 item survey are less consistent. A Bayesian method estimated that the standard error of the prevalence of 72 risk behaviors is at median 3.3 times larger than conventional estimates of standard error (95% CI (1.9, 5.0)). The majority of questions on the YRBS yield unreliable estimates of risk behavior prevalence, but questions which cannot be interpreted in multiple ways may be more reliable than ambiguous questions. Conventional measures of uncertainty underestimate the error on the prevalence of adolescent risk behaviors.

The Accuracy of Reported Insurance Status in the MEPS
Steven Hill (Agency for Healthcare Research and Quality, USA)

Four sources of validation data were used to assess the quality of respondents' reports of private insurance and uninsurance in the 1996 Medical Expenditure Panel Survey Household Component (MEPS). For the insured, the MEPS asked for insurance cards, which provide physical evidence of coverage. Most people reporting private insurance were also asked for policy booklets, which are also evidence of coverage. Employers and insurance companies identified by the households were asked whether individuals in the household had insurance from the establishment. Medical providers identified by households were surveyed about sources of payments, especially insurance, for health services received by MEPS sample members. Three analyses were conducted. The first examined whether private insurance was overreported, using three types of validation data (insurance cards, policy booklets, and employer and insurance companies) to compensate for weaknesses in any one source. Private insurance is the major source of insurance for the nonelderly but has been little studied. The second and third analyses examined whether insurance was underreported using information from employers and providers. Only some sample members had data from employers and providers, so differences in socioeconomic characteristics and survey compliance between the analytic and full samples were analyzed. While members of the analytic samples were more compliant, noncompliance was not associated with disagreements between the household and employer or provider respondents. Probit and complementary log-log regression analysis suggests recall errors account for some underreporting among the uninsured, low-income respondents overreport private insurance but more accurately report uninsurance, and a few respondents report only dental insurance when they have physician/hospital insurance. Using the regression coefficients to project the rates of

disagreement to the full sample suggests that 1.7% of the reportedly privately insured likely did not have it and 8.3% of the reportedly uninsured were likely privately insured.

Quality of Income Data in Household Surveys: Lessons From a Comparative Analysis
Gabrielle Denmead (Denmead Services, USA), John Czajka and Robert Weathers (Mathematica Policy Research, Inc., USA), and Joan Turek (U.S. Department of Health and Human Services)

Most major household surveys sponsored by the Federal government include measures of income, because income is a critical classification variable for policy-related analyses. This paper presents some of the major findings to date from a project for the office of the Assistant Secretary for Planning and Evaluation, Department of Health and Human Services, that is conducting a comprehensive and systematic examination of income data across eight federally sponsored surveys: the Annual Social and Economic Supplement to the Current Population Survey, the Survey of Income and Program Participation, the American Community Survey, the National Health Interview Survey, the Medical Expenditure Panel Survey Household Component, the Medicare Current Beneficiary Survey, the Health and Retirement Study, and the Panel Study of Income Dynamics.

Estimates of income and other policy-related measures for calendar year 2002 are made using the same definitions, the same time periods, the same units of analysis and the same income measures, to the extent possible, for each survey. Estimates cover a range of demographic groups and all levels of the income distribution, and are supplemented by ad hoc tabulations and simulations to measure various attributes of individual surveys. The project is also collecting detailed information on overall survey design, methodology, timing, questionnaire content, and data processing procedures that that may contribute to differences in the estimates. Findings offer a number of lessons about the measurement of income in household surveys.

CONCURRENT SESSION IX-A:
CHALLENGES AND STRATEGIES IN QUESTIONNAIRE DESIGN

Effects of Language and Culture on Interpretation of Translated "Confidentiality" and "Mandatory" Survey Messages
Yuling Pan and Ashley Landreth (U.S. Census Bureau), Marjorie Hinsdale-Shouse and Hyunjoo Park (RTI International, USA), and Alisú Schoua-Glusberg (Research Support Services, USA)

Obtaining informed consent involves educating potential respondents about a survey effort so they may make an informed decision regarding their participation. Topics addressed in the consent materials are shaped by ethical standards associated with human subjects research (National Research Council, 2003) and sometimes mandated by federal law (for research funded in whole or part by federal monies). These messages typically include the survey's purpose, confidentiality assurances, and notification of whether the survey is voluntary or mandatory. Effectively conveying these messages in English can be challenging. Having to translate the same messages into other languages further complicates the endeavor. Yet, to carry out ethical research and ensure data quality, it is critical for non-English speakers to understand these messages in translated survey materials (e.g., survey letter and informational brochures). Concerted efforts to create and pretest meaningful translations were pursued with survey materials for the U.S. Census Bureau's American Community Survey. Materials translated into Chinese, Korean, Russian, and Spanish were cognitively pre-tested with monolingual respondents in these languages. This paper describes the cause for respondents problematic interpretations of two important informed consent messages: confidentiality and the mandatory nature of the survey. Both are prominent within the survey letter and both refer to federal law. The mandatory message states that the law requires a response to the survey, and the confidentiality message states that survey responses are protected by law. Interpretation difficulties were associated with literal translation practices or by the presence/lack of cultural context for these concepts. This paper documents the unique and shared communicative difficulties observed across cultural groups. We anticipate building on the analysis method outlined here to examine the full suite of informed consent messages embedded in the materials that were cognitively pre-tested for this study.

Asking for Numbers and Quantities - The Design of Answer Boxes to Frequency Questions and Its Impact on Data Quality
Marek Fuchs (University of Kassel, Germany)

In almost every self-administered survey respondents are asked to report numbers, frequencies or quantities. Since closed ended questions with response categories are prone to scale effects, in many surveys open questions are used to collect this kind of data. Usually, the size and shape of the answer box as well as the position of associated labels are not considered as relevant aspects of a survey that might alter the responses collected.

In this paper results form several field experiments embedded in paper & pencil surveys as well as in Web surveys are reported. The experimental conditions vary the size and the appearance of the boxes. Also, use of labels has experimentally been varied.

Results indicate that the design of the answer boxes to open questions has a significant impact on the answers provided by respondents. Large boxes yield more detailed information; however, once the answer space becomes unreasonable large the proportion of overly detailed responses, explicit estimates and unrelated information increases as well. By contrast, small response boxes yield more hyping or bunching. Also, a label associated to the response box increases the accuracy of the response. The findings will be discussed in the light of visual design effects.

From Start to Pilot: A Multi-Method Approach to the Comprehensive Redesign of an Economic Survey Questionnaire

Alfred Tuttle, Rebecca Morrison, and Diane Willimack (U.S. Census Bureau)

In 2004, the Census Bureau began work on a long-term project with the Bureau of Economic Analysis (BEA), to aid in the redesign of two of their surveys collecting foreign direct investment (FDI) data from U.S. companies with foreign owners, one annual and the other quarterly. The surveys were often burdensome for many respondents, and data verification was labor-intensive for BEA analysts. Some of the problems with data quality stemmed from difficulties comprehending the FDI concept and other particulars of the surveys, which differ subtly but significantly from the ways respondents think about their companies and keep their records. In order to maximize the effectiveness of our redesign, we undertook a comprehensive process using multiple research methods to gain a thorough understanding of the conceptual underpinnings of the surveys and the problems respondents encountered when completing them. Our initial steps involved conducting focus group and individual interviews with BEA analysts, observations of telephone calls to respondents, and debriefings with recent respondents. We parlayed the knowledge we gained into a model for redesigning the forms based on best practices in questionnaire design, along with some innovations in graphical design and placement of instructions. We evaluated our design recommendations in an intensive period of iterative cognitive testing. Results were used to design a complete form, which was fielded in a pilot test with a sub-sample of respondents. We conducted debriefings with respondents who completed the pilot, and conducted statistical and qualitative analyses comparing edit failures from the pilot form with the legacy form. This paper will describe the various questionnaire testing and evaluation methods used at each stage of the project, their complementary nature, and major results.

Background and Planning for Incorporating an Event History Calendar Into the Re-Engineered SIPP

Jason Fields (U.S. Census Bureau) and Mario Callegaro (Knowledge Networks, USA)

The Census Bureau is currently re-engineering the Survey of Income and Program Participation (SIPP). The re-engineered SIPP will be an integrated data system with multiple components and methods: a survey-based data collection; an American Community Survey-based sampling frame; administrative records; and statistical modeling. Together, these components will yield a fuller, stronger, more richly detailed program of estimates of economic well-being. The re-engineered SIPP survey component will be evaluated to incorporate improvements to the current SIPP. The survey's mission is to provide a nationally representative sample for evaluating: 1) annual and sub-annual income dynamics, 2) movements into and out of government transfer programs, 3) family and social context of individuals and households, and 4) interactions between these items. These survey objectives are pursued in the context of several goals - cost reduction and improved accuracy, relevance, timeliness, and accessibility. We present background information about the methodological choices that led to the current design. The re-engineered SIPP survey is being designed around annual data collection that collects sub-annual dynamics in income and government program use. Annual survey administration is the primary source of significant cost savings compared with prior SIPP panels. Moving from the current design, with interviewing 3 times per year, to an annual survey, raises significant concerns about recall accuracy and the ability for respondents to report the sub-annual dynamics SIPP traditionally measured. The current preferred option is to design a new survey based on more streamlined content, minimal topical module-based content, and to utilize an Event History Calendar (EHC) module to improve respondents' ability to recall the sub-annual dynamics the survey is tasked with collecting. We discuss the background for choosing an EHC module, additional design considerations, and propose methodological evaluations that may be incorporated into strategies for evaluating EHC collected data.

CONCURRENT SESSION IX-B:
DISCLOSURE II

Recent Developments in the use of Noise for Protecting Magnitude Data Tables: Balancing to Improve Data Quality and Rounding That Preserves Protection
Paul Massell and Jeremy Funk (U.S. Census Bureau)

Over the years, many program managers in the economic directorate of the U.S. Census Bureau have been interested in exploring alternatives to cell suppression for protecting released tables, i.e., protecting the confidentiality of microdata that underlie the tables. The goal is to produce tables with fewer suppressions that are still fully protected. With fewer suppressions, the tables would likely be of greater value to users, even if the cell values were perturbed a bit from their original values. A decade ago, a method was developed by researchers at the Census Bureau Evans-Zayatz-Slanta (J. Official Statistics,1998) that involves adding noise to the microdata. Recently this method has been applied to survey and census tables with distinctive features. In some programs there may be a table that is considered the primary table, and one may wish to fine-tune the noise method, using a technique called 'noise balancing', so that the data quality is as high as possible for this table while maintaining the quality in other tables. There are also differences in the type of rounding that is applied to the raw data as it is transformed into microdata and perhaps later when microdata values are summed to form cell values. Rounding often adds uncertainty about the pre-rounded value, so it would seem to increase disclosure protection. However, when noisy values are rounded, this sequence has the potential of reducing the protection level of the noise. To ensure that this does not happen requires enhancement of some of the ordinary rounding methods.

Model Based Disclosure Avoidance for Data on Veterans
Sam Hawala and Jeremy Funk (U.S. Census Bureau)

The U.S. Department of Veterans Affairs (VA) routinely requests from the U.S. Census Bureau special tabulations on veterans to study socio-economic characteristics, demographic characteristics and the geographic distribution of the veteran population. Due to confidentiality concerns the Census Bureau may apply a number of different disclosure avoidance techniques to protect the data collected through the variety of its surveys. Techniques include rounding, top-coding, data swapping, imposing geographic thresholds, introducing random noise, cell suppression and complementary cell suppression. In the case of VA data the techniques we use induce enough loss of detail in the data necessitating some counter-measures, such as de-rounding. VA employs de-rounding to adapt the tabulations they obtain from the Bureau to their needs. The confidentiality edits also result in discrepancies between the estimates obtained from the special tabulations and what the Census Bureau publicizes on the web. We examine the use of partially synthetic data and to a limited extent other disclosure avoidance methods to produce a veteran microdata file that is disclosure proof, from which VA will be able to obtain all cross tabulations and run analyses, without the need for de-rounding. The estimates from the file will be compared to other Census Bureau benchmarks. The file will contain several versions of the microdata for some variables and some records in order to assess the uncertainty of the estimates using multiple imputation techniques.

Microdata Risk Assessment in an NSI Context
Jane Longhurst and Paul Vickers (Office for National Statistics, United Kingdom)

Dissemination of microdata is becoming more common in the UK. ONS has developed a range of criteria (statistical, logistical and strategic) and licence arrangements (end user licence and special licence) to decide how much detail can be released to its users under strict conditions. This paper discusses the statistical risk assessment that underpins these decisions and the statistical methods supporting this assessment. In particular, it will focus on estimating risk using a probabilistic disclosure risk method using the Poisson distribution and Log linear models as outlined by Elamir and Skinner. The application of this method on a range of key variables taken from a variety of different social surveys carried out by a National Statistical Institute will be covered as well as the practical issues related to implementation within a microdata release strategy. This statistical analysis will be relevant to any statistical organisation contemplating disseminating microdata for the purposes of policy and research. It will have applications for a range of data sources from social surveys. Elamir, E. and Skinner, C. (forthcoming) Record-Level Measures of Disclosure Risk for Survey Microdata, Journal of Official Statistics.

CONCURRENT SESSION IX-C:
SAMPLE DESIGN

Properties of Alternative Sample Designs and Estimation Methods for the Consumer Expenditure Surveys
John Eltinge (Bureau of Labor Statistics, USA)

The U.S. Consumer Expenditure Survey (CE) is a large-scale household survey conducted to collect information on expenditures at a relatively fine level of classification. Data currently are collected through quarterly interviews (the Consumer Expenditure Diary Survey). Due to interest in reduction of respondent burden, improvement of data quality, or possible reduction in the cost of data collection, the Bureau of Labor Statistics is carrying out some preliminary exploration of alternatives to the current designs of the diary and interview surveys. These alternatives include, but are not necessarily limited to: administration of subsets of the current instrument to sample consumer units selected for the interview survey; use of global expenditure questions in place of more detailed expenditure questions for some sections of the current instrument, for some sample consumer units selected for the interview survey; changes in the periodicity, number of reference periods for the diary or interview surveys. If implemented, any of these alternatives would require substantial changes in CE sample design and estimation methods. This paper explores the mathematical properties of some of the resulting survey procedures, with primary emphasis on the following topics. (1) Possible use of adaptive, two-phase or matrix-sampling methods to assign global questions or instrument sections to sample consumer units. (2) Modification of selection probabilities to increase the estimation efficiency for specified expenditure means. (3) Modification of estimation weights and variance estimation procedures to account for (1) and (2). (4) Trade-offs in estimation efficiency for, respectively, simple quantities (e.g., expenditure means) and more complex parameters (e.g., the coefficients of a generalized linear model). Discussion of (1)–(4) will include mathematical characterization of practical operational constraints and risks; and the role of cost information in the optimization of alternative survey procedures.

The American Community Survey Sample Design: An Experimental Springboard
Megha Joshipura and Steven Hefter (U.S. Census Bureau)

In an ongoing effort to improve the quality of the American Community Survey (ACS) estimates by reducing non-sampling error, the Census Bureau routinely solicits ideas for various projects including pre-testing and field testing of proposed new content, and revisions to the current set of ACS questions. In 2007, two experiments with different objectives and data collection strategies will test two questionnaire formats, grid and sequential, and a new Field of Degree question proposed by the National Science Foundation (NSF), respectively. The requirements of each experiment necessitate unique sample designs. The Grid/Sequential Format Test is mail-out only. The sample was selected using constant overall sampling rates for each treatment. The NSF Content Test will collect data via mail-out, Computer Assisted Telephone Interview, and Computer Assisted Personal Interview in sequential months. This sample selection employed several sampling rates, which are proportional to the 2007 ACS probabilities of selection. Each sample consists of 30,000 addresses selected in pairs. Of particular interest is the adjustment of the selection probabilities to account for previous samples, including the production ACS sample, drawn from the same frame at different points in time. This paper discusses the unique opportunities provided by the ACS production sample design and selection for experimental sample designs and sampling frames to be derived from it. The added benefit of using the ACS sample, given the readily available production response data, which can be used as experimental controls, is also highlighted. For each experimental design we focus on the statistical requirements, including reliability calculations, the selection probability derivations accounting for the differential, two-phase sample design of the ACS, and provide an accounting of the samples by experimental panel.

An Adaptive Sample Allocation for a Multiple Objectives Survey of Business
Daniela Golinelli, Gregory Ridgeway, and John Adams (RAND Corporation, USA)

Researchers often have multiple objectives that they wish to accomplish with one survey. Classic sample allocation techniques, however, often focus on optimizing the sample allocation to address a single objective. In 2006 RAND fielded the National Computer Security Survey to provide estimates of computer incidents by industry, by company size, and by whether they occurred at a critical infrastructure company. The optimal sample allocation to maximize precision across industries would not give adequate precision across company size categories. Therefore, we developed an adaptive sampling algorithm that iteratively allocates sampling effort to company strata so that the effective sample sizes across all estimates of interest are fairly homogeneous. We present the algorithm and demonstrate that if the interest is in being able to make inference at the industry level and/or at company size level, this allocation tends to provide more homogenous effective sample sizes

and precisions across industries and company size classes than, for example, an allocation that is proportional to the stratum's size. Therefore, this allocation represents the right compromise when the analyst has multiple goals with a given survey. Other advantages of this allocation are that the number of observations allocated to every stratum is an integer number and, therefore, a sample size of exactly n units can be allocated. Lastly, the allocation algorithm can easily be modified to give, for example, more weight to certain industries or company size classes if higher precision is desired for those analyses.

CONCURRENT SESSION X-A:
WEB APPLICATIONS

Developments in Electronic Survey Design for Establishment Surveys
Grace O'Neill (U.S. Census Bureau)

As electronic-based surveys grow in popularity, many agencies have developed their own sets of electronic form design practices. Some of these practices developed from research while other practices developed from field-based experience. This paper acts to synthesize a variety of electronic form design practices used by the establishment surveys at the U.S. Census Bureau. These practices are based on research and findings from cognitive and usability testing. Because these findings evolved out of respondent testing, the Census Bureau has been able to develop electronic survey design features for establishment surveys that integrate well with the activities and tasks respondents perform in to complete surveys. This provides for a more respondent-centered electronic survey design. In this paper, we discuss topics such as working with internal Information Technology staff, addressing the visual differences between paper and electronic forms, respondents use of paper associated with electronic response, and the use of real-time edits and other electronic-only features. The paper will also discuss the potential implications of the Census Bureau's findings for both household and establishment-based surveys, and outline areas that need more research.

Increasing Response Rates: Pre-Notification
Dominic Lusinchi (Farwest Research Statistical Consulting, USA)

Although web-based survey research has serious limitations (e.g., non-coverage, lack of a sampling plan), it has been embraced by many who do survey research. It has been shown to be particularly useful when the population surveyed is almost universally connected to the Internet, and is proficient with the medium and the tools to use the medium (computing devices: e.g., desk top computers, p.d.a.'s, etc.). In this experiment conducted on a panel of electronics engineers and engineering managers, we explore issues that may affect response rates, and response quality. Respondents are asked to report on the process of designing an integrated circuit. First, we examine the effect of a time-honored method on both unit response rates and speed of returns: one half of the panel received a notification in advance of the survey invitation, letting potential respondents know that they would be asked in a few days to participate; the other half did not. Second, we analyze what effect two different formats of the same instrument have on response quality. One half of the panel was randomly assigned to a scroll-type web questionnaire, in which the instrument is loaded onto one web page and the respondent scrolls down from one question to the next; the other half of the panel received a page-type questionnaire, in which the respondents click on a button to move from one question to the next. Is one type more advantageous than the other? Measures of quality used will be the break-off rate (the proportion of respondents that abandon the survey) and item response rate (looking both at errors of omission, failing to answer an appropriate question; and errors of commission, answering a question forwhich the respondent is not eligible).

Enhancing Web-Based Data Collection Using Excel Spreadsheets
Dan Jackson and Michele Eickman (Bureau of Labor Statistics, USA)

The Current Employment Statistics (CES) Survey, conducted by the Bureau of Labor Statistics (BLS), ratio adjusts (benchmarks) its sample-based employment estimates to independent population controls at various levels of geography and industry details. For the most part, these independent population controls are obtained from the Bureau's Quarterly Census of Employment and Wages (QCEW) program. QCEW covers all employees subject to Unemployment Insurance (UI) laws. Part of the benchmarking process is the collection of presumed non-covered (PNC) employees; defined as workers not subject to UI laws or the Unemployment Compensation for Federal Employees (UCFE) program and therefore are not included in employment counts from the BLS QCEW program. Examples of PNCs are railroad employees, insurance carriers, and clergy.

PNC employment data obtained from the Census Bureau's County Business Patterns are supplemented by information provided by State partners. In the past, States submitted PNC data to BLS on paper forms. Early

versions of a Web-based form proved to be burdensome if States chose to submit PNC data at finer levels. As a result, many States only reported at the statewide level; the minimum level of detail that BLS would accept. To reduce burden and increase the level of detail reported by the States, BLS enhanced the functionality of the system using Excel spreadsheets. States can now download a template spreadsheet containing initial PNCs, copy and paste data to the spreadsheet from different sources to update it for the current year, and upload the revised spreadsheet to the web system. The spreadsheet can be uploaded and downloaded multiple times, and the database is overwritten with each upload. The addition of the spreadsheet has allowed BLS to capture more monthly PNC data than in previous years, increasing the quality of model-based estimates for selected State and area series.

CONCURRENT SESSION X-B:
VARIANCE ESTIMATION II

On X11 Seasonal Adjustment and Estimation of Its Variance
Michail Sverchkov and Stuart Scott (Bureau of Labor Statistics, USA) and Danny Pfeffermann (The Hebrew University of Jerusalem, Israel, and University of Southampton, United Kingdom)

Most official seasonal adjustments use the (nonparametric) X-11 method and its extensions, available for instance in the Census Bureau'roximation to X-11 and accounts for the effect of sampling error. Extending this approach are the methods of Pfeffermann (1994) and Bell & Kramer (1999). In this talk we show that under mild assumptions that can be tested from the observed data the seasonal and the trend components can be defined in such a way that (1) X11 estimates of these components are almost unbiased and (2) the Pfeffermann (1994) method gives unbiased estimates of the variance of X11 estimates. The results are illustrated in a simulation study.

Diagnostic Process to Assess the Effects of Truncating Extreme BRFSS Sampling Weights
Henry Roberts, Elizabeth Hughes, and Ruth Jiles (Centers for Disease Control and Prevention, USA) and Robert Woldman (North Carolina Department of Health and Human Services, USA)

This research focuses on developing and implementing a diagnostic process to study the effects of truncating extreme BRFSS sampling weights; and the extent that extreme sampling weights affect the robustness and precision of BRFSS prevalence estimates.

The 2004 North Carolina BRFSS database consisting of 15,052 completed interviews were analyzed to assess the effects of truncating BRFSS sampling weights. This diagnostic process sequentially assesses the: variability, robustness, and precision of BRFSS prevalence estimates. The truncation of extreme sampling weights resulted in reducing the variance of the pre-truncation sampling weights for Hispanics and adults ages 18-34 by 298% and 181%, respectively. Estimates that were robust or non-robust prior to the truncation of extreme sampling weights remained robust or non-robust. The largest decrease (-23%) in the standard errors were observed for the estimated prevalence of adult males having access to healthcare. The impact of truncating BRFSS sampling weights was: no overall effect on the robustness of BRFSS prevalence estimates; large decreases in standard errors; and large deceases in the variability of the sampling weights.

An Examination of Alternative Variance Estimators
Laura Ozcoskun and Samson Adeshiyan (U.S. Census Bureau)

The Medical Expenditure Panel Survey (MEPS) survey is a stratified one-stage sample design that employs an iterative multi-stage weighting procedure that accounts first for unit non-response before post-stratifying to outside control totals. This weighting procedure is both time-consuming and resource intensive, with twenty-one separate iterations. Currently, variance estimates for MEPS are produced using the method of random groups and are constructed from the final adjusted weights, thus leading to biased variance estimates. Full replication of the weighting procedure is statistically preferable, but difficult to implement. Moreover, while theoretically pleasing, the random group variance estimator is known to have high variance in without-replacement samples; replicate estimates constructed from more sampled units such as the delete-a-group or stratified jackknife have been repeatedly shown to have better properties for a highly stratified design. Through empirical calculation and a simulation study, we address two concerns for this survey: the decision to fully or partially replicate a weighting procedure and the appropriate choice of variance estimator.

CONCURRENT SESSION X-C:
IMPUTATION

Multiple Imputation and Estimating Aggregate Productivity Growth in Manufacturing
Kirk White (U.S. Census Bureau), Amil Petrin (University of Chicago, USA), and Jerome Reiter (Duke University, USA)

Tim Dunne (1998) and others have noted that the quality of imputed data is a problem for researchers using the plant-level Annual Survey of Manufactures and Census of Manufacturers data available at the Census Bureau's Center for Economic Studies. Futhermore, missing data became a more important problem for some types of analysis in the 1990s when the ASM stopped collecting capital asset variables in non-Census years. We apply methods of multiple imputation (Rubin, 2004) to replace the missing and imputed data in the manufacturing datasets. The 20002 CMF data have detailed item impute flags. Using the detailed impute flags in these years we develop a model of missing data in the CMF and ASM. We use sequential multivariate regression imputation (Raghunathan et al., 2001), a technique that has been used for imputation in the National Health Interview Survey and in imputation models for data confidentiality in several Census RDC projects. This technique, coupled with advances in imputation based on semi-parametric models and predictive mean matching facilitates imputation of multivariate, possibly non-normal data. We apply these methods to the U.S. manufacturing data and then estimate and decompose aggregate productivity from the plant-level data.

Multiple Imputation of Right-Censored Wages in the German IAB Employment Register Considering Heteroscedasticity
Thomas Buettner and Susanne Raessler (Institute for Employment Research, Germany)

In many large data sets of economic interest, some variables, as wages, are top-coded or right-censored. In order to analyze wages with the German IAB employment register we first have to solve the problem of censored wages at the upper limit of the social security system. We treat this problem as a missing data problem and use multiple imputation approaches to impute the censored wages by draws of a random variable from a truncated distribution, based on Markov chain Monte Carlo technique. In general, the dispersion of income is smaller in lower wage categories than in higher categories and the assumption of homoscedasticity in an imputation model is highly questionable. Therefore, we suggest a new multiple imputation method, which does not presume homoscedasticity of the residuals. Finally, in a simulation study different imputation approaches are compared under different situations and the necessity as well as the validity of the new approach is confirmed.

Imputing Missing Values in the Common Core of Data for Use in Computing the Averaged Freshman Graduation Rate
Jack Buckley, Marilyn McMillen Seastrom, and Chris Chapman (National Center for Education Statistics, USA)

During the recent debate about the estimation of dropout and/or on-time graduation rates, NCES proposed the Averaged Freshman Graduation Rate (AFGR) as a potential solution. The AFGR is conceptually simple, can be computed at the school district level from the existing repeated cross-sectional population data in the Common Core of Data (CCD), and appears to be more accurate than many, if not all, competing rates. However, the AFGR requires enrollment counts in the eight, ninth, and tenth grades, as well as the diploma count, for a given cohort, and some of these quantities are missing in the CCD for many school districts, particularly once stratified by gender or race/ethnicity. This paper explores several ways of imputing these missing values. We focus in particular on a hybrid approach and demonstrate its effectiveness via simulation.

FCSM RESEARCH CONFERENCE AUTHOR AND COAUTHOR INDEX

NAME	SESSION
Megha Joshipura	IX-C
Sarah Kalsbeek	Technical Demo
Mallika Kasturirangan	IV-A
Andrew Keller	V-B
Carol King	II-C
Chris Krebs	IV-B
Jon Krosnick	Keynote Speaker
Herbert Lacayo	VII-B
Ashley Landreth	IX-A
Andrew Ledger	V-C, VI-C
Jason Legg	VIII-A
Kerry Levin	II-A
Michael Link	II-C, III-C
Stephen Litavecz	Technical Demo
Jun Liu	VI-B
Xijian Liu	II-C
Jane Longhurst	IX-B
Martin Luppes	IV-A
Victoria Lynch	VII-A
Dominic Lusinchi	X-A
James Mabli	V-C
Agustin Maravall	Technical Demo
Ellen Marks	VI-B
Katherine Mason	Technical Demo
Paul Massell	IX-B
Tracy Mattingly	VIII-A
Tucker McElroy	II-B
Loretta McKenzie	III-B
Jatinder Mehta	IV-C
Ana Militino	IV-C
Ali Mokdad	II-C, III-C
Brain Monsell	II-B
Edward Morgan	V-A
Rebecca Morrison	IX-A
Jeri Mulrow	V-B
Krish Muralidhar	III-A
Shadana Myers	II-C
Balgobin Nandram	VII-B
Nagaraj Neerchal	VII-B
Charles Nelson	VI-A
Justin Nguyen	IV-A
Thuy Trang Nguyen	II-C
Iain Noble	V-C
Barry Nussbaum	VII-B
Amy O'Hara	VI-C
Broderick Oliver	III-B
Grace O'Neill	X-A

NAME	SESSION
Larry Osborn	II-C, III-C
Laura Ozcoskun	III-B, X-B
Dick Paddock	Technical Demo
Yuling Pan	IX-A
Swamy Paravastu	IV-C
Hyunjoo Park	IX-A
Reginald Pendergraph	Technical Demo
Alan Peterson	VI-A
Amil Petrin	X-C
Danny Pfeffermann	X-B
Polly Phipps	II-A
Susanne Raessler	III-A, X-C
David Raglin	VI-A
Dylan Rassier	V-A
Jerome Reiter	X-C
Dean Resnick	VII-A
Bryan Rhodes	VI-B
Gregory Ridgeway	IX-C
Mary-Helen Risler	II-A
Henry Roberts	X-B
Conrad Roesch	V-A
Janet Rosenbaum	VIII-B
Michelle Rotermann	VIII-B
Smanchai Sae-Ung	VIII-A
Christopher Sanders	VII-C
Cathlin Sarafin	II-A
Rathindra Sarathy	III-A
Monica Scannapieco	VI-C
Stuart Scott	II-B, X-B
Alisú Schoua-Glusberg	IX-A
Marilyn McMillen Seastrom	X-C
Jun Shao	V-B
Pedro Silva	III-B
Michelle Simard	II-A
Catherine Simile	III-C
Avinash Singh	V-B
C.Dennis Sissel	VIII-A
Eric Slud	V-C
Helen Smith	Technical Demo
Tom W. Smith	VI-B, VIII-A
Jane Staveley	VII-A
Martha Stinson	VII-A
Barbara Stussman	III-C
Michail Sverchkov	X-B
Cynthia Taeuber	VII-A
Yves Thibaudeau	V-B
Steven Thomas	II-A

NAME	SESSION
Katherine Jenny Thompson	III-B, X-B
Vanessa Thornburg	Technical Demo
Richard Tiller	II-B
Laura Tosco	VI-C
Bac Tran	IV-A
Clyde Tucker	II-A
Tiziana Tuoto	VI-C
Joan Turek	VIII-B
Alfred Tuttle	IX-A
Maria Ugarte	IV-C
Paul Vickers	IX-B
Robert Weathers	VIII-B
Kirk White	X-C
Diane Willimack	IX-A
Ellen Wilson	VI-A
Kevin Wilson	Technical Demo
Robert Woldman	X-B
Jeanne Woodward	VI-A
Ajay Yesupriya	VII-C
Cindy Yu	VIII-A
Robert Yuskavage	V-A
Tamara Zimmerman	IV-A, IV-C

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

Copies of these working papers may be ordered from NTIS Document Sales, 5285 Port Royal Road, Springfield, VA 22161; telephone: 1-800-553-6847. The Statistical Policy Working Paper series is also available electronically from FCSM's Web site <http://www.fcsm.gov>.

NOTES

NOTES

NOTES

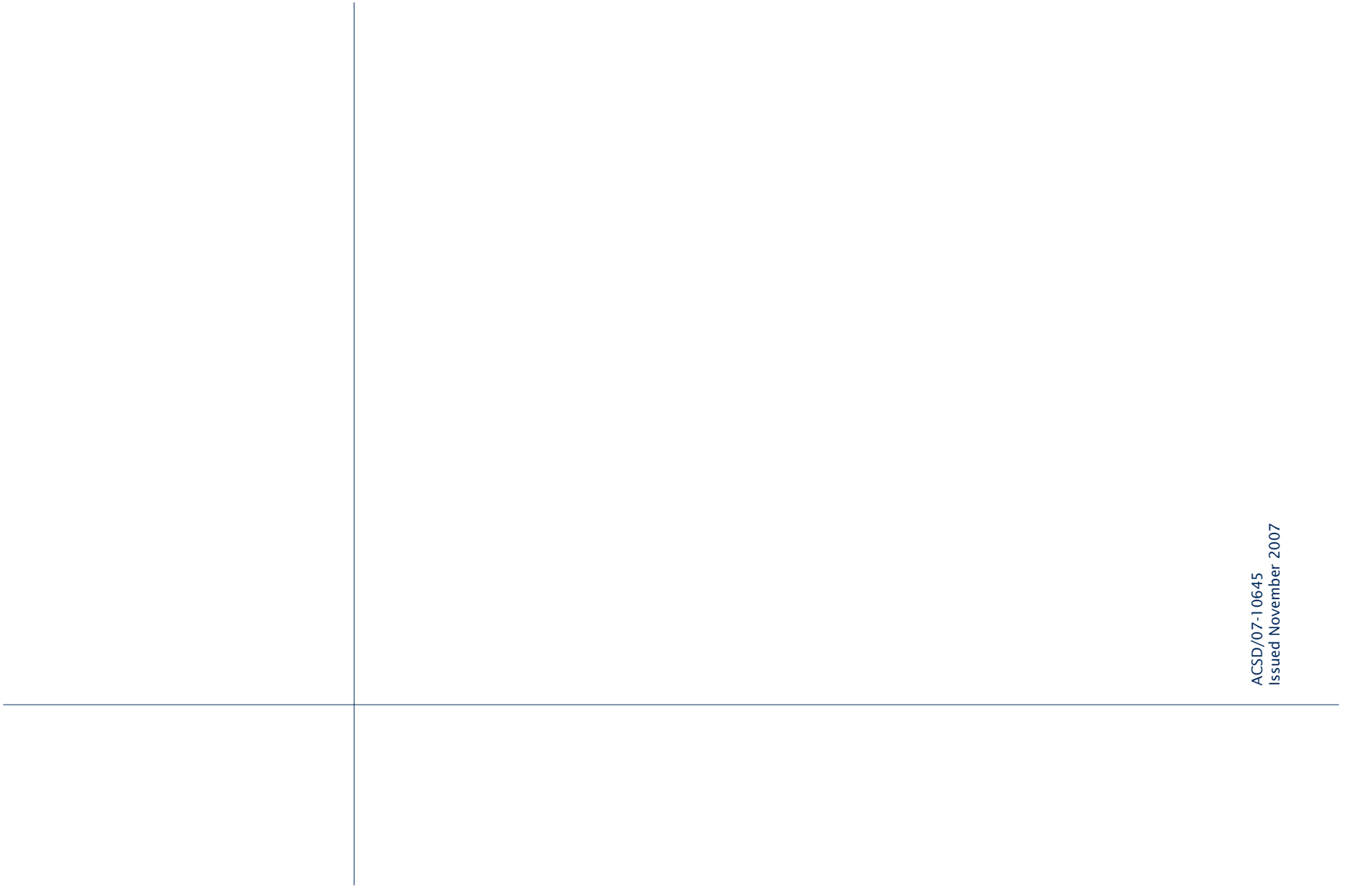
NOTES

NOTES

NOTES

NOTES

NOTES



ACSD/07-10645
Issued November 2007