Federal Committee on Statistical Methodology 2021 Research and Policy Conference

November 2-4, 2021

Full Program

Updated: Nov 5, 2021

Sponsored By:

Hosted By:

Federal Sponsors

The 2021 FCSM Research and Policy Conference received generous support from the following federal agencies.

FCSM 2021 Research and Policy Conference Program Committee

Darius Singpurwalla, Co-Chair, *National Center for Science and Engineering Statistics*
Jessica Graber, Co-Chair, *National Center of Health Statistics*
Ellen Galantucci, Co-Chair, *Bureau of Labor Statistics*
David Kashihara, Co-Chair, *Agency for Healthcare Research and Quality*
Paul Schroeder, Co-Chair, *Council of Professional Associations on Federal Statistics*
Maura Bardos, *U.S. Energy Information Administration*
Scott Boggess, *U.S. Census Bureau*
Yang Cheng, *National Agricultural Statistics Service*
Stephanie Coffey, *U.S. Census Bureau*
William Hahn, *Economic Research Service*
Diba Kahn, *National Center of Health Statistics*
Richard Kluckow, *Bureau of Justice Statistics*
Amy Lauger, *Bureau of Justice Statistics*
Tiandong Li, *Health Resources and Services Administration*
Aaron Maitland, *National Center of Health Statistics*
Brett Matsumoto, *Bureau of Labor Statistics*
Pam McGovern, *National Agricultural Statistics Service*
Valerie Testa, *Internal Revenue Service*
Laura Tiehen, *Economic Research Service*
Victoria Udalova, *U.S. Census Bureau*
Steven Wallander, *Economic Research Service*
Scott Wentland, *Bureau of Economic Analysis*
Doug Williams, *Bureau of Labor Statistics*

Federal Committee on Statistical Methodology

Rochelle (Shelly) Martinez, *Office of Management and Budget (Chair)*
Stephen Blumberg, *National Center for Health Statistics*
Chris Chapman, *National Center for Education Statistics*
John Eltinge, *U.S. Census Bureau*
John Finamore, *National Center for Science and Engineering Statistics*
Dennis Fixler, *Bureau of Economic Analysis*
Michael Hawes, *U.S. Census Bureau*
Wendy Martinez, *Bureau of Labor Statistics*
Jaki McCarthy, *National Agricultural Statistics Service*
Jennifer Nielsen, *National Center for Education Statistics (Secretary)*
Jennifer Ortman, *U.S. Census Bureau*
Anne Parker, *Internal Revenue Service*
Jennifer D. Parker, *National Center for Health Statistics*
Polly Phipps, *Bureau of Labor Statistics*
Mark Prell, *Economic Research Service*
Joseph Schafer, *U.S. Census Bureau*
Rolf R. Schmitt, *Bureau of Transportation Statistics*
Marilyn M. Seastrom, *National Center for Education Statistics*
Robert Siviniski, *Office of Management and Budget*
G. David Williamson, *Agency for Toxic Substances and Disease Registry*
Linda Young, *National Agriculture Statistics Service*

The Council of Professional Associations on Federal Statistics (COPAFS) is devoted to educational activities and preserving the public good represented by federal statistical collections. Since 1980, COPAFS has provided an open dialog between those who use federal statistics in professional contexts and the Federal statistical agencies that produce those statistics for the public good. Supporting organizations include professional associations, businesses, research institutes, and others that help to produce and/or use federal statistics.

Our Goal: Advancing Excellence in Federal Statistics.

COPAFS' objectives are to:

- Increase the level and scope of knowledge about developments affecting Federal statistics;
- Encourage discussion within and among professional organizations to respond to important issues in Federal statistics and bring the views of professional associations to bear on decisions affecting Federal statistical programs.

In support of these objectives, COPAFS:

- Obtains information on developments in statistics through discussions with officials, attendance at congressional hearings and meetings of statistical advisory committees, engaging with the broader statistical community, and reviewing recent reports or directives affecting the Federal statistical system;
- Disseminates information and encourages discussion and action on developments in federal statistics through correspondence and presentations at COPAFS and professional association meetings, direct calls for action via email, and announcements on social media; and
- Plans and presents educational programs on uses of statistics in policy formulation, public and private decision- making, research, the distribution of products, and the allocation of resources.

COPAFS helps:

- Professional associations and other organizations obtain and share information about developments affecting federal statistical programs;
- Federal agencies to disseminate information on developments of interest to the professional community and to obtain advice about professional societies' concerns and priorities;
- Congressional offices to clarify issues and questions about the federal statistical system, to plan hearings related to federal statistical programs, and to identify experts to testify; and
- The public to learn more about the federal statistical agencies, to communicate views of data users concerning Federal statistical activities, and to obtain a better understanding of how policy and budget are likely to affect the availability of federal statistics.

Member associations and affiliates appoint representatives to serve on the Council and to attend its quarterly meetings. The representatives are responsible for establishing COPAFS' priorities and guiding its activities. At each quarterly meeting, members discuss significant cross-cutting issues, hear from statistical agencies and other producers and users of data, and make suggestions for further action.

The Board of Directors, comprised of elected officers and four at-large members, facilitates the work of COPAFS between meetings. The Executive Director is responsible for the day-to-day operations. Financial support for COPAFS' on-going programs comes from annual member dues.

2021 FCSM Research and Policy Conference

The Federal Committee on Statistical Methodology (FCSM) is an interagency committee dedicated to improving the quality of federal statistics. This conference helps the committee achieve their major goals, which are to:

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

The 2021 FCSM Research and Policy Conference will focus on trust as the cornerstone of federal statistics and evidence building, and the role of the Federal Statistical System in data collection and outcomes research that supports evidence-based policy making. The conference provides a forum for experts and practitioners from around the world to discuss and exchange current methodological knowledge and policy insights about topics of current and critical importance to federal agencies as well as the Federal Statistical System as a whole.

Each day of the conference will offer papers on a wide range of topics relevant to the production, quality and use of federal statistics. Attendees from a range of backgrounds will find sessions of interest, including statistical methods, administrative data, questionnaire design, program evaluation, policy making, and more.

Sessions feature presentations by government, private sector, and academic researchers from multiple countries. All sessions will include an open discussion and some sessions will include a formal discussion. Presentations will be made available on the conference website following the conference.

KEYNOTE SPEAKER

Helen Nissenbaum, *Professor of Information Science at Cornell Tech*

Schedule Overview

Tuesday, November 2

8:30 - 9:30 a.m. **Welcoming Remarks and PLENARY SESSION**

9:30 - 10:00 a.m. **Break**

10:00 - 11:45 a.m. **CONCURRENT SESSION A**
A-1. Survey Design Challenges During COVID-19
A-2. Re-inventing and Integrating Annual Economic Survey Programs at the US Census Bureau
A-3. Recent Advances in Disclosure Limitation and Data Privacy
A-4. Data Quality: Communication of Uncertainty in Official Statistics
A-5. Using Multiple Survey Modes or Administrative Data to Improve Estimates

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

1:15 – 3:00 p.m. **CONCURRENT SESSION B**
B-1. Measuring Sexual Orientation & Gender Identity
B-2. Impact of the COVID-19 Pandemic on National Center for Health Statistics Data Collections
B-3. Sampling & Calibration for Data Versatility
B-4. Equity
B-5. Topics in Survey Data Quality

3:00 – 3:15 p.m. **Break**

3:15 – 5:00 p.m. **CONCURRENT SESSION C**
C-1. Recruiting and Surveying Victims Using Social Media and Online Platforms: Promising Practices and Lessons Learned
C-2. Evidence & Data Policy
C-3. Sampling Issues, Estimation, and Testing
C-4. The Standard Application Process: A Coordinated Effort across the Federal Statistical System to Implement an Evidence Act Requirement
C-5. Leveraging Administrative Data for Survey Methods and Research

Wednesday, November 3

8:30 - 10:15 a.m. **CONCURRENT SESSION D**
D-1. Data Collection Challenges During COVID-19
D-2. Innovations in Health Insurance Data Collection & Measurement Across Federal Surveys
D-3. Enhancing Transparency, Reproducibility and Privacy
D-4. A Call for Federal Statistical Coordination on Climate Change Data Needs
D-5. Data Linkage for Government Health & Safety Statistics: Methods & Applications

10:15 – 10:45 a.m. **Break**

10:45 a.m. – 12:30 p.m. **CONCURRENT SESSION E**
E-1. Strategies for Recruiting & Identifying Subpopulations
E-2. Creative Problems in Social Science
E-3. Advances in Disclosure Limitation & Publication Standards
E-4. Communicating Fitness for Use
E-5. Collecting Spending Data during a Pandemic – an Evaluation of Quality and Response in the Consumer Expenditure Surveys

12:30 p.m. - 1:45 p.m.
Lunch on Your Own

1:45 – 3:30 p.m. **CONCURRENT SESSION F**
F-1. Supplementing data collection and research at the National Center for Health Statistics using the Research and Development Survey
F-2. From Data Collection to Estimation: Highlights of Survey Lifecycle Issues from FoodAPS
F-3. Privacy and Sample Surveys
F-4. Considerations for Calculating and Communicating the Value of Federal Statistics
F-5. Web Scraping

3:30 – 3:45 p.m. **Break**

3:45 - 5:30 p.m. **CONCURRENT SESSION G**
G-1. Reply YES to Innovate: Text Messages for Federal Surveys
G-3. Statistical Methods for Improving Small Domain and Key Survey Estimates
G-4. Data Quality-Nonresponse Bias
G-5. Science and Engineering Indicators: Measuring S&E Education, Work Force, and Research Output

Thursday, November 4

8:30 - 10:15 a.m. **CONCURRENT SESSION H**
H-1. Data Collection in the Post-Pandemic Era
H-2. Using Administrative Data to Examine Food Assistance Program Effectiveness
H-3. Modernizing Data Dissemination
H-5. Assessing & Improving the Quality of Prescription Drug Data from Surveys

10:15 - 10:30 a.m. **Break**

10:30 a.m. – 12:15 p.m. **CONCURRENT SESSION I**
I-1. Survey Design & Measurement
I-3. Price Indices and Consumer Demand
I-4. A Discussion of the Final Report of the Interagency Technical Working Group on Evaluating Alternative Measures of Poverty
I-5. Machine Learning Applications

12:15 p.m. - 1:30 p.m.
Lunch on Your Own

1:30 – 3:15 p.m. **CONCURRENT SESSION J**
J-1. Modernizing Federal Survey Data Collections through Modularized Design
J-2. Perseverance & Resilience: The Medical Expenditure Panel Survey in the Pandemic
J-3. After the Fact - Data Revision, Imputation, and Analysis Methods
J-5. The Reporting, Analysis, and Mitigation of Nonresponse Bias in Federal Surveys

3:15 – 3:30 p.m.
Break

3:30 - 5:15 p.m. **CONCURRENT SESSION K**
K-1. Previously Reported Data & Dependent Interviewing in Official Statistics: Existing Practices & New Findings
K-4. Academic-Statistical Agency Collaboratives to Create Data Infrastructure for Evidence-Building
K-5. Leveraging Data Science to Improve Survey Operations

Plenary Session
Tuesday, November 2nd
8:30 – 9:30 AM

"All the parameters matter!" Privacy as Contextual Integrity for Statistical Reporting
Helen Nissenbaum, *Professor of Information Science at Cornell Tech*

Dr. Nissenbaum is director of the Digital Life Initiative and Professor of Information Science at Cornell Tech. Her books include Obfuscation: A User's Guide to Privacy and Protest (with Finn Brunton), Values at Play in Digital Games (with Mary Flanagan), and Privacy in Context: Technology, Policy, and the Integrity of Social Life, where she lays out the theory of contextual integrity. Her research spans topics of privacy, security, accountability, bias, data concentration, and values in design as manifest in digital technologies. Beyond academic publication, Nissenbaum works on free software tools, Adnauseam and TrackMeNot, defending privacy, autonomy, and freedom online. She holds a BA (Hons) in mathematics and philosophy from the University of Witwatersrand and a Ph.D. in philosophy from Stanford University.

Concurrent Session A
Tuesday, November 2nd
10:00 - 11:45 AM

Session A-1: Survey Design Challenges During COVID-19

Session Liaison: Aaron Maitland (National Center for Health Statistics)
Session Chair: Aaron Maitland (National Center for Health Statistics)

Analysis of Open-text Time Reference Web Probes on a COVID-19 Survey
Kristen Cibelli Hibben, Valerie Ryan, Travis Hoppe, Paul Scanlon and Kristen Miller (National Center for Health Statistics)

Applying Client Relationship Management Techniques to an Institutional Contacting Guide Redesign
Robert McCracken, R. Suresh, Brandon Peele, Jason Kennedy, Shawn Cheek, Joe Nofziger (RTI International) and Peter Tice (Agency for Healthcare Research and Quality)

Cross Agency Collaboration in the Rapid Development of COVID-19 Questionnaire Items for the Medicare Current Beneficiary Survey (MCBS)
Andrea Mayfield, Kylie Carpenter, Rachel Carnahan, Elise Comperchio and Felicia LeClere (NORC at the University of Chicago)

Missing Price Imputation in the Producer Price Index
Yoel Izsak and Monica Moleres (Bureau of Labor Statistics)

National Ambulatory Medical Care Survey: Redesigning for the Future Health Care Landscape
Sonja N. Williams, Jill J. Ashman, and Brian W. Ward (National Center for Health Statistics)

Session A-2: Re-inventing and Integrating Annual Economic Survey Programs at the U.S. Census Bureau

Session Liaison: Scott Boggess, (U.S. Census Bureau)
Session Chair: Kimberly P. Moore (U.S. Census Bureau)
Session Organizer: Diane K. Willimack (U.S. Census Bureau)

Developing an Integrated Annual Survey
Jenna Morse, James Burton, and Kimberly Moore (U.S. Census Bureau)

Aiding a Holistic View of Businesses through Intensive Respondent Research
Diane K. Willimack (U.S. Census Bureau)

Determining and Harmonizing Content in Developing the Annual Integrated Economic Survey
Blynda Metcalf, Heidi St. Onge, and Kimberly Moore (U.S. Census Bureau)

Developing a Unified Sample Design for the Integrated Annual Survey
Katherine Jenny Thompson, James Burton, James Hunt, and Amy Newman Smith (U.S. Census Bureau)

Discussant: Edward T. Morgan (Bureau of Economic Analysis)

Session A-3: Recent Advances in Disclosure Limitation and Data Privacy

Session Liaison: Darius Singpurwalla (National Center for Science and Engineering Statistics)
Session Chair: Katherine Jenny Thompson (U.S. Census Bureau)
Organizer: Harrison Quick (Drexel University)

Disclosure Limitation in the Census of Fatal Occupational Injuries
Ellen Galantucci (Bureau of Labor Statistics)

Synthetic Public Use File of Administrative Tax Data: Methodology, Utility, and Privacy Implications
Claire Bowen (Urban Institute)

Accuracy Gains from Privacy Amplification through Sampling for Differential Privacy
Jingchen Hu (Vassar College), Joerg Drechsler (University of Maryland) and Hang J. Kim (University of Cincinnati)

A Latent Class Modeling Approach for DP Synthetic Data Contingency Tables
Andres Felipe Barrientos (Florida State University), Michelle Pistner Nixon (Pennsylvania State University), Jerome P. Reiter (Duke University), and Aleksandra Slavkovic (Pennsylvania State University)

Improving the Utility of Poisson-Distributed, Differentially Private Synthetic Data via Prior Predictive Truncation with an Application to CDC WONDER
Harrison Quick (Drexel University)

Session A-4: Data Quality: Communication of Uncertainty in Official Statistics

Session Liaison: Linda Young (National Agriculture Statistics Service)
Session Chair: John L. Eltinge (United States Census Bureau)

Evaluating Uncertainty in Multiple Dimensions of Data Quality
John L. Eltinge (United States Census Bureau)

More Fully Capturing Uncertainty Associated with Official Estimates
Linda Young (National Agriculture Statistics Service)

Tailored Transparency: Public Trust vs. Reproducibility
Peter V. Miller (Northwestern University and United States Census Bureau (Retired))

Discussant: Jeffrey M. Gonzalez (Economic Research Service)

Session A-5: Using Multiple Survey Modes or Administrative Data to Improve Estimates

Session Liaison: Amy Lauger (Bureau of Justice Statistics)
Session Chair: Katie Genadek (U.S. Census Bureau)

It's All a Matter of Degrees: Comparing Survey and Administrative Educational Attainment Data
Andrew Foote and Larry Warren (U.S. Census Bureau)

Implementing a Multi-Mode Approach to Household Eligibility Screening for 2021-2022 NHANES
Juliana McAllister, Allan Uribe, Jessica Graber, Denise Schaar, and Chia-Yih Wang (National Center for Health Statistics)

A Meta-Analysis of the Impact of Number of Contact Attempts on Response Rates and Web Completion Rates in Multimode Surveys
Ting Yan (Westat)

Discussant: Katie Genadek (U.S. Census Bureau)

Concurrent Session B
Tuesday, November 2nd
1:15 - 3:00 PM

Session B-1: Measuring Sexual Orientation & Gender Identity

Session Liaison: Doug Williams (Bureau of Labor Statistics)
Session Chair: Doug Williams (Bureau of Labor Statistics)

Best Practices for Collecting Gender and Sex Data
Wendy Martinez (Bureau of Labor Statistics), Suzanne Thornton (Swarthmore), Dooti Roy, Stephen Perry, Donna LaLonde, Renee Ellis, and David J. Corliss

Measuring Sexual Orientation and Gender Identity Among College Graduates: Results from a Methodological Experiment Using a Non-Probability Sample
Rebecca L. Morrison (National Center for Science & Engineering Statistics), Jesse Chandler (Mathematica), Flora Lan, and Karen S. Hamrick (National Center for Science & Engineering Statistics)

Assessing Measurement Error in Sex and Gender Identity Measures in an NCES Survey
Elise Christopher (National Center for Education Statistics)

Session B-2: Impact of the COVID-19 Pandemic on National Center for Health Statistics Data Collections

Session Liaison: Jessica Graber (National Center for Health Statistics)
Session Chair: Brian W. Ward (National Center for Health Statistics)

Impact of the Pandemic on National Health Interview Survey Data Collection
Stephen J. Blumberg (National Center for Health Statistics)

COVID-19 Pandemic Impact on the NCHS Provider-Based Surveys
Carol DeFrances, Brian W. Ward, and Manisha Sengupta (National Center for Health Statistics)

The National Health and Nutrition Examination Survey (NHANES): The Impact of the COVID-19 Pandemic on Data Collection and Release
Ryne Paulose-Ram, Jessica E. Graber, and Namanjeet Ahluwalia (National Center for Health Statistics)

Expanding the NCHS Research and Development Survey During the COVID-19 Pandemic
Jennifer D. Parker (National Center for Health Statistics)

Discussant: Denys T. Lau (National Committee for Quality Assurance)

Session B-3: Sampling & Calibration for Data Versatility

Session Liaison: Laura Tiehen, (Economic Research Service)
Session Chair: Laura Tiehen, (Economic Research Service)

An Easy Way to Calibrate on Partly Known Overlapping Multiple Totals in Frequency Tables with Application to Real Data
Michael Sverchkov (Bureau of Labor Statistics)

Designing a Probability Sample to Produce a Large Number of Key Estimates
Philip Kott and Darryl V. Creel (RTI International)

Sample-Based Calibration of Multiple Surveys
Jean Opsomer and Weijia Ren (Westat), John Foster (NOAA)

Sampling and Estimation for Multipurpose Surveys
Yang Cheng, Jeff Bailey, (National Agricultural Statistics Service), Eric Slud (U.S. Census Bureau/University of Maryland) and Lu Chen (National Agricultural Statistics Service and National Institute of Statistical Sciences)

Sampling Using Multiple Measures of Size: A Simulation Study
Tim Keller, Mark Apodaca Franklin Duan, and Peter Quan (National Agricultural Statistics Service)

Session B-4: Equity

Session Liaison: Linda Young (National Agriculture Statistics Service)
Session Chair: Linda Young (National Agriculture Statistics Service)
Organizer: Robert Sivinski (Office of Management and Budget) and Chris Chapman (National Center for Education Statistics)

Policy Review of Initiatives to Improve Equity Information
Jamie Keene, *Executive Office of the President*

Educational Equity: Identifying and Presenting Information within New Online Resources
Ross Santy, *National Center for Education Statistics*

Healthy People: Exploring Disparities in the Nation's Health
David Huang, *National Center for Health Statistics*

Delivering Equity in Federal Forms and Surveys: LGBTQ+ Data Collection
Amy Paris, *Department of Health and Human Services*

Session B-5: Topics in Survey Data Quality

Session Liaison: Brett Matsumoto (Bureau of Labor Statistics)
Session Chair: Brett Matsumoto (Bureau of Labor Statistics)

Survey Isolation during COVID-19: The Effects of Suddenly Relying on Address-Matched Phone Numbers for Interviewing Households
Jonathan Eggleston, Yarissa Gonzalez, Tim Trudell, John Voorheis (U.S. Census Bureau)

The New Non-employer Business Demographics Statistics: Responding to 20th-Century Survey-Based Statistics Challenges while Addressing 21st- Century Needs
Adela Luque , Ken Rinz, James Noon, and Michaela Dillon(U.S. Census Bureau)

Participation Metrics for Accelerometer-Based Research
Christopher Antoun (University of Maryland) and Alexander Wenz (University of Mannheim)

Developing State Personal Income Distribution Statistics
Dirk van Duym and Christian Awuku-Budu (Bureau of Economic Analysis)

Concurrent Session C
Tuesday, November 2nd
3:15 - 5:00 PM

Session C-1: Recruiting and Surveying Victims Using Social Media and Online Platforms: Promising Practices and Lessons Learned

Session Liaison: Paul Schroeder (Council of Professional Associations on Federal Statistics)
Session Chair: Heather Brotsos (Bureau of Justice Statistics)

Evaluating Crime Survey Responses and Engagement among Juveniles and their Parents using Social Media and Online Platforms
Jenna Truman and Grace Kena (Bureau of Justice Statistics) and Chris Krebs (RTI International)

Testing Improvements to the NCVS Hate Crime Items using Online Survey Panels
Grace Kena (Bureau of Justice Statistics), Lynn Langton and Chris Krebs (RTI International)

Using Online Survey Panels to Enhance Identity Theft Measurement
Lynn Langton and Chris Krebs (RTI International), Erika Harrell and Grace Kena (Bureau of Justice Statistics)

Discussant: Heather Warnken (Department of Justice)

Session C-2: Evidence & Data Policy

Session Liaison: Steven Wallander (Economic Research Service)
Session Chair: Steven Wallander (Economic Research Service)

Automated Collection of Publicly Available Data from the Internet
Anup Mathur, Michael Castro, and Sumit Khaneja (U.S. Census Bureau)

Creating Policy Tools for Broadband Subsidy programs: Combining Spatial Regression Discontinuity Designs and Bayesian Wombling
Aritra Halder and Joshua R. Goldstein (University of Virginia), John Pender (Economic Research Service), Devika Mahoney-Nair, Aaron Schroeder, Neil Kattampallil, Stephanie S. Shipp, and Sallie Keller (University of Virginia)

Evidence Act Standard Application Process: Stakeholder Engagement and Observations
Jon Desenberg (The MITRE Corporation), Heather Madray (U.S. Census Bureau), and Sue Collin (The MITRE Corporation)

Six Crises or One Dozen Opportunities in Public-Stewardship Statistics
John Eltinge (U.S. Census Bureau)

Session C-3: Sampling Issues, Estimation, and Testing

Session Liaison: Scott Wentland (Bureau of Economic Analysis)
Session Chair: Scott Wentland (Bureau of Economic Analysis)

Business Sample Revision Universe Extraction Measure of Size Determination and Validation
Kenishea Donaldson, Katrina Washington, Erica Wong, Olivia Brozek, Jacob Sandruck, Micah Tyler (U.S. Census Bureau)

Influential Unit Treatment in the Annual Survey of Local Government Finances
Noah Bassel (U.S. Census Bureau)

Probabilistic Classification of a Policy Relevant Subpopulation: The Case of High-Tech Start-ups Operating in Innovation Markets
Timothy R. Wojan, John Jankowski, and Audrey Kindlon (National Center for Science and Engineering Statistics)

Roots from Trees: A Machine Learning Approach to Unit Root Decision
Gary Cornwall (Bureau of Economic Analysis), Jeff Chen (University of Cambridge), Beau Sauley (Murray State University)

Understanding the Characteristics of Unresolved Matched Records in Capture- Recapture Methodology
Denise A. Abreu (National Agricultural Statistics Service)

Session C-4: The Standard Application Process: A Coordinated Effort across the Federal Statistical System to Implement an Evidence Act Requirement

Session Liaison: Darius Singpurwalla (National Center for Science and Engineering Statistics)
Session Chair: John Finamore (National Center for Science and Engineering Statistics)

SAP Pilot: A One-Stop Application Portal
Heather Madray (U.S. Census Bureau)

SAP Policy Guidance Development
Mark Prell (Economic Research Service)

SAP Stakeholder Engagement and Outreach Efforts
Vipin Arora (National Center for Science and Engineering Statistics)

The Fully Functional SAP Portal
John Finamore (National Center for Science and Engineering Statistics)

Session C-5: Leveraging Administrative Data for Survey Methods and Research

Session Liaison: Scott Boggess, (U.S. Census Bureau)
Session Chair: Scott Boggess, (U.S. Census Bureau)

Comparing the 2019 American Housing Survey to Contemporary Sources of Property Tax Records: Implications for Survey Efficiency and Quality
Ariel J. Binder (US Census Bureau), Emily Molfino (Department of Housing and Urban Development) and John Voorheis (US Census Bureau)

Determining Household Obesity Status Using Scanner Data
Elina T. Page, Sabrina K. Young, Megan Sweitzer, and Abigail Okrent (Economic Research Service)

Measuring the Distributional Effects of Climate Change and Environmental Inequality with Linked Survey, Census and Administrative Data
John Voorheis (US Census Bureau)

An Evaluation of the Gender Wage Gap using Linked Census and Administrative Records
Brad Foster and Marta Murray-Close (US Census Bureau), Christin Landivar and Mark deWolf (US Department of Labor)

Identifying Undercounted Children Using Birth Records
Gloria Aldana (US Census Bureau)

Concurrent Session D
Wednesday, November 3rd
8:30 – 10:15 AM

Session D-1: Data Collection Challenges During COVID-19

Session Liaison: Valerie Testa (Internal Revenue Service)
Session Chair: Valerie Testa (Internal Revenue Service)

Implications of the Coronavirus Pandemic on the National Public Education Financial Survey (NPEFS) and the School District Finance Survey (F-33)
Stephen Q. Cornman (U.S. Department of Education), Malia Howell, Osei Ampadu, and Stephen Wheeler (U.S. Census Bureau)

Measuring the Impacts of the Coronavirus Pandemic on Higher Education R&D Expenditures and Research Space: Experience from Survey Question Development at the National Center for Science and Engineering Statistics
Michael T. Gibbons (National Center for Science and Engineering Statistics)

Patterns of Response During Covid-19 in a National Survey of Businesses: A Look at the Medical Expenditure Panel Survey - Insurance Component (MEPS-IC)
David Kashihara (Agency for Healthcare Research and Quality)

The 2020 COVID-19 Module on the Survey of Graduate Students and Postdocs in Science and Engineering
Caren Arbeit and Pat Green (RTI International) and Mike Yamaner (National Center for Science and Engineering Statistics)

Session D-2: Innovations in Health Insurance Data Collection & Measurement Across Federal Surveys

Session Liaison: Victoria Udalova (U.S. Census Bureau)
Session Chair: Laryssa Mykyta (U.S. Census Bureau)

Two times a charm? Verifying Reports of Uninsurance in a National Survey
Paul Jacobs and Patricia Keenan (Agency for Healthcare Research and Quality)

Improving Measurement of VA Health Coverage among Military Veterans on the National Health Interview Survey
Robin A. Cohen and Carla E. Zelaya (National Center for Health Statistics)

Decomposing Data Processing Improvements on Estimates Health Insurance Coverage in the Current Population Survey Annual Social and Economic Supplement (CPS ASEC)
Laryssa Mykyta, Amy Steinweg, and Katherine Keisler-Starkey (U.S. Census Bureau)

Using Insurance Claims Data in the Medical Price Indexes
Brian Parker, John Bieler, Caleb Cho, and Brett Matsumoto (Bureau of Labor Statistics)

Discussant: Steven Cohen (RTI)

Session D-3: Enhancing Transparency, Reproducibility and Privacy

Session Liaison: Amy Lauger (Bureau of Justice Statistics)
Session Chair: Amy O'Hara (Georgetown University)

A Federal Tiered Access Model to Promote Evidence-Based Policymaking
Amy O'Hara and Lahy Amman (Georgetown University)

Sharing Student Data Across Organizational Boundaries Using Secure Multiparty Computation
Stephanie Straus (Georgetown University), David Archer (Galois, Inc.), Amy O'Hara (Georgetown University) and Rawane Issa (Boston University)

The Privacy Paradox: How well do Respondent Attitudes and Concerns about Privacy Predict Privacy-Related Behaviors?
Casey Eggleston, Aleia Clark Fobia, Jennifer Hunter Childs (U.S. Census Bureau)

Discussant: Keenan Dworak-Fisher (Bureau of Labor Statistics)

Session D-4: A Call for Federal Statistical Coordination on Climate Change Data Needs

Session Liaison: David Kashihara (Agency for Healthcare Research and Quality)
Session Chair: Eileen O'Brien (National Agricultural Statistics Service)
Organizer: Eileen O'Brien (National Agricultural Statistics Service)

Panelists:

- Mindy Selman – Senior Analyst, Office of Energy and Environmental Policy, OCE, U.S. Department of Agriculture
- Carla Frisch – Principal Deputy Director, Office of Policy, U.S. Department of Energy
- Richard Allen – Chief Data Officer and Strategist, U.S. Environmental Protection Agency
- Nick Hart – President, Data Foundation

Session D-5: Data Linkage for Government Health & Safety Statistics: Methods & Applications

Session Liaison: Ellen Galantucci (Bureau of Labor Statistics)
Session Chairs: Ellen Galantucci (Bureau of Labor Statistics) and Victoria Udalova (U.S. Census Bureau)

Assessing Linkage Eligibility Bias in the National Health Interview Survey
Jonathan Aram, Crescent B. Martin, and Lisa B. Mirel (National Center for Health Statistics)

Using Administrative Data to Supplement and Assess Occupational Health and Safety Statistics
Ellen Galantucci (Bureau of Labor Statistics)

Social Determinants of Emergency Department Utilization in Utah
David Powers, Sara Robinson, Edward Berchick, J. Alex Branham, Lucinda Dalzell, Lorelle Dennis, Kristi Eckerson, Alfred Gottschalck, Joanna Motro, John Posey, Andrew Verdon, and Victoria Udalova (U.S. Census Bureau)

Examining Earnings of U.S. Physicians Using Tax Return Information
Victoria Udalova (U.S. Census Bureau)

Concurrent Session E
Wednesday, November 3rd
10:45 - 12:30 PM

Session E-1: Strategies for Recruiting & Identifying Subpopulations

Session Liaison: Doug Williams (Bureau of Labor Statistics)
Session Chair: Doug Williams (Bureau of Labor Statistics)

Administrative Data Limitations and the Need for Continued Improvement of Face to Face Interviewing for Non-English Speakers in National Surveys
Patricia Goerman (U.S. Census Bureau), Alisu Schoua Glusberg (Research Support Services), and Leticia Fernandez (U.S. Census Bureau)

Hard to Get: Understanding Why People Take Our Surveys
Mina Muller, Seth Messinger and Randall K. Thomas (Ipsos Public Affairs)

Making Data Collection More Efficient? Optimizing Incentives and Reminder Modes
Jerry Timbrook, Antje Kirchner, and Emilia Peytcheva (RTI International)

Recruiting a Probability Sample of 18 year olds for a Longitudinal Study on Interpersonal Violence
David Cantor and Reanne Townsend (Westat)

Using Crowdsourcing for Survey Administration: A Study of Innovation Activity among Individuals
Audrey Kindlon (National Center for Science and Engineering Statistics), Jesse Chandler (MPR) and Rebecca Morrison (National Center for Science and Engineering Statistics)

Session E-2: Creative Problems in Social Science

Session Liaison: Steven Wallander (Economic Research Service)
Session Chair: Steven Wallander (Economic Research Service)

Assessing the Drivers of U.S. Food Expenditures
Eliana Zeballos, Wilson Sinclair, Timothy Park (Economic Research Service)

Estimation of Deaths among Health Care Personnel (HCP) with COVID-19 using Capture-Recapture Methods — United States, March 17–April 29, 2020
Jennifer Rammon, Kerui Xu, Matt Stuckey, Michelle Hughes, Reid Harvey, Sherry Burrer, Sophia Chiu, Jess Rinsky, and Matthew Groenewold (Centers for Disease Control and Prevention)

Measuring the US Space Economy
Tina Highfill, Annabel Jouard and Connor Franks (Bureau of Economic Analysis)

Session E-3: Advances in Disclosure Limitation & Publication Standards

Session Liaison: Ellen Galantucci (Bureau of Labor Statistics)
Session Chair: Ellen Galantucci (Bureau of Labor Statistics)

Comparative Analysis of Differential Privacy and Swapping Methods in the Context of the U.S. Census
Miranda Christ and Sarah Radway (Columbia University)

Evaluating Publication Rules for the County Agricultural Production Survey Using Simulation
Andrew Dau, Nathan Cruze, Joe Parsons, and Linda Young (NASS)

Posterior Risk and Utility from Private Synthetic Weighted Survey Data
Quentin Brummet (NORC at the University of Chicago) and Jeremy Seeman (Pennsylvania State University)

Private Tabular Survey Data Products through Synthetic Microdata Generation
Terrance Savitsky (Bureau of Labor Statistics), Matthew Williams (National Center for Science and Engineering Statistics), and Jingchen Hu (Vassar College)

Session E-4: Communicating Fitness for Use

Session Liaison: Jessica Graber (National Center for Health Statistics)
Session Chair: Jennifer Parker (National Center for Health Statistics)

Using Data Visualizations, Short Articles, and Social Media to Communicate Complex Data Simply
Jay Meisenheimer (Bureau of Labor Statistics)

Journalists Communicating Statistical Information
Regina Nuzzo (American Statistical Association)

Short Communication as a Medium: Is Engagement a Substitute for Efficacy?
Travis Hoppe (National Center for Health Statistics)

Discussant: G. David Williamson (Centers for Disease Control and Prevention)

Session E-5: Collecting Spending Data during a Pandemic – an Evaluation of Quality and Response in the Consumer Expenditure Surveys

Session Liaison: Scott Boggess, (U.S. Census Bureau)
Session Chair: Laura Erhard (Bureau of Labor Statistics)

An Examination of Nonresponse Bias in the Consumer Expenditures Survey during the COVID-19 Period
Stephen Ash, Brian Nix, Barry Steinberg, and David Swanson (Bureau of Labor Statistics)

Consumer Expenditure Interview Survey: Data Quality Assessment Pre vs. Post COVID-19
Yezzi Angi Lee and David Biagas (Bureau of Labor Statistics)

Evaluating Diary Collection Mode Changes in the Context of the COVID-19 Pandemic
Brett McBride and Nikki Graf (Bureau of Labor Statistics)

COVID-19's Effect on the Consumer Expenditure Surveys' Estimates
Scott Curtin, Bryan Rigg, and Brett Creech (Bureau of Labor Statistics)

Discussant*:* Stephanie Eckman (RTI)

Concurrent Session F
Wednesday, November 3rd
1:45 - 3:30 PM

Session F-1: Supplementing Data Collection and Research at the National Center for Health Statistics using the Research and Development Survey

Session Liaison: Jessica Graber (National Center for Health Statistics)
Session Chair: Katherine Irimata (National Center for Health Statistics)

Introduction to the Research and Development Survey and Calibration Approaches for Panel Surveys
Katherine Irimata (National Center for Health Statistics)

An Overview of the 2019 Research and Development Survey (RANDS)
Li-Yen Rebecca Hu, Paul Scanlon, Kristen Miller, Yulei He, Katherine Irimata, Guangyu Zhang, and Kristen Cibelli Hibben (National Center for Health Statistics)

Comparison of Mental Health Estimates by Sociodemographic Characteristics in the Research and Development Survey and National Health Interview Survey, 2019
Leanna Moron, Katherine Irimata, and Jennifer Parker (National Center for Health Statistics)

Estimates from Selected Variables in Three Rounds of RANDS during COVID-19 Pandemic
Rong Wei, Yulei He, Van Parsons, and Paul Scanlon (National Center for Health Statistics)

Discussant: Michael Yang (NORC at the University of Chicago)

Session F-2: From Data Collection to Estimation: Highlights of Survey Lifecycle Issues from FoodAPS

Session Liaison: Paul Schroeder (Council of Professional Associations on Federal Statistics)
Session Chair: Joseph Rodhouse (National Agricultural Statistics Service)

Is "Proof of Purchase" Really Proof?
Adam Kaderabek and Brady T. West, (University of Michigan), John A. Kirlin (Kirlin Analytic Services), Elina T. Page and Jeffrey M. Gonzalez (Economic Research Service)

Usability Evaluation of Smartphone-based Data Collection Instrument
Lin Wang, Anthony Schulzetenberg, and Alda G. Rivas, (U.S. Census Bureau) Heather Ridolfo (National Agricultural Statistical Service) and Shelley Feuer (U.S. Census Bureau)

Examination of the Data Quality Properties of USDA and Proprietary Databases with Information on Food Item and Food Retailers to Reduce Nonsampling Errors in FoodAPS-2
Clare Milburn (The George Washington University) and Jeffrey M. Gonzalez, Linda Kantor and Elina T. Page (Economic Research Service)

Using the Weighted Finite Population Bayesian Bootstrap to Account for Complex Sample Design Features when Estimating State- and Sub-State-Level Food Insecurity Prevalence
Katherine Li, Yajuan Si and Brady T. West (University of Michigan), John A. Kirlin (Kirlin Analytic Services), Xingyou Zhang (U.S. Bureau of Labor Statistics)

Discussant: Stephanie Zimmer (RTI International)

Session F-3: Privacy and Sample Surveys

Session Liaison: Darius Singpurwalla (National Center for Science and Engineering Statistics)
Session Chair: Rolando A. Rodríguez (U.S. Census Bureau)

Adapting Surveys for Formal Privacy: Where We Are, Where We Are Heading
Aref Dajani (U.S. Census Bureau)

Controlling Privacy Loss in Survey Sampling
Mark Bun (Boston University), Joerg Drechsler (University of Maryland) Marco Gaboardi (Boston University) Audra McMillan (Apple), and Jayshree Sarathy (Harvard University)

Leveraging Public Data for Practical Private Query Release
Terrance Liu (Carnegie Mellon University), Giuseppe Vietri (University of Minnesota), Thomas Steinke (Google), Jonathan Ullman (Northeastern University), and Steven Wu (Carnegie Mellon University)

Discussant: Jerry Reiter (Duke University)

Session F-4: Considerations for Calculating and Communicating the Value of Federal Statistics

Session Liaison: David Kashihara (Agency for Healthcare Research and Quality)
Session Chair: Jonathan Auerbach (George Mason University)
Session Organizer: Steve Pierson (American Statistical Association)

Panel Discussion: Reflections on Census Quality Indicators Taskforce

- Constance F. Citro - Senior Scholar, Committee on National Statistics
- Julia Lane, Professor - NYU and Cofounder, Coleridge Initiative
- Joseph Salvo – Institute Fellow, Social and Data Analytics Division, University of Virginia Biocomplexity Institute; Senior Advisor, National Conference on Citizenship

Discussant: Andrew Reamer – Research Professor, George Washington Institute of Public Policy, George Washington University

Session F-5: Web Scraping

Session Liaison: Maura Bardos (U.S. Energy Information Administration)
Session Chair: Maura Bardos (U.S. Energy Information Administration)

Detecting and Measuring Product Innovation in News Articles Using Natural Language Processing Methods
Gizem Korkmaz and Neil Alexander Kattampallil (University of Virginia) and Gary Anderson (National Science Foundation)

Machine-Learning Based Identification of Emerging Research Topics Using Research & Development Administrative Data
Eric J. Oh, Kathryn Linehan, Joel Thurston, Stephanie Shipp, and Sallie Keller, (University of Virginia) and John Jankowski, Audrey Kindlon (National Center for Science and Engineering Statistics)

Using Web Scraping and Network Analysis to Study International Collaboration in Open Source Software
Brandon Kramer, Gizem Korkmaz, José Bayoán Santiago Calderón, and Carol Robbins (University of Virginia)

Progress in the Use of Web-Scraped List Frames and Capture-Recapture Methods: Insights from a National Farmers Markets Managers Survey
Michael Jacobsen, Linda J. Young (National Agricultural Statistics Service)

Using a Web-Scraped List Frame for an Agricultural Survey
Habtamu Benecha (National Agricultural Statistics Service), Bruce A. Craig (Purdue University), Grace Yoon, Zachary Terner, Denise A. Abreu, Linda J. Young (National Agricultural Statistics Service)

Concurrent Session G
Wednesday, November 3rd
3:45 - 5:30 PM

Session G-1: Reply YES to Innovate: Text Messages for Federal Surveys

Session Liaison: Stephanie Coffey (U.S. Census Bureau)
Session Chair: Maura Spiegelman (National Center for Education Statistics)

Texting Interviewers to Encourage Proper Protocols in the Survey of Income and Program Participation
Kevin Tolliver (U.S. Census Bureau)

How Should We Text You? Designing and Testing Text Messages for the 2021-22 Teacher Follow-Up Survey (TFS) and Principal Follow-Up Survey (PFS)
Jonathan Katz, Kathleen Kephart, Jasmine Luck, and Jessica Holzberg (U.S. Census Bureau)

Yes, I consent to receive text messages: Conducting Follow-up Text Surveys with Principals and Teachers
Maura Spiegelman (National Center for Education Statistics) and Allison Zotti (U.S. Census Bureau)

The Future of SMS and Email in Federal Surveys
Jennifer Hunter Childs (U.S. Census Bureau)

Discussant: Frederick Conrad (University of Michigan)

Session G-3: Statistical Methods for Improving Small Domain and Key Survey Estimates

Session Liaison: Stephanie Coffey (U.S. Census Bureau)
Session Chair: Stephanie Coffey (U.S. Census Bureau)

A Practical Framework for Area-Level Small Area Estimation
Stephanie Zimmer, Dan Liao, and Rachel Harter (RTI International)

Using American Community Survey Data to Improve Estimates from Smaller U. S. Surveys through Bivariate Small Area Estimation Models
William R. Bell (U.S. Census Bureau) and Carolina Franco (National Opinion Research Center)

Using Bayesian Regression Models in Small Sample Size Contexts to Support System- Level Education Decision-Making
Bradley Rentz and Christina Tydeman (REL Pacific at McREL International)

Utilizing Occupational Employment and Wage Statistics (OEWS) Survey to Improve Small Domain Estimation (SDE) in the Occupational Requirements Survey (ORS)
Xingyou Zhang, Erin McNulty, Ellen Galantucci, Patrick Kim, Joan Coleman, and Tom Kelly (Bureau of Labor Statistics)

Statistical Data Integration using Multilevel Models to Predict Employee Compensation
Andreea L. Erciulescu, Jean D. Opsomer, and Benjamin J. Schneider (Westat)

Session G-4: Data Quality-Nonresponse Bias

Session Liaison: Tiandong Li (Health Resources and Services Administration)
Session Chair: Tiandong Li (Health Resources and Services Administration)

2017 Census of Agriculture Non-Response Sample
Mark Apodaca, Peter Quan, Franklin Duan, and Tim Keller (National Agricultural Statistics Service)

Subsampling to Reduce Nonresponse Bias in FoodAPS: A Simulation Study
Jeffrey M. Gonzalez (Economic Research Service), Joseph Rodhouse and Darcy Miller (National Agricultural Statistics Service)

Using Process Data to Study Interviewer Effects on Measurement Error and Nonresponse in the Consumer Expenditure Survey
John Dixon and Erica Yu (Bureau of Labor Statistics)

Who's Left Out?: Nonresponse Bias Assessment for an Online Probability-based Panel Recruitment
Frances Barlas and Randall K. Thomas (Ipsos)

Session G-5: Science and Engineering Indicators: Measuring S&E Education, Work Force, and Research Output

Session Liaison: Darius Singpurwalla (National Center for Science and Engineering Statistics)
Session Chair: Karen White (National Center for Science and Engineering Statistics)

Publications Output: U.S. Trends and International Comparisons
Karen White (National Center for Science and Engineering Statistics)

Science and Engineering Indicators Academic Research and Development
Josh Trapani (National Center for Science and Engineering Statistics)

Science and Engineering Indicators: Trends in U.S. Elementary and Secondary STEM Education
Susan Rotermund (RTI) and Amy Burke (National Center for Science and Engineering Statistics)

The STEM Workforce of Today: Scientists, Engineers and Skilled Technical Workers
Abigail Okrent and Amy Burke (National Center for Science and Engineering Statistics)

Discussant: Megan Fasules (Micronomics)

Concurrent Session H
Thursday, November 4th
8:30 - 10:15 AM

Session H-1: Data Collection in the Post-Pandemic Era

Session Liaison: Linda Young (National Agriculture Statistics Service)
Session Chair: Linda Young (National Agriculture Statistics Service)

Adapting Data Collection in a Pandemic – What adaptations will endure?
Barbara R. Rater (National Agricultural Statistics Service)

Beyond the Pandemic: Building New Programs at BTS
Rolf Schmitt (Bureau of Transportation Statistics)

COVID-19 Effects on the Health of Education Data and Implications for Future Education Data Collection
Chris Chapman (National Center for Education Statistics)

Session H-2: Using Administrative Data to Examine Food Assistance Program Effectiveness

Session Liaison: Paul Schroeder (Council of Professional Associations on Federal Statistics)
Session Chair: Laura Tiehen (Economic Research Service)

Estimating SNAP Eligibility and Access Using Linked Survey and Administrative Records
Renuka Bhaskar, Brad Foster, Brian Knop, and Maria Perez-Patron (U.S. Census Bureau)

Assessing SNAP Unit Simulations with Linked Survey and Administrative Data
Karen Cunnyngham and John Czajka (Mathematica Policy Research)

Investigating the Factors Behind High SNAP Participation Rate Estimates using Linked SNAP Administrative Data, CPS ASEC Data, and TRIM3 Microsimulation Model Estimates
Laura Wheaton (Urban Institute), Nancy Wemmerus and Tom Godfrey (Decision Demographics)

Using Administrative Data to Examine Cross-program Participation in SNAP and WIC
Leslie Hodges and Laura Tiehen (Economic Research Service)

Discussant: Constance Newman (USDA)

Session H-3: Modernizing Data Dissemination

Session Liaison: Darius Singpurwalla (National Center for Science and Engineering Statistics)
Session Chair: Joseph L. Parsons (National Agricultural Statistics Service)

Moving to a More User Centered Design for Data Dissemination at the US Department of Agriculture's National Agricultural Statistics Service
Bryan Combs, Jackie Ross, Elvera Gleaton, and King Whetstone (National Agricultural Statistics Service)

Modernizing Data Dissemination at the U.S. Bureau of Labor Statistics
Clayton Waring (Bureau of Labor Statistics)

Developing an Enterprise-wide Dissemination Program at the U.S. Census Bureau
Zachary Whitman (U.S. Census Bureau)

Discussant: Cynthia Parr (USDA, Research Education and Economics Mission Area)

Session H-5: Assessing & Improving the Quality of Prescription Drug Data from Surveys

Session Liaison: Jessica Graber (National Center for Health Statistics)
Session Chair: Sharon Arnold (Office of the Assistant Secretary for Planning & Evaluation)
Organizer: Steven C. Hill (Agency for Healthcare Research and Quality)

Improving Self-Reported Prescription Medicine Data Quality with a Commercial Database Lookup Tool and Claims Matching
Kali Defever, Becky Reimer, Michael Trierweiler, and Elise Comperchio (NORC at the University of Chicago)

Exploring Potential Benefits of Enumerating All Prescribed Medicines as a Tool for Estimating Opioid Use in the Medicare Current Beneficiary Survey (MCBS)
Becky Reimer, Elise Comperchio, Andrea Mayfield, and Jennifer Titus (NORC at the University of Chicago)

Evaluating Alternative Benchmarks to Improve Identification of Outlier Drug Prices for MEPS Prescribed Medicines Data Editing
Yao Ding and Steven C. Hill (Agency for Healthcare Research and Quality)

Discussant*:* Geoffrey Paulin (Bureau of Labor Statistics)

Concurrent Session I
Thursday, November 4th
10:30 AM- 12:15 PM

Session I-1: Survey Design & Measurement

Session Liaison: Yang Cheng (National Agricultural Statistics Service)
Session Chair: Tim Trudell (U.S. Census Bureau)

Blind to the Consequences of Measurement?: Response Format Effects on Self- Reported Disability
Megan A. Hendrich, Randall K. Thomas, and Frances M. Barlas (Ipsos Public Affairs)

Evaluating the National Agricultural Statistics Service's Grain Stocks Program: Results from Behavior Coding
Ashley Thompson and Heather Ridolfo (National Agricultural Statistics Service)

Examining Proxy Response Bias in a Large-Scale Survey of People with Disabilities
Eric Grau and Jason Markesich (Mathematica)

Integrating Previously Reported Data into The Census of Agriculture
Gavin Corral, Greg Lemmons, Linda J. Young, and Denise Abreu (National Agricultural Statistics Service)

Session I-3: Price Indices and Consumer Demand

Session Liaison: William Hahn (Economic Research Service)
Session Chair: William Hahn (Economic Research Service)

Consumer Prices During a Stay-In-Place Policy: Theoretical Inflation for Unavailable Products
Rachel Soloveichik (Bureau of Economic Analysis)

Democratic Aggregation: Issues and Implications for Consumer Price Indexes
Robert Martin (Bureau of Labor Statistics)

Session I-4: A Discussion of the Final Report of the Interagency Technical Working Group on Evaluating Alternative Measures of Poverty

Session Liaison: Paul Schroeder (Council of Professional Associations on Federal Statistics)
Session Chair and Organizer: Kevin Corinth (University of Chicago)

Presenters:

- Bruce Meyer, University of Chicago
- Thesia Garner, Bureau of Labor Statistics
- Liana Fox, U.S. Census Bureau

Discussants:

- Gary Burtless, Brookings Institution
- Laura Wheaton, Urban Institute

Session I-5: Machine Learning Applications

Session Liaison: Pam McGovern (National Agricultural Statistics Service)
Session Chair: Nathan Cruze (National Agricultural Statistics Service)

Machine Learning Assisted Complex Survey Weights
Stas Kolenikov (Abt Associates)

Using Machine-learning Algorithms to Improve Imputation in the Medical Expenditure Panel Survey
Chandler McClellan, Emily Mitchell, Jerrod Anderson, and Samuel Zuvekas (AHRQ)

An Overview of Business Establishment Automated Classification of NAICS (BEACON) for the Economic Census
Daniel Whitehead and Brian Dumbacher (U.S. Census Bureau)

Supplementing Cognitive Interviewing with Natural Language Processing Approaches from Data Science
Katherine Blackburn, Peter Baumgartner, Stephanie Eckman, David Henderson, Y. Patrick Hsieh, Patricia Green (RTI International) and Kelly Kang (National Center for Science and Engineering Statistics (NCSES)

For What It's Worth: Measuring Land Value in the Era of Big Data and Machine Learning
Scott Wentland and Gary Cornwall (Bureau of Economic Analysis), and Jeremy Moulton (University of North Carolina - Chapel Hill)

Concurrent Session J
Thursday, November 4th
1:30 - 3:15 PM

Session J-1: Modernizing Federal Survey Data Collections through Modularized Design

Session Liaison: Darius Singpurwalla (National Center for Science and Engineering Statistics)
Session Chair: Jennifer Sinibaldi (National Center for Science and Engineering Statistics)

NCSES' BAA Program and Emphasis on Modular Design Research
Jennifer Sinibaldi (National Center for Science and Engineering Statistics)

Investigating Modular Designs for the Survey of Doctorate Recipients
Ai Rene Ong (University of Michigan)

Developing and Evaluating Methodology for Split Questionnaire Design in the National Survey of College Graduates
Andy Peytchev (RTI)

Co-Designing a Smartphone App with Target Audience Members
Chris Antoun (University of Maryland)

Session J-2: Perseverance & Resilience: The Medical Expenditure Panel Survey in the Pandemic

Session Liaison: Tiandong Li (Health Resources and Services Administration)
Session Chair: Brad Edwards (Westat)

Managing Survey Change during the Pandemic
Brad Edwards and Rick Dulaney (Westat)

Assessment of the Effects of COVID-19 on Data Collected by the Medical Expenditure Panel Survey
Alisha Creel, Ralph DiGaetano, Hanyu Sun, Alexis Kokoska, and David Cantor (Westat)

From One, Many: Hatching a Multimode, Multiple-Respondent Supplement via a Household Interview
Darby Steiger and Angie Kistler (Westat) and Sandra Decker (Agency for Healthcare Research and Quality)

Interviewer Training for Classroom versus Distance Learning: Initial Skill Gains and Measures of Drift
Hanyu Sun. Angie Kistler, Ryan Hubbard, Brad Edwards, and Marcia Swinson-Vick (Westat)

Discussant: Joel Cohen (Agency for Healthcare Research and Quality)

Session J-3: After the Fact - Data Revision, Imputation, and Analysis Methods

Session Liaison: Richard Kluckow (Bureau of Justice Statistics)
Session Chair: Richard Kluckow (Bureau of Justice Statistics)

A Kriging Approach for Representing Crop Progress and Condition at Small Domains
Arthur Rosales (National Agricultural Statistics Service)

Coming Clean: Does Data Cleaning Reduce or Increase Bias in Sub-groups?
Randall K. Thomas, Frances M. Barlas, and Megan Hendrich (Ipsos Public Affairs)

Growing a Modern Edit and Imputation System
Darcy Miller,Vito Wagner, Travis Smith, Jeff Beranek, Lori Harper, Pamela Coleman, Karl Brown, James Johanson, and Megan Lipke (National Agricultural Statistics Service)

How Large are Long-run Revisions to U.S. Labor Productivity?
Peter B Meyer, John Glaser, Kendra Asher, Jay Stewart, and Jerin Varghese (Bureau of Labor Statistics)

Session J-5: The Reporting, Analysis, and Mitigation of Nonresponse Bias in Federal Surveys

Session Liaison: Maura Bardos (U.S. Energy Information Administration)
Session Chair: Stephen Blumberg (National Center for Health Statistics)

A Systematic Review of Nonresponse Bias Studies in Federally Sponsored Surveys
Tala H. Fakhouri (U.S. Food and Drug Administration), Jennifer Madans (National Center for Health Statistics), Peter Miller (U.S. Census Bureau) Morgan Earp (National Center for Health Statistics), Kathryn Downey Piscopo (US Substance Abuse and Mental Health Services Administration), Steven M. Frenk (National Institutes of Health), and Elise Christopher (National Center for Education Statistics)

FCSM Best Practices for Nonresponse Bias Reporting
Morgan Earp (National Center for Health Statistics),Jennifer Madans (National Center for Health Statistics), Elise Christopher (National Center for Education Statistics), Jenny Thompson (U.S. Census Bureau), Tala Fakhouri (U.S. Food and Drug Administration), Robert Sivinski (Office of Management and Budget), Kathryn Downey Piscopo (US Substance Abuse and Mental Health Services Administration), Joseph Schafer (U.S. Census Bureau), and Stephen Blumberg (National Center for Health Statistics)

Nonresponse Bias Analysis Methods: A Taxonomy and Summary
James Wagner (University of Michigan)

Nonresponse Bias Mitigation Strategies
Andy Peytchev (RTI)

Discussant: Robert Sivinski (Office of Management and Budget)

Concurrent Session K
Thursday, November 4th
3:30 - 5:15 PM

Session K-1: Previously Reported Data & Dependent Interviewing in Official Statistics: Existing Practices & New Findings

Session Liaison: Scott Wentland (Bureau of Economic Analysis)
Session Chair: Jeffrey M. Gonzalez (Economic Research Service)
Organizer: Joseph Rodhouse (National Agricultural Statistics Service)

Displaying Previously Reported Data to Respondents in the Quarterly Census of Employment and Wages Program at the Bureau of Labor Statistics
Emily Thomas (Bureau of Labor Statistics)

Testing Dependent Interviewing on a Self-Administered Survey
Jennifer Sinibaldi (National Center for Science and Engineering Statistics)

Evaluating Previously Reported Data in the Census of Agriculture: Results from Usability Testing
Heather Ridolfo and Kenneth Pick (National Agricultural Statistics Service)

Utilizing Respondents' Previously Reported Data in a Census of Establishments: Results from an Experiment in the Census of Agriculture's 2020 Content Test
Joseph Rodhouse, Kathy Ott and Zachary Terner (National Agricultural Statistics Service)

Discussant: Matt Jans (ICF)

Session K-4: Academic-Statistical Agency Collaboratives to Create Data Infrastructure for Evidence-Building

Session Liaison: Paul Schroeder (Council of Professional Associations on Federal Statistics)
Session Chair: Maggie Levenstein (ICPSR)

Re-Engineering Statistics using Economic Transactions (RESET)
Matthew D. Shapiro, Gabriel Ehrlich, and David Johnson (University of Michigan), John Haltiwanger (University of Maryland), and Ron Jarmin (U.S. Census Bureau)

Criminal Justice Administrative Records System: A Collaborative Approach to Building Next Generation Criminal Justice Data Infrastructure
Michael Mueller-Smith (University of Michigan) and Keith Finlay (U.S. Census Bureau)

Decennial Census Digitization and Linkage Project
Katherine Genadek (U.S. Census Bureau) and J. Trent Alexander (University of Michigan)

Agricultural and Food Data Systems for 2021 and Beyond
Brent Hueth (Economic Research Service) and Mark Denbaly (Economic Research Service)

Discussant: Daniel Goroff (Sloan Foundation)

Session K-5: Leveraging Data Science to Improve Survey Operations

Session Liaison: David Kashihara (Agency for Healthcare Research and Quality)
Session Chair: Carla Medalia (U.S. Census Bureau)

Machine Learning and the Commodity Flow Survey
Christian Moscardi (U.S. Census Bureau)

Parsing the Code of Federal Regulations for the Commodity Flow Survey's Hazardous Materials Supplement
Krista Chan and Christian Moscardi (U.S. Census Bureau)

Using PDF Extraction and Web Scraping Tools to Collect Government Health Insurance Plan Information
Virginia Gwengi (U.S. Census Bureau)

Using Open Source Tools to Build a Custom Data Entry Application for Creating Truth Data
Cecile Murray and Katie Genadek (U.S. Census Bureau)

Discussant: Kevin Deardorff (U.S. Census Bureau)

ABSTRACT BOOKLET

Session A-1
Survey Design Challenges During COVID-19

Analysis of Open-text Time Reference Web Probes on a COVID-19 Survey
Kristen Cibelli Hibben, *National Center for Health Statistics*
Valerie Ryan, *National Center for Health Statistics*
Travis Hoppe, *National Center for Health Statistics*
Paul Scanlon, *National Center for Health Statistics*
Kristen Miller, *National Center for Health Statistics*

To address the issue of time reference for Coronavirus pandemic related survey questions, we analyzed three open-ended web probes about when respondents think the pandemic began, when it first affected their lives, and why. We also assessed the quality of responses and if this differs by key socio-demographics. Data are from the Research and Development Survey (RANDS) during Covid-19 created at the National Center for Health Statistics. The National Opinion Research Center (NORC) at the University of Chicago collected data from June 9 to July 6, 2020 using their AmeriSpeak® Panel, a probability-based panel representative of the US adult English-speaking non-institutionalized population. The sample was supplemented with data from a non-probability online only opt-in sample (Dynata). All three probes were open text. A rule-based machine learning approach automated data cleaning. Hand review, topic modeling, and other computer-assisted approaches were used to examine response content and quality. Results show there is no uniform understanding of when the pandemic began and there is little alignment between when respondents think it began and when it first affected their lives. Preliminary data quality findings indicate most respondents gave valid answers to the two date probes, but the third probe had a wider range in response quality and variation among key population subgroups. This work sheds light on use of "since the Coronavirus pandemic began" as a time reference and the use of data science approaches for the analysis of open-ended web probes.

Applying Client Relationship Management Techniques to an Institutional Contacting Guide Redesign
Robert McCracken, *RTI International*
R. Suresh, *RTI International*
Brandon Peele, *RTI International*
Jason Kennedy, *RTI International*
Shawn Cheek, *RTI International*
Joe Nofziger, *RTI International*
Peter Tice, *Agency for Healthcare Research and Quality*

In 2019, RTI began redesign of a system used since 2010 to guide contacting of medical providers to help improve cooperation rates and provide more complete data. The MEPS MPC contact guide instrument collects medical providers' points of contact (POC) information for obtaining medical and billing records. Traditionally, these instruments were treated as survey instruments. We applied Client Relationship Management techniques to make it more efficient and intuitive. The redesign had three goals. First, to build POC confidence in the ease of the data collection process by providing a more POC-/client-focused system. Second, to improve flow at key points to increase efficiency and reduce errors. Several originally external components were integrated into the instrument. Third, to upgrade the instrument from a web-based platform to an application-based platform to align it with other project systems. In February of 2020, the MEPS MPC began using the redesigned POC module. In this presentation, we will discuss the conversion process and describe the results. Anecdotal feedback indicates that tracking of records status by POC is easier with the new system, but the flow improvements have had mixed results. Changes brought about by the COVID-19 pandemic have made full evaluation of the redesign difficult.

Cross Agency Collaboration in the Rapid Development of COVID-19 Questionnaire Items for the Medicare Current Beneficiary Survey (MCBS)
Andrea Mayfield, *NORC at the University of Chicago*
Kylie Carpenter, *NORC at the University of Chicago*
Rachel Carnahan, *NORC at the University of Chicago*
Elise Comperchio, *NORC at the University of Chicago*
Felicia LeClere, *NORC at the University of Chicago*

The Medicare Current Beneficiary Survey (MCBS) is a survey of a nationally representative sample of the Medicare population sponsored by the Centers for Medicare & Medicaid Services (CMS) through a contract with NORC at the University of Chicago. Starting in Summer 2020, CMS leveraged the MCBS design to assess the impact of the COVID-19 pandemic on beneficiaries' lives by planning rapid response surveys to supplement the main MCBS. The surveys included items related to beneficiaries' experiences with COVID-19, such as preventive health behaviors, access to and use of telemedicine, COVID-19 testing, attitudes towards vaccination, and, starting in Winter 2021, vaccination uptake. Collaboration with the National Center for Health Statistics, who was also developing questions for their Research and Development Survey, facilitated rapid development and subsequent improvements of the MCBS COVID-19 Supplement. This presentation will discuss the cross-agency collaborative process used to ensure the rapid implementation of tested COVID-19 items on the MCBS. We will also share estimates for measures from the Supplements and explore differences by factors such as age, sex, and race/ethnicity.

Missing Price Imputation in the Producer Price Index
Yoel Izsak, *Bureau of Labor Statistics*

The Producer Price Index at the Bureau of Labor Statistics currently uses cell mean imputation for missing price data. In the time since the implementation of the current process, multiple imputation methods have become much easier to use on large data sets. In this study, we investigate alternatives to the current procedure. We examined a few different multiple imputation methods with packages in R, including: MICE (Multiple Imputation using Chained Equations), Random Forest, AMELIA (bootstrap EM algorithm), and MI. We tested each method over many simulated data sets, using the past few years of PPI price data. Success of imputation for the missing prices was measured by RMSE (Root Mean Squared Error), as well as Index estimates. Results from the study will be discussed.

National Ambulatory Medical Care Survey: Redesigning for the Future Health Care Landscape
Sonja N. Williams, *National Center for Health Statistics*
Jill J. Ashman, *National Center for Health Statistics*
Brian W. Ward, *National Center for Health Statistics*

The National Ambulatory Medical Care Survey (NAMCS), conducted by the National Center for Health Statistics, is designed to meet the need for objective, reliable information about the provision and use of ambulatory medical care services in the United States. Since NAMCS began fielding in 1973, drastic changes have occurred in the provision of health care. A lot of care is now provided by an array of advanced practice providers and large health networks, across more diverse care settings, and relies heavily on the use of electronic health records (EHRs). These changes, along with declining response rates and the COVID-19 pandemic, have created a need for NAMCS to be redesigned to remain the premier source of data on ambulatory care in the United States. This presentation will discuss the ongoing redesign of NAMCS, including consultation from stakeholders and ongoing changes being implemented that will provide more timely, accurate, and nimble data. Success of some early changes to the survey will be shown, including the release of the survey's first ever preliminary estimates that focus on physicians' experiences providing care during the COVID-19 pandemic.

Session A-2
Panel: Re-Inventing and Integrating Annual Economic Survey Programs at the U.S. Census Bureau

Developing an Integrated Annual Survey
Jenna Morse, *U.S. Census Bureau*
James Burton, *U.S. Census Bureau*
Kimberly Moore, *U.S. Census Bureau*

The U.S. Census Bureau maintains a suite of annual economic surveys that present challenges to controlling respondent burden and creating cross-program data products due to their existing structure. The Bureau requested a NAS Panel study to obtain recommendations for reengineering the surveys to address these limitations and better meet future goals. The panel's overarching recommendation was that the Bureau develop and implement an Integrated Annual Survey.

This paper outlines the scope of the NAS Panel study and its key recommendations and provides an overview of efforts to meet the scope of the project based on methodological research and extensive testing, along with feedback from data users. Objectives include research and implementation of improved frame, sample and survey methodology, including standardization of survey units, reduction and harmonization of content, creation of new collection modules, increased use of alternative data, , and coordination of data collection at a company level. The paper also points out challenges and provides target goals and milestones. The other papers in this session will go into detail on some of these individual topics.

Aiding a Holistic View of Businesses through Intensive Respondent Research
Diane K. Willimack, *U.S. Census Bureau*

A holistic view of businesses is fundamental to the Integrated Annual Survey. This requires in-depth understanding of business survey respondents' reporting behaviors, perspectives, and response burden, based on multi-method research into:

- individualized reporting arrangements with large complex companies that annually receive six or more different surveys
- business roundtables and focus groups to discern beliefs and attitudes that impact response motivation
- exploratory interviews with medium-sized companies investigating record-keeping practices and data availability at different company units/levels
- testing field periods and contacts coordinated across multiple annual surveys to gain practical insights for successful implementation.

Synthesis of results aided decisions about content harmonization, determination of sampling units, reporting units, and collection units, and development of a modular data collection strategy. Extensive cognitive testing is underway to evaluate topic-based structure and timing of questionnaire modules. Research will culminate with field testing the integrated survey, and full-scale implementation is planned for calendar year 2024.

Determining and Harmonizing Content in Developing the Annual Integrated Economic Survey
Blynda Metcalf, *U.S. Census Bureau*
Heidi St. Onge, *U.S. Census Bureau*
Kimberly Moore, *U.S. Census Bureau*

As part of an effort to transform and modernize statistical programs, the U.S. Census Bureau is re-engineering six annual economic surveys to develop a new integrated annual program, with implementation planned for calendar year 2024, for reference year 2023. Economic survey analysts engaged in a rigorous review and evaluation of existing content collected in our annual surveys. Large scale efforts were performed to harmonize and standardize content that overlapped surveys, identify data gaps through consultations with stakeholders, and design new economy-wide data products. This paper will provide a high-level overview of the Census Bureau's process to engage stakeholders, determine integrated questionnaire content, and harmonize metadata across trades while considering data user needs. We will outline the proposed collection approach to obtain a more holistic view of companies through company-centric survey design and collection instruments. The paper will also highlight proposed new economy-wide data products, including expanded geographic and North American Product Classification System (NAPCS) data across the manufacturing, retail, services, and wholesale sectors.

Developing a Unified Sample Design for the Integrated Annual Survey
Katherine Jenny Thompson, *U.S. Census Bureau*
James Burton, *U.S. Census Bureau*
James Hunt, *U.S. Census Bureau*
Amy Newman Smith, *U.S. Census Bureau*

This paper provides a high-level overview of the proposed sample design for the Integrated Annual Survey. The unified sample replaces our directorate's existing practice of independently developing sampling frames and sampling procedures for a suite of separate annual surveys. The replacement design is a "work-in-progress," developed in parallel with the other re-engineering and research projects discussed in this session. As such, the design requirements are informed by the user community's data needs including the addition of sub-national (geographic) tabulations in selected sectors and harmonized item definitions as well as the respondent research on collection units. Before sketching out the sample design, we briefly discuss the challenges of establishing unique sampling units along with a single measure-of-size for a multi-purpose business survey whose collection covers a wide range of economic sectors, then describe issues in sample maintenance, expected reliability, and respondent attribution that might be mitigated via the sample design. We conclude by describing ongoing research activities commissioned to inform and support any design decisions.

Session A-3
Recent Advances in Disclosure Limitation and Data Privacy

Disclosure Limitation in the Census of Fatal Occupational Injuries
Ellen Galantucci, *Bureau of Labor Statistics*

The Census of Fatal Occupational Injuries (CFOI) and the Bureau of Labor Statistics (BLS) collects a complete census of all fatalities that occur during a calendar year as the result of workplace injuries. CFOI data present unique disclosure challenges because there are a small number of fatalities per year (5,333 in 2019) and many fatalities are reported publicly. For any information that BLS collects from confidential sources, it is required to protect the collected data. With the recognition that CFOI was not adequately protecting its confidential data, we set out to update the methodology. In this presentation, I outline the options that we considered for disclosure limitation and the benefits and drawbacks of each. I will also present the results of a reconstruction attack attempted on the CFOI data that was conducted to evaluate disclosure risk.

Synthetic Public Use File of Administrative Tax Data: Methodology, Utility, and Privacy Implications
Claire McKay Bowen, *Urban Institute*

US government agencies possess data that could be invaluable for evaluating public policy, but often may not be released publicly due to disclosure concerns. For instance, the Statistics of Income division (SOI) of the Internal Revenue Service releases an annual public use file of individual income tax returns that is invaluable to tax analysts in government agencies, nonprofit research organizations, and the private sector. However, SOI has taken increasingly aggressive measures to protect the data in the face of growing disclosure risks, such as a data intruder matching the anonymized public data with other public information available in nontax databases. In this talk, we describe our approach to generating a fully synthetic representation of the income tax data by using sequential Classification and Regression Trees and kernel density smoothing. We also tested and evaluated the tradeoffs between data utility and disclosure risks of different parameterizations using a variety of validation metrics by comparing to the confidential data and the original sanitized PUF.

Accuracy Gains from Privacy Amplification Through Sampling for Differential Privacy
Jingchen Hu, *Vassar College*
Joerg Drechsler, *University of Maryland*
Hang J. Kim, *University of Cincinnati*

Recent research in differential privacy demonstrated that (sub)sampling can amplify the level of protection. For example, for ε-differential privacy and simple random sampling with sampling rate r, the actual privacy guarantee is approximately rε, if a value of ε is used to protect the output from the sample. In this talk, we present results from our study about whether this amplification effect can be exploited systematically to improve the accuracy of the privatized estimate. Specifically, assuming the agency has information for the full population, we ask under which circumstances accuracy gains could be expected, if the privatized estimate would be computed on a random sample instead of the full population. We find that accuracy gains can be achieved for certain regimes. However, gains can typically only be expected, if the sensitivity of the output with respect to small changes in the database does not depend too strongly on the size of the database.

A Latent Class Modeling Approach for DP Synthetic Data Contingency Tables
Andres Felipe Barrientos, *Florida State University*
Michelle Pistner Nixon, *Pennsylvania State University*
Jerome P. Reiter, *Duke University*
Aleksandra Slavkovic, *Pennsylvania State University*

We present an approach to construct differentially private synthetic data for contingency tables. The algorithm achieves privacy by adding noise to selected summary counts, e.g., two-way margins of the contingency table, via the Geometric mechanism. We posit an underlying latent class model for the counts, estimate the parameters of the model based on the noisy counts, and generate synthetic data using the estimated model. This approach allows the agency to create multiple imputations of synthetic data with no additional privacy loss, thereby facilitating estimation of uncertainty in downstream analyses. We illustrate the approach using a subset of the 2016 American Community Survey Public Use Microdata Sets.

Improving the Utility of Poisson-Distributed, Differentially Private Synthetic Data via Prior Predictive Truncation with an Application to CDC WONDER
Harrison Quick, *Drexel University*

CDC WONDER is a web-based tool for the dissemination of epidemiologic data collected by the National Vital Statistics System. While CDC WONDER has built-in privacy protections, they do not satisfy formal privacy protections such as differential privacy and thus are susceptible to targeted attacks. Given the importance of making high-quality public health data publicly available while preserving the privacy of the underlying data subjects, we aim to improve the utility of a recently developed approach for generating Poisson-distributed, differentially private synthetic data by using publicly available information to truncate the range of the synthetic data. Specifically, we utilize county-level population information from the U.S. Census Bureau and national death reports produced by the CDC to inform prior distributions on county-level death rates and infer reasonable ranges for Poisson-distributed, county-level death counts. In doing so, the requirements for satisfying differential privacy for a given privacy budget can be reduced by several orders of magnitude, thereby leading to substantial improvements in utility. To illustrate our proposed approach, we consider a dataset comprised of over 26,000 cancer-related deaths from the Commonwealth of Pennsylvania belonging to over 47,000 combinations of cause-of-death and demographic variables such as age, race, sex, and county-of-residence and demonstrate the proposed framework's ability to preserve features such as geographic, urban/rural, and racial disparities present in the true data.

Session A-4
Data Quality: Communication of Uncertainty in Official Statistics

Evaluating Uncertainty in Multiple Dimensions of Data Quality
John L. Eltinge, *U.S. Census Bureau*

Evaluation of the quality of official statistics generally requires balanced consideration of multiple dimensions, e.g., accuracy, relevance, timeliness, granularity, comparability, interpretability and accessibility. In-depth study of the "accuracy" dimension for sample surveys has led to a substantial literature on "Total Survey Error" (TSE) models. More recently, several authors have explored extensions of TSE models to cases involving the integration of data from multiple sources, e.g., administrative records and commercial transactions, as well as sample surveys. This paper moves beyond the "accuracy" dimension to consider further extensions of the TSE approach that account for the other abovementioned dimensions of data quality in the integration of multiple sources. Strengths and limitations of both qualitative (descriptive) and quantitative (model oriented) approaches are considered. These ideas are illustrated with examples that involve appending administrative record data to sample survey units ("relevance" and "comparability" dimensions); extensions of multiple-frame, multiple-mode methods to the integration of administrative or commercial records ("timeliness," "granularity" and "interpretability" dimensions); and data subject to disclosure-protection methods ("accessibility" dimension).

More Fully Capturing Uncertainty Associated with Official Estimates
Linda J. Young, *National Agriculture Statistics Service*

The Total Survey Error approach has made practitioners aware of sources survey error, such as sampling variability, interviewer effects, frame errors, response bias, and non-response bias. However, most measures of uncertainty published with estimates, including official estimates, fail to capture all of these sources of error, which generally results in overly optimistic reports of uncertainty. For the 2012 Census of Agriculture, a capture-recapture methodology was adopted to produce official estimates at all levels of geography. The estimates were based on four logistic models. In addition, the modeled results were calibrated to known population values, and the final results were rounded to integers. The uncertainty associated with fitting each model and the modeling error were captured, but the additional uncertainty from calibration and rounding was not. In 2017, the impact of calibration and rounding was incorporated in all measures of uncertainty. In this paper, the effect of accounting for these additional sources of variability is presented. Additional sources of uncertainty that have yet to be incorporated in the published results are reviewed, and the potential for including these in the future is discussed.

Tailored Transparency: Public Trust vs. Reproducibility
Peter V. Miller, *Northwestern University and U.S. Census Bureau (Retired)*

The OMB Standards and Guidelines for Statistical Surveys, the Census Bureau's Statistical Quality Standards and other published principles recognize the importance of transparency in survey research. The AAPOR Code states that, "Good professional practice imposes the obligation on all public opinion and survey researchers to disclose sufficient information about how the research was conducted to allow for independent review and verification of research claims." Enabling expert review is a key purpose of this provision, but disclosure also is intended to facilitate better general understanding of surveys among less sophisticated audiences, and to foster trust in the research enterprise. The different purposes and audiences for methodological information suggest a tailored approach to transparency. Such an approach becomes increasingly useful as survey methods evolve and incorporate other data sources, increasing the scope and complexity of disclosure practice. The very different transparency demands of the general public and a scientific community focused on information needed to reproduce survey estimates also impel a tailored approach to disclosure. This paper discusses tailored disclosure of survey information that takes into account different purposes and audiences for the documentation. The discussion considers surveys conducted to

produce official statistics, with recommendations on how transparency may be approached for the wide variety of consumers of information from these key research efforts.

Session A-5
Using Multiple Survey Modes or Administrative Data to Improve Estimates

It's All a Matter of Degrees: Comparing Survey and Administrative Educational Attainment Data
Andrew Foote, *U.S. Census Bureau*
Larry Warren, *U.S. Census Bureau*

This paper measures the accuracy of survey responses on educational attainment and field of study in the American Community Survey, using linked administrative data on degree attainment. We find that 5-10% of respondents with bachelor's degrees under-report their education level, while an average of 20% of respondents mis-report their field of study. We also show that misreporting is higher for Black non-Hispanic and Hispanic respondents, and that some of the misreporting is due to the imputation procedures used by ACS to correct for item non-response. We also show that total error in response increased in 2013 and is driven by increases in the use of the hot-deck imputation procedure.

Implementing a Multi-Mode Approach to Household Eligibility Screening for 2021-2022 NHANES
Juliana McAllister, *National Center for Health Statistics*
Allan Uribe, *National Center for Health Statistics*
Jessica Graber, *National Center for Health Statistics*
Denise Schaar, *National Center for Health Statistics*
Chia-Yih Wang, *National Center for Health Statistics*

The National Health and Nutrition Examination Survey (NHANES) is a nationally representative survey that historically determined eligibility through in-person household screening. In the 2021-2022 survey period, NHANES introduced a multi-mode screening approach to also include self-response (web and mail) and interviewer-administered (telephone and in-person) modes. Reasons for implementing the multi-mode screening were two-fold; minimizing in-person contact between respondents and field staff due to safety concerns in the COVID-19 environment; and increasing field efficiency by reducing the resources spent on in-person screening, leaving more time available to focus on tasks targeted to identified eligible respondents such as gaining cooperation, refusal conversion, and interview administration. This presentation describes the implementation process and challenges in operationalizing a multi-mode screening and presents response rates by each of the four screener modes (web, paper, telephone, in-person) using preliminary data collected from one county during the period of about 2 months. The impact of each subsequent mailing will be examined along with paradata from the web screener responses. Findings on respondent burden, by mode, will also be summarized.

A Meta-Analysis of the Impact of Number of Contact Attempts on Response Rates and Web Completion Rates in Multimode Surveys

Ting Yan, *Westat*

Due to decreasing response rates and increasing data collection cost, surveys employing a multimode design are on the rise. Multimode surveys use multiple modes (such as mail, web, telephone, or in-person) to contact potential respondents and collect data from respondents. Multimode surveys are used to improve coverage, increase response rates, reduce costs, and improve measurement (De Leeuw, 2018; Tourangeau, 2017). Researchers and practitioners employing multimode surveys need to make informed decisions on what modes are to be used in the design and in what order, whether multiple modes are offered concurrently or sequentially, and how many contacts are to be attempted, and so on. The number of contact attempts is an important design decision for multimode surveys employing web and mail. The primary objective of this paper is to use meta-analytic methods to examine the trend in response rates to multimode surveys using web and mail. In particular, we are interested in quantifying the effect of the number of contact attempts on response rates and web completion rates in multimode surveys.

To achieve this goal, we conducted a literature search and identified empirical studies employing a web and mail design. We coded various survey characteristics for each study including survey topic, sampling frame, population of interest, type of multimode design, number of contact attempts, use of incentives, whether a pre-notification letter/email was sent, overall response rate, and web completion rate. Then we will compute an overall effect size of the number of contact attempts on response rates and web completion rates. We will then conduct moderator analysis to see if the effect size varies by other survey characteristics (such as the use of incentives). The findings will inform the survey field of optimal designs of multimode surveys employing mail and web.

Session B-1
Measuring Sexual Orientation & Gender Identity

Best Practices for Collecting Gender and Sex Data
Wendy Martinez, *Bureau of Labor Statistics*
Suzanne Thornton, *Swarthmore*
Dooti Roy
Stephen Perry
Donna LaLonde,
Renee Ellis,
David J. Corliss

Researchers and practitioners look to statisticians to provide useful and practical guidelines for collecting data on sex and gender. In 2020, a group representing diverse areas of statistical practice came together to create a document that compiled definitions of terms and examined the necessary context to understand gender identities and sexual characteristics. Through the use of examples, the work presents the statistical and ethical considerations of inclusive data collection. The authors also discuss the important statistical considerations that should inform the data collection plan. In this talk, I will provide an overview of the document and will focus on how it relates to other efforts to collect SOGI data in the federal government.

Measuring Sexual Orientation and Gender Identity Among College Graduates: Results from a Methodological Experiment Using a Non-Probability Sample
Rebecca L. Morrison, *National Center for Science & Engineering Statistics*
Jesse Chandler, *Mathematica*
Flora Lan, *National Center for Science & Engineering Statistics*
Karen S. Hamrick, *National Center for Science & Engineering Statistics*

Federal agencies are facing increasing calls for data concerning sexual and gender minorities in American society and its workforce. In this presentation, NCSES will report results from two methodological experiments focusing on sexual orientation and gender identity (SOGI) questions; the study was conducted on a convenience sample recruited from Amazon MTurk in Spring 2021. In the first experiment, all respondents received the same question concerning sexual orientation, but received one of four sets of response options. The purpose of this experiment was to examine possible measurement error effects resulting from the response options. In the second experiment, which examines possible context effects for the two-step gender identity questions, all respondents received a question asking for sex at birth, and current gender identity. However, the order of those two questions varied across respondents. Finally, the authors will present findings from a question that asked about the pronouns used by respondents (e.g., she/her, he/him, they/them), as well as findings from follow-up questions aimed at measuring comprehension and sensitivity of the SOGI items.

Assessing Measurement Error in Sex and Gender Identity Measures in an NCES Survey
Elise Christopher, *National Center for Education Statistics*

Over several decades, the National Center for Education Statistics has collected data from high school students and their families, including demographic information such as gender. In previous iterations of longitudinal surveys, sample members may have updated their demographics between rounds. It is not known the extent to which these updates reflect measurement error or changes in self-identification of sample members. In the case of gender, NCES evaluated and employed a two-step measure of sex at birth and gender identity on a later round of a longitudinal survey, the High School Longitudinal Study of 2009 (HSLS:09). To understand the potential for measurement error, several types of data were compared, including sample member responses across modes, between rounds, and administrative records. Analyses in

this paper focus on whether measurement error decreased with the two-step method, which is a primary goal of collecting gender information in this way.

Session B-2
Impact of the COVID-19 Pandemic on National Center for Health Statistics Data Collections

Impact of the Pandemic on National Health Interview Survey Data Collection
Stephen J. Blumberg, *National Center for Health Statistics*

The National Health Interview Survey (NHIS) is the longest-running household-based health survey in the US. On March 19, 2020, NHIS temporarily changed from an in-person survey to a telephone survey. Where possible, sample addresses were matched to telephone numbers using commercial lists and additional searches. Response rates declined and the sample skewed toward older and more affluent households. Personal visits resumed in selected areas in July and in all areas in September. However, cases were still attempted by telephone first. NHIS also fielded the 2020 questionnaire with a parallel sample: some adult respondents who completed the NHIS in 2019 were recontacted by phone and asked to participate again. These follow-up interviews permit NHIS to look at health, health care, and well-being from before and during the pandemic, for the same individuals. Challenges to be discussed include how weighting and estimation techniques were used to produce official 2020 estimates from the disparate pieces — normal production in Quarter 1, telephone-only in Quarter 2, telephone-first in Quarters 3 and 4, and the longitudinal follow-up — each with its own coverage and nonresponse issues.

COVID-19 Pandemic Impact on the NCHS Provider-Based Surveys
Carol DeFrances, *National Center for Health Statistics*
Brian W. Ward, *National Center for Health Statistics*
Manisha Sengupta, *National Center for Health Statistics*

To continue generating data from ambulatory, hospital, and long-term care providers during the midst of the COVID-19 pandemic, the National Center for Health Statistics (NCHS) had to modify survey operations for several of its provider-based surveys. This included quickly adding survey questions that captured the experiences of providing care during the pandemic. This presentation provides an overview of the effect of the pandemic on each of these provider surveys: the National Ambulatory Medical Care Survey, National Electronic Health Records Survey, National Hospital Ambulatory Medical Care Survey, National Hospital Care Survey, and National Post-acute and Long-term Care Study. Highlighted are some key challenges that impacted data collection activities for these surveys, as well as the measures taken to minimize the disruption in data collection, and optimize the disseminating of quality data in a timely manner including the release of preliminary estimates on the COVID-19 pandemic for selected surveys.

The National Health and Nutrition Examination Survey (NHANES): The Impact of the COVID-19 Pandemic on Data Collection and Release

Ryne Paulose-Ram, *National Center for Health Statistics*
Jessica E. Graber, *National Center for Health Statistics*
Namanjeet Ahluwalia, *National Center for Health Statistics*

NHANES is a unique source of national data on the health and nutritional status of the US population, collecting data through interviews, in-person examinations and biospecimen collection. Due to the COVID-19 pandemic, data collection in 2019-2020 could not be completed on the target sample. A statistical approach was developed to create a 4-year nationally representative sample, by combining the 2019-2020 data with NHANES 2017-2018 data, when possible. These 2017-2020 pre-pandemic data files are planned for release in 2021. New content collected in 2019-2020 will also be released as convenience samples via the Research Data Center. Due to COVID-19, the NHANES 2021-2022 survey was also delayed. This cycle was redesigned to focus on key components, emphasizing staff and participants' safety in a COVID-19 environment. Changes include alternative data collection modes, reduced survey time and respondent burden, and new COVID-related content. Physical design changes to the mobile examination center to increase social distancing were also made. The challenges and implications of the release of NHANES 2017-2020 pre-pandemic data and the redesigned 2021-2022 survey will be presented.

Expanding the NCHS Research and Development Survey During the COVID-19 Pandemic

Jennifer D. Parker, *National Center for Health Statistics*

The National Center for Health Statistics' (NCHS) Research and Development Survey (RANDS) is a series of commercial panel surveys collected for methodological research. In response to the COVID-19 pandemic, NCHS expanded the use of the RANDS platform, fielding three rounds of RANDS during COVID-19 to evaluate COVID-19 related survey question interpretation and performance and to produce a set of experimental estimates for public release on work loss due to illness with COVID-19, telemedicine access and use, and reduced access to health care. This presentation will describe the use of RANDS to provide information during the pandemic, including selected COVID survey question evaluations and the calibration to NCHS' National Health Interview Survey (NHIS) to adjust for potential bias in the panel, and will discuss implications of the release of experimental estimates.

Concurrent Session B-3
Sampling and Calibration for Data Versatility

An Easy Way to Calibrate on Partly Known Overlapping Multiple Totals in Frequency Tables with Application to Real Data
Michael Sverchkov, *Bureau of Labor Statistics*

Deville and Sarndall (1992, Section 4) considered calibration on the known counts (cell counts or marginal counts) of a frequency table in any number of dimensions (generalized raking procedure). In this paper, we show that a similar procedure can be applied to the case of partly known overlapping counts. As an example we consider calibration of area-month-year unemployment estimates to month-year totals from a time series model of State estimates from the Current Population Survey and area-year totals from the American Community Survey

Designing a Probability Sample to Produce a Large Number of Key Estimates
Phillip Kott, *RTI International*
Darryl V. Creel, *RTI International*

Probability proportional-to-size (pps) sampling can be an efficient sampling methodology for a survey when the measure of size (mos) of a sampling unit is roughly proportional to the variable of interest. In practice, however there are many variables of interest collected on a survey but only a single mos for pps sampling. How can we find a viable sampling approach when there may be many equally important variables of interest with its own associated mos? A popular sampling approach is to pick a potential mos very roughly proportional to several of the variables of interest, but this can be complicated. Our solution is to use maximal probability proportional-to-size (mpps) sampling. In mpps sampling, for each variable of interest and associated mos, one first calculates a selection probability for a sampling unit under pps sampling. Then, for each sampling unit, one chooses the largest of these variable-specific selection probabilities: the maximal probability proportional-to-size. This paper describes the theoretical background of, addresses some of the practical issues involved with, and examines an example of mpps sampling.

Sample-based Calibration of Multiple Surveys
Jean Opsomer, *Westat*
Weijia Ren, *Westat*
John Foster, *NOAA Fisheries*

NOAA Fisheries' Marine Recreational Information Program (MRIP) combines data from multiple surveys to estimate the volume and characteristics of fish caught by recreational anglers. These data are used in stock assessment modeling and in setting allowable catch limits for marine fish species. A key step in the creation of survey datasets is the calibration of the weights from intercept surveys collecting information on individual trips, to control totals from off-site surveys estimating the number of trips. In order to capture the control total variability into the variance estimation procedures for the intercept surveys, we developed a sample-based calibration approach that is similar to that proposed in Fuller (1998) in the two-phase context. The method is readily applicable to other survey calibration settings in which control totals are random. We describe some of the potential pitfalls encountered in implementing sample-based calibration as well as proposed solutions.

Sampling and Estimation for Multipurpose Surveys
Yang Cheng, *National Agricultural Statistics Service*
Jeff Bailey, *National Agricultural Statistics Service*
Eric Slud*, U.S. Census Bureau and University of Maryland*
Lu Chen, *National Agricultural Statistics Service and National Institute of Statistical Sciences*

Many agricultural surveys have multipurpose. Each population unit has many study variables. When samples are drawn from the population, each sample unit may not contain all study variables. As a result, some study variables may not have sufficient responses to meet the survey precision requirement. Since 1996, the National Agricultural Statistics Service (NASS) has adopted the Multivariate Probability Proportional to Size (MPPS) sampling design to deal with this challenge. A ratio or regression estimator for the totals or means of several study variables is employed, and a Delete-A-Group Jackknife (DAGJK) method is used to develop measures of uncertainty. In this paper, the current NASS sample design, estimation, and variance estimates are investigated. An alternative approach is proposed. Data from an agricultural survey are used to evaluate the effectiveness of the alternative strategy.

Sampling Using Multiple Measures of Size: A Simulation Study
Tim Keller, *National Agricultural Statistics Service*
Mark Apodaca, *National Agricultural Statistics Service*
Franklin Duan, *National Agricultural Statistics Service*

The efficiency of probability proportional to size (PPS) sample designs for estimating population totals is well documented. However, it is often the case that a single measure of size is not equally effective for estimating population totals for multiple items of interest. The conventional approach to address this situation is known as multivariate probability proportional to size (MPPS) sampling, which generalizes the PPS method by taking the inclusion probabilities to be the maximum of the inclusion probabilities corresponding to multiple measures of size.
The imperative of reducing respondent burden means that one seeks to achieve targets for the precision of estimates with the smallest possible sample size. In the context of multiple measures of sizes, there may be more efficient sample designs than the conventional MPPS. Alternative sample designs for which a PPS sample based on a composite measure of size, which is a function of several given measures of size, are explored and compared with the MPPS design using simulated data.

Session B-4
Equity

Policy Review of Initiatives to Improve Equity Information
Jamie Keene, *Executive Office of the President*

Educational Equity: Identifying and Presenting Information within New Online Resources
Ross Santy, *National Center for Education Statistics*

Present on development of the Education Equity Dashboard, what motivated the development and what informed it

Healthy People: Exploring Disparities in the Nation's Health
David Huang, *National Center for Health Statistics*

Present on development of the Healthy People 2020 and 2030 Disparities Tools, what motivated the development and what informed it

Delivering Equity in Federal Forms and Surveys: LGBTQ+ Data Collection
Amy Paris, *Department of Health and Human Services*

Concurrent Session B-5
Topics in Survey Data Quality

Survey Isolation during COVID-19: The Effects of Suddenly Relying on Address-Matched Phone Numbers for Interviewing Households

Jonathan Eggleston, *U.S. Census Bureau*
Yarissa Gonzalez, *U.S. Census Bureau*
Tim Trudell, *U.S. Census Bureau*
John Voorheis, *U.S. Census Bureau*

The COVID-19 pandemic greatly impacted data collection for household surveys, curtailing in-person interviewing across the world. For example, the U.S. Census Bureau conducted numerous in-person surveys before the pandemic, but temporarily switched to phone data collection for these surveys in March of 2020. To contact newly-sampled households, the Census Bureau relied on vendor-provided phone numbers matched to addresses. This sudden change in interviewing protocol and generally the effects of the pandemic on respondent behavior may have profound effects on who is able to be contacted for Census Bureau surveys (noncontact bias) and who is willing to respond conditional on being contacted (refusal bias). To investigate these biases, we use the Current Population Survey and the American Community Survey. We first use administrative data matched to sampled addresses to see how the characteristics of respondents and nonrespondent in 2020 and 2021 compare to prior years. Next, we focus specifically on the extent to which phone number match quality can explain changes in noncontact bias, as these phone numbers were the primary method of contact. Preliminary results suggest that the lower a household's income is, the less likely the household is to have their phone numbers correctly matched to their current address. This correlation potentially introduces a nonresponse bias with respect to income that was not present before in these surveys. We discuss the reasons for some households having lower match quality as well as a possible correction for this bias through weighting.

The New Non-employer Business Demographics Statistics: Responding to 20th-century Survey-based Statistics Challenges while Addressing 21st-century Needs

Adela Luque, *U.S. Census Bureau*
Ken Rinz, *U.S. Census Bureau*
James Noon, *U.S. Census Bureau*
Michaela Dillon, *U.S. Census Bureau*

The new Nonemployer Statistics by Demographics series or NES-D is the Census Bureau's response to the challenges faced by 20th-century survey-based statistics while addressing 21st-century needs for more frequent and timely high-quality data, at lower cost and no additional respondent burden. NES-D is not a survey; rather, it is an annual statistical series that exclusively uses existing administrative and census records to provide demographics for the universe of nonemployer businesses by geography, industry, receipt size class and legal form of organization. Its first release was December, 2020. NES-D replaces the nonemployer component of the quinquennial Survey of Business Owners. Coupled with the new Annual Business Survey (ABS), which provides demographics for employer businesses, Census now provides annual business owner demographics through a blended-data approach that combines AR-derived estimates for nonemployer firms and survey-derived estimates for employer firms. In the near future, NES-D will be enhanced with characteristics relevant to understanding nonemployers' behavior and dynamics, such as characteristics related to the gig economy, household characteristics and transitions to employer status. NES-D exemplifies what results can be accomplished with well-researched administrative records and census data, the application of sound methodologies, the drive to address users' needs, and strong collaborations with stakeholders.

Participation Metrics for Accelerometer-Based Research
Christopher Antoun, *University of Maryland*
Alexander Wenz, *University of Mannheim*

Researchers are increasingly using accelerometer devices rather than surveys to measure physical activity (PA). However, a key challenge in accelerometer-based studies is nonparticipation. Individuals may decline to participate, not wear the devices throughout the full measurement period, or not return the devices, among other things. This will reduce the accuracy of sensor-based results if the subset of individuals who do participate are not representative of the broader population to which the results are intended to generalize. Perhaps because researchers from multiple disciplines are using accelerometer-based methods, no common terminology has emerged for different types of non-participation in this context. Our paper attempts to propose standardized participation metrics that, if used, would enable PA researchers to find common ground on which to transparently report participation rates. We also propose ways for researchers to compare the characteristics of those who participated and those who did not participate to assess whether their PA estimates might be subject to selection bias. Finally, we illustrate our participation metrics by considering their application to the National Health and Nutrition Examination Survey (NHANES) accelerometer studies conducted in 2011-2012 and 2013-2014.

Developing State Personal Income Distribution Statistics
Dirk van Duym, *Bureau of Economic Analysis*
Christian Awuku-Budu, *Bureau of Economic Analysis*

This paper provides new statistics on income inequality by state, retaining consistency with Bureau of Economic Analysis (BEA) data on both State Personal Income aggregates and the national personal income distribution. We distribute BEA State Personal Income to households to show how those with different income levels share in prosperity and growth, using CPS microdata as the principal source data. To address concerns about sample size at the state level, we use a three-year pooled sample of CPS households. We use a number of other data sources to improve the distribution for specific income components and the top of the income distribution, including data from the IRS Statistics of Income, the Survey of Consumer Finances, the Medical Expenditure Panel Survey, and the American Community Survey. Once allocators have been chosen for all BEA income components, in effect we have complete microdata for all CPS households, for all BEA income components. We can then generate bottom-up inequality statistics that are consistent with published and unpublished BEA aggregates of detailed personal income components, and permit examination of trends by state and over the 2009-2018 time period.

Session C-1
Recruiting and Surveying Victims Using Social Media and Online Platforms: Promising Practices and Lessons Learned

Evaluating Crime Survey Responses and Engagement among Juveniles and their Parents using Social Media and Online Platforms
Jenna Truman, *Bureau of Justice Statistics*
Grace Kena, *Bureau of Justice Statistics*
Chris Krebs, *RTI International*

Youth make up a key demographic of interest in the Bureau of Justice Statistics (BJS)' National Crime Victimization Survey (NCVS). As part of its efforts to modernize the NCVS, in 2020, BJS undertook testing to evaluate youth comprehension of the redesigned NCVS instrument as well as well the quality of data collected from youth, and to better understand from parents and youth reasons for nonresponse and strategies for increasing youth participation. This presentation will highlight recruiting methods and strategies for conducting virtual interviews.

Testing Improvements to the NCVS Hate Crime Items using Online Survey Panels
Grace Kena, *Bureau of Justice Statistics*
Lynn Langton, *RTI International*
Chris Krebs, *RTI International*

In the fall of 2020, the Bureau of Justice Statistics (BJS) and RTI International collected data from over 4,000 survey respondents and conducted cognitive interviews with over 30 to assess respondent understanding of key terms and refine measurement of core hate crime constructs in the National Crime Victimization Survey (NCVS). This presentation will describe aspects of the testing methodology, including initial challenges with data falsification, as well as data collection considerations for relatively rare events and hard to find populations.

Using Online Survey Panels to Enhance Identity Theft Measurement
Lynn Langton, *RTI International*
Chris Krebs, *RTI International*
Erika Harrell, *Bureau of Justice Statistics*
Grace Kena, *Bureau of Justice Statistics*

In 2020, the Bureau of Justice Statistics (BJS) and RTI International successfully used online testing as part of BJS's efforts to modernize and enhance measurement for the Identity Theft Supplement (ITS) of The National Crime Victimization Survey. The ITS provides person-level data on identity theft for household persons age 16 or older. BJS and RTI conducted cognitive testing, and a pilot test that used three sources of sample and included over 30,000 respondents that were diverse across demographic characteristics and identity theft crime types. This presentation will include discussion of the methodology and testing components, including limitations and lessons learned. Evaluating crime survey responses and engagement among juveniles and their parents using social media and online platforms. Youth make up a key demographic of interest in the Bureau of Justice Statistics (BJS)' National Crime Victimization Survey (NCVS). As part of its efforts to modernize the NCVS, in 2020, BJS undertook testing to evaluate youth comprehension of the redesigned NCVS instrument as well as well the quality of data collected from youth, and to better understand from parents and youth reasons for nonresponse and strategies for increasing youth participation. This presentation will highlight recruiting methods and strategies for conducting virtual interviews.

Concurrent Session C-2
Evidence and Data Policy

Automated Collection of Publicly Available Data from the Internet
Anup Mathur, *U.S. Census Bureau*
Michael Castro, *U.S. Census Bureau*
Sumit Khaneja, *U.S. Census Bureau*

The advent of new data collection methods such as automated web scraping and web crawling may provide opportunities for the Census Bureau to improve the statistical products we produce for the American public, reduce the burden we place on our business, government, and individual respondents, and lower the cost of data collection. The Census Bureau is exploring ways to leverage these rapidly evolving technologies in a manner that is consistent with our commitment to scientific integrity, ethical and legal responsibilities, and our Privacy Principles. To this end, the Census Bureau has established a policy for the automated collection of data from the internet and associated governance in the form of the Automated Internet Data Collection Review Board (AIDCRB) . In this article we will delve into the need and origins of the policy and describe our implementation of the policy and the associated governance.

Creating Policy Tools for Broadband Subsidy programs: Combining spatial regression discontinuity designs and Bayesian Wombling.
Aritra Halder, *University of Virginia*
Joshua R. Goldstein, *University of Virginia*
John Pender, *Economic Research Service*

The USDA Rural Utilities Service (RUS) administers grant and loan programs to improve infrastructure access and quality of life in rural areas. We are collaborating with the USDA Economic Research Service to develop a statistical framework to study the impacts of RUS broadband programs on residential property values. Our approach features a Bayesian hierarchical spatial modeling of property prices at its core. The main programmatic focus will be on impacts of the ReConnect Program (RCP) – the most recently established and largest USDA rural broadband program at present – and selected other USDA programs. These methods can be used by program managers and stakeholders to measure the effects of broadband programs in for policy analysis and evaluation.

Evidence Act Standard Application Process: Stakeholder Engagement and Observations
Jon Desenberg, *The MITRE Corporation*
Heather Madray, *U.S. Census Bureau*
Sue Collin, *The MITRE Corporation*

The 2018 Foundations for Evidence-Based Policymaking Act brought Congressional attention to both the unlocked value and current challenges in using Federal data. The law requires a standard application process for all restricted-use data requests that fall under the Confidential Information Protection and Statistical Efficiency Act. To ensure the process meets the needs of the data user community, the legislation also requires engagement with external stakeholder groups during requirements gathering, development, testing, and eventual roll-out. The U.S. Census Bureau and the MITRE Corporation engaged with data users and stakeholders to learn about their experiences with current application processes. From December 2020 to January 2021, data users and stakeholders representing a broad array of viewpoints were interviewed. Data users in the research and evaluation community as well as statistical policy experts and senior Congressional staff were all interviewed for a unique understanding of today's data access challenges and the future vision for a streamlined application process. The proposed poster will capture data users and stakeholders perspectives on transparency and access as well as issues of privacy and security during this important step in implementing the Evidence Act.

Six Crises or One Dozen Opportunities in Public-Stewardship Statistics
John Eltinge, *U.S. Census Bureau*

In recent years, the methodological literature, advisory groups and statistical agencies have identified a number of perceived crises in the processes for production, dissemination and use of public-stewardship statistical information. Examples include degradation of some dimensions of data quality; the reproducibility crisis and other inferential issues; risks to privacy and confidentiality; reductions in availability of discretionary resources; changing expectations about public goods; and effects of the general decline in trust in science, expertise and public institutions.
This presentation explores these perceived crises through application and extension of some customary statistical models for measures of quality, risk and cost. These models point to some constructive responses that have the potential to produce substantial improvements in public-stewardship statistical work. One dozen opportunities receive special attention, including six opportunities for improved understanding of our operational procedures; and six opportunities for improved design of those procedures.

Concurrent Session C-3
Sampling Issues, Estimation, and Testing

Business Sample Revision Universe Extraction Measure of Size Determination and Validation
Kenishea Donaldson, *U.S. Census Bureau*
Katrina Washington, *U.S. Census Bureau*
Erica Wong, *U.S. Census Bureau*

Universe extraction is the initial step in the process of reselecting the samples for the monthly, quarterly, and annual Retail, Wholesale and Services Surveys. These samples are redrawn approximately every 5 years following completion of the Economic Census. All multiunit and singleunit establishments that are in-scope to the surveys are extracted from the Census Bureau's Business Register (BR). Administrative and Census data are used to compute several estimates of each establishment's measure of size (MOS, i.e., sales or receipts) as well as an estimate of each establishment's inventory (for Retail and Wholesale). For the 2017 Business Sample Revision, up to 6 MOS for sales were computed for each establishment using receipts and payroll data from the Census and BR. The best sales MOS was then determined by testing the ratios of the data items constituting each of the competing MOS against a hierarchy of ratio edits, as well as performing additional editing and correction methodologies on the collection of calculated MOS. This paper outlines the research conducted to improve the MOS determination and validation process in preparation for the next sample revision.

Influential Unit Treatment in the Annual Survey of Local Government Finances
Noah Bassel, *U.S. Census Bureau*

The Annual Survey of Local Government Finances is a multipurpose survey selected every five years that publishes annual estimated totals of local government expenditures, revenues, debts, and assets for individual states and at the national level. As in many repeated, multipurpose economic surveys the design weights of this survey will not necessarily be strongly predictive of all variables of interest in all survey years. It is therefore possible for small units with large survey weights to report unusual values for key variables making classical design-based estimators unstable. In this study we investigate several methods for treating such influential units, including Winsorized estimators, model assisted, and model-based approaches. We also investigate the question of preserving consistency across domains in production.

Probabilistic Classification of a Policy Relevant Subpopulation: The Case of High-Tech Start-ups Operating in Innovation Markets
Timothy R. Wojan, *National Center for Science and Engineering Statistics*
John Jankowski, *National Center for Science and Engineering Statistics*
Audrey Kindlon, *National Center for Science and Engineering Statistics*

NCSES and the Census Bureau have been collecting data on R&D performing microbusinesses since 2017. Despite accounting for only 1.5% of R&D expenditures, research points to a "division of innovative labor" where high-risk radical innovation is pursued predominantly by small start-ups and large incumbents specialize in incremental innovation. Differentiating R&D performing microbusinesses pursuing radical innovation for eventual licensing or acquisition by incumbents from small R&D performers that are either providing R&D services or are in the early stages of entering product markets is the objective of this research. Latent class analysis (LCA) will be used with the Annual Business Survey to probabilistically assign membership to at least three classes: 1) innovation market participants, 2) early-stage product market participants, and 3) R&D service providers. Categorical variables in the LCA include whether revenues come predominantly from product sales or from grants and licensing; whether founders have advanced STEM degrees; the value placed on intellectual property protection; and whether firms used or tested esoteric technologies. The analysis will provide the first estimate of the size of the microbusiness innovation market in the US which has only been studied in specific sectors to date.

Roots from Trees: A Machine Learning Approach to Unit Root Detection
Gary Cornwall, *Bureau of Economic Analysis*
Jeff Chen, *University of Cambridge*
Beau Sauley, *Murray State University*

In this paper we have updated the hypothesis testing framework by drawing upon modern computational power and classification models from machine learning. We show that a simple classification algorithm such as a boosted decision stump can be used to fully recover the full size-power trade-off for any single test statistic. This recovery implies an equivalence, under certain conditions, between the basic building block of modern machine learning and hypothesis testing. Second, we show that more complex algorithms such as the random forest and gradient boosted machine can serve as mapping functions in place of the traditional null distribution. This allows for multiple test statistics and other information to be evaluated simultaneously and thus form a pseudo-composite hypothesis test. Moreover, we show how practitioners can make explicit the relative costs of Type I and Type II errors to contextualize the test into a specific decision framework. To illustrate this approach we revisit the case of testing for unit roots, a difficult problem in time series econometrics for which existing tests are known to exhibit low power. Using a simulation framework common to the literature we show that this approach can improve upon overall accuracy of the traditional unit root test(s) by seventeen percentage points, and the sensitivity by thirty six percentage points.

Understanding the Characteristics of Unresolved Matched Records in Capture-Recapture Methodology
Denise A. Abreu, *National Agricultural Statistics Service*

The National Agricultural Statistics Service (NASS) conducts a Census of Agriculture (COA) every 5 years, in years ending in 2 and 7. The census uses a list frame. The 2017 COA used capture-recapture to adjust the COA for undercoverage, nonresponse, and misclassification of farms/non-farms. NASS's June Area Survey (JAS) was used as the independent survey in the capture-recapture approach. The JAS is conducted annually in June. It is based on an area frame and the data are collected via in-person interviews. Capture-recapture requires a matched dataset consisting of all matches of a COA record to a JAS record. This dataset is the foundation for modeling the probability that a JAS farm is captured by the COA. A farm is a place with $1000 or more of sales or potential sales. In the dataset, the farm status based on the JAS and the COA agree in most cases. However, in other cases, a record is identified as a farm (non-farm) on the JAS and a non-farm (farm) on the Census. These records have unresolved farm status. Resolving the farm status is important to the accuracy of the COA published estimates. The characteristics of the records with unresolved farm status are described.

Session C-4
The Standard Application Process: A Coordinated Effort Across the Federal Statistical System to Implement an Evidence Act Requirement

SAP Pilot: A One-Stop Application Portal
Heather Madray, *U.S. Census Bureau*

This presentation discusses the Pilot effort at creating a one-stop application portal and the resulting lessons learned.

SAP Policy Guidance Development
Mark Prell, *Economic Research Service*

This presentation will discuss the policy guidance that established the SAP and enabled the development of technical requirements to build an SAP portal.

SAP Stakeholder Engagement and Outreach Efforts
Vipin Arora, *National Center for Science and Engineering Statistics*

This presentation will discuss ongoing stakeholder engagement efforts and how this information has and will help shape the SAP development

The Fully Functional SAP Portal
John Finamore, *National Center for Science and Engineering Statistics*

This presentation discusses the development of the SAP portal, highlights the role that the Pilot, stakeholder engagement, and policy guidance played in its creation, and provides a demonstration of the SAP portal beta release.

Session C-5
Leveraging Administrative Data for Survey Methods and Research

Comparing the 2019 American Housing Survey to Contemporary Sources of Property Tax Records: Implications for Survey Efficiency and Quality

Ariel J. Binder, *US Census Bureau*
Emily Molfino, *Department of Housing and Urban Development*
John Voorheis, *US Census Bureau*

This memo summarizes the Census Bureau's research efforts, in collaboration with the Department of Housing and Urban Development, to understand the potential uses of administrative property tax data in enhancing the American Community Survey. Such data could contribute to the construction of the sampling frame, sample universes and skip patterns, response editing and allocation, potential question removal, or the independent production of small-area statistics. As a starting point, it is important to understand the extent to which the property tax records can be linked to, and contain similar information as, existing AHS records. Accordingly, this memo contains six sets of results. First, it reports linkage rates between two contemporary property tax data sources, and how these rates vary across states. Second, it reports the fractions of 2019 AHS housing units that could be linked (via address information) to each property tax data source. Third, it reports agreement rates between the AHS record and the property tax record for each of 11 AHS variables. Fourth, it investigates heterogeneity in linkage and agreement rates across variables, states, and property tax data sources. Fifth, it shows how the inclusion of "fuzzier" linkages based on geographic (i.e. lat/long) information affects agreement rates across states and data sources. Sixth, it reports national-level variation in variable agreement rates by metro status and by structure type.

Determining Household Obesity Status Using Scanner Data

Elina T. Page, *Economic Research Service*
Sabrina K. Young, *Economic Research Service*
Megan Sweitzer, *National Center for Science and Engineering Statistics*
Abigail M. Okrent, *National Center for Science and Engineering Statistics*

Diet is the primary contributing factor to the ongoing obesity epidemic, with devastating negative health, economic, and social impacts for individuals and society. The IRI Consumer Network and the IRI MedProfiler surveys are unique datasets that allow researchers to link household food purchases to self-reported height and weight of household members to better study and understand the relationship between diet and obesity. However, self-reported height and weight are often misreported in survey data, and therefore it is necessary to consider the quality of these data when calculating body mass index (BMI) and classifying household members by body weight status– i.e., as normal weight, overweight, and obese. In this study, we use data from 2011 to 2018 to compare the two datasets and assess methods for correcting for self-reported measurement error in the IRI MedProfiler data using self-reported and measured data from the National Health and Nutrition Examination Survey (NHANES). We then assess methods for classifying household body weight status and compare household food expenditures on fruit and vegetables among normal weight, overweight, and obese household.

Measuring the Distributional Effects of Climate Change and Environmental Injustice with Linked Survey, Census and Administrative Data
John Voorheis, *U.S. Census Bureau*

There has been increasing interest in understanding the social, economic and health impacts of climate change and other environmental hazards, and in how these impacts are distributed across race, income, and other demographic groups. Research to better understand these impacts has largely not taken advantage of rich microdata available at the US Census Bureau, however. This project describes Census Bureau efforts to fill this gap by facilitating research on the impact of the environment on households and firms by combining confidential microdata and cutting-edge breakthroughs in environmental measurement. We present three case studies of how this data can be used to better understand these issues: 1) using satellite derived data on ambient air pollution and linked survey and administrative records to measure how the gap in pollution exposure varies by race and ethnicity, 2) using meteorological modelling, administrative tax data and survey data from the American Community Survey to measure the distributional impacts of hurricanes, and 3) using high resolution information on smoke plumes and the extent of fires combined with detailed demographic data to understand the distributional impacts of increasingly severe forest fires.

An Evaluation of the Gender Wage Gap using Linked Census and Administrative Records
Brad Foster, *U.S. Census Bureau*
Marta Murray-Close, *U.S. Census Bureau*
Christin Landivar, *U.S. Department of Labor*
Mark deWolf, *U.S. Department of Labor*

The narrowing of the gender wage gap has slowed in recent decades, even as women's human capital characteristics increasingly resemble men's. Recent scholarship suggests that the key to understanding the remaining gender gap in wages is the measurement of raw and residual wage gaps across and within detailed occupation categories, but this detailed analysis is not possible using publicly available data sources. Our research links American Community Survey and Current Population Survey – Annual Social and Economic Supplement responses to Internal Revenue Service W-2 and Social Security Administration earnings records, respectively, to (1) measure the contemporary gender wage gap in 316 detailed occupation categories, (2) decompose these gaps to attain occupation-specific residuals, and (3) model residual gender wage gaps as a function of O*NET occupation characteristics. Using an hourly wage measure derived from recent administrative earnings and survey-reported hours and weeks worked, we find a contemporary wage gap of 18 percent among full-time, year-round workers. We demonstrate that this gap varies significantly across occupations: gaps are as large as 45% in some occupations but have closed completely in others. Furthermore, we show that wage gaps tend to be larger in more competitive and hazardous occupations, but smaller in occupations granting employees more autonomy and requiring more communication among workers and with clients. Taken together, results communicate the quality of administrative data on earnings, particularly when linked with survey responses, and demonstrate how such data can facilitate a better understanding of difficult-to-measure economic concepts.

Identifying Undercounted Children Using Birth Records
Gloria Aldana, *U.S. Census Bureau*

The Census Bureau acknowledges the well-documented undercount of children in Census Bureau surveys, including the Decennial Census and the American Community Survey (ACS). To reduce the undercount of children, the Census Bureau has focused research on evaluating the coverage of children in surveys and understanding the causes of undercounting. This presentation will cover progress so far on a case study using California birth records to examine differences in counted versus non-counted young children in the Decennial Census. Given that birth certificate records should provide a complete record of all births, birth record data will be used to examine trends in coverage rates and misreported child ages, based on demographic and socioeconomic variables. The project will later use IRS records to examine coverage rates and misreported child ages in the Decennial Census and the ACS.

Session D-1
Data Collection Challenges During COVID-19

Implications of the Coronavirus Pandemic on the National Public Education Financial Survey (NPEFS) and the School District Finance Survey (F-33)
Stephen Q. Cornman, *U.S. Department of Education*
Malia Howell, *U.S. Census Bureau*
Osei Ampadu, *U.S. Census Bureau*

The National Public Education Financial Survey (NPEFS) and School District Finance Survey (F-33) collect finance data on elementary and secondary education in the United States. As a direct result of the COVID–19 circumstances, many districts closed school buildings and began remote learning instruction, disrupting the collection of attendance data. The Coronavirus Aid, Relief, and Economic Security (CARES) Act provided $30.75 billion to public PK-12 and higher education school systems. The CARES Act also established and appropriated $150 billion to the Coronavirus Relief Fund. NCES and the Census Bureau received approval from OMB to add CARES Act revenue and expenditure data items to the surveys and modify instructions. Recommendations were considered from an expert panel of State Fiscal Coordinators and school district personnel; 51 State Fiscal Coordinators; and various offices within the U.S. Department of Education. The question of whether there is a match between the new CARES Act variables and data that State Fiscal Coordinators can report on the surveys will be reviewed in terms of burden, data quality, and resources being expended by states and the federal government.

Measuring the Impacts of the Coronavirus Pandemic on Higher Education R&D Expenditures and Research Space: Experience from Survey Question Development at the National Center for Science and Engineering Statistics
Michael T. Gibbons, *National Center for Science and Engineering Statistics*

The National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation reviewed two of its establishment surveys, the Higher Education Research and Development (HERD) survey and the Survey of Science and Engineering Research Facilities (Facilities), to determine whether metrics on the impact to R&D from the Coronavirus pandemic could be effectively added to the FY 2020 HERD and FY 2021 Facilities questionnaires. NCSES considered respondent burden versus the benefits of understanding the impacts of the pandemic on R&D data trends. NCSES interviewed respondents from each survey to gain an initial understanding of pandemic impacts and relevant data tracking by the institutions. Subsequent HERD survey question development culminated in the addition of 3 qualitative questions on the FY 2020 HERD survey. NCSES decided not to add questions to the FY 2021 Facilities survey after the initial interviews. This presentation will highlight the broader questions NCSES attempted to address, the question development process (including interviews, a literature review, a respondent webinar, question finalization, and OMB clearance) and the preliminary results from each data collection.

Patterns of Response During Covid-19 in a National Survey of Businesses: A Look at the Medical Expenditure Panel Survey - Insurance Component (MEPS-IC)
David Kashihara, *Agency for Healthcare Research and Quality*

In the Spring of 2020, Covid-19 had spread enough in the United States to force many businesses to alter their operations. At that time, the Medical Expenditure Panel Survey – Insurance Component (MEPS-IC) was in final preparations. Due to the immense infrastructure necessary to field a large Federal survey, the process began roughly on schedule with some major changes to data collection procedures. The MEPS-IC is an annual survey that produces national and state-level estimates on topics including the percentage of employers offering health insurance and the premiums and deductibles of the plans. Prior to the pandemic, participants were able to respond via the web, mail, or telephone (follow-up). However, with the onset of the pandemic, challenges were presented as survey operations shut down, businesses closed, and teleworking became more prominent. This presentation describes survey response patterns in each data collection operation of the 2020 MEPS-IC. The patterns will be analyzed by comparing to prior years using select business characteristics such as industry and firm size. Changes that may have affected response patterns will also be discussed. This presentation will expand the field of knowledge about how businesses respond to surveys, especially during an impactful pandemic. These results can be a valuable aid to other business surveys as they plan their data collection methods.

The 2020 COVID-19 Module on the Survey of Graduate Students and Postdocs in Science and Engineering
Caren Arbeit, *RTI International*
Pat Green, *RTI International*
Mike Yamaner, *National Center for Science and Engineering Statistics*

Every year, over 700 institutions provide information on science, engineering and health graduate student enrollment, postdocs and nonfaculty researchers to the Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS), sponsored by the National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation (NSF) and by the National Institutes of Health (NIH). We saw many coordinators struggle to complete the 2019 data collection in March of 2020 due to the COVID-19 pandemic. Articles about the impact of the covid-19 pandemic on graduate enrollment, higher education budgets, and visa-holding students implied that we should expect to see large changes in enrollment and potentially the capacity of our respondents to complete the survey. Working with NCSES leadership, we fielded a 26 question (including 5 open-ended text responses) module on the impact of the COVID-19 Pandemic hosted within the main GSS web application. The module included items on: ability to complete the GSS survey; the impact on graduate student enrollment (master's and doctoral) and funding in Fall 2020, as well as the impact on postdocs and NFRs, including asking about hiring freezes. Several questions highlighted asked about temporary visa holding students and postdocs. The module was fielded after the launch of the GSS, with a relatively short turnaround in order to gather data that could be released prior to the full 2020 data release. Even with the short turnaround, we had excellent response rates. In our presentation, we will present our key findings and discuss our plans for a Fall 2021 follow-up

Concurrent Session D-2
Innovations in Health Insurance Data Collection and Measurement Across Federal Surveys

Two times a charm? Verifying Reports of Uninsurance in a National Survey
Paul Jacobs, *Agency for Healthcare Research and Quality*
Patricia Keenan, *Agency for Healthcare Research and Quality*

To improve health insurance coverage estimates, the Medical Expenditure Panel Survey Household Component (MEPS-HC) added a verification series. The series confirms whether individuals who did not initially report coverage actually had health insurance coverage at some time during the survey round. The MEPS verification series was based on the Current Population Survey (CPS) verification questions and adapted to the MEPS context. In this analysis, we use 2017 and 2018 MEPS-HC data to examine the impact of the verification questions on coverage. Using the same respondents, we compare insurance coverage rates before and after taking into account responses to the verification series. We examine the percent reporting coverage through verification, the difference in coverage rates, overall and by coverage types (public versus private, employer, nongroup, Medicaid/CHIP, and Medicare). In particular, we explore whether the verification series helped address the Medicaid undercount, or whether it resulted in increases in other categories, such as private coverage rates (as was found for the CPS verification question). We also examine whether responses varied by sub groups, such as educational attainment, income, age, and family size or structure.

Improving Measurement of VA Health Coverage among Military Veterans on the National Health Interview Survey
Robin A. Cohen, *National Center for Health Statistics*
Carla E. Zelaya, *National Center for Health Statistics*

In 2017, the Department of Veterans Affairs (VA) estimated that approximately 8.8 million veterans enrolled in, and 6.0 million utilized the VA health care system. However, in the National Health Interview Survey (NHIS), estimates of VA health care coverage through self-report fell short (3.0 million) of these administrative statistics.
Since 1993, VA health care coverage has been included as a response option in the NHIS health insurance section. However, veterans who do not consider the VA a primary source of care, only use VA health care occasionally or for particular health services (e.g., mental health care), or a veteran who has enrolled but never utilized VA health care, may fail to report their VA coverage in this section.
Therefore, in 2018, we added a new question to the veteran section (positioned after the health insurance section) of the NHIS: "[Have you/has {person}] ever used or enrolled in VA health care?" Our strategy to add a targeted probe mirrored a similar approach taken to address undercounts of Medicare and Medicaid coverage in 2004. Results from 2018 and 2019 show an increase in the reporting of VA health care coverage to 8.8 million, the same as the 2018 VA estimate. Respondents who benefited from the extra probe to elicit the desired information were more likely to have private or public health coverage, be under age 65, employed, or be in fair or poor health, and they were less likely to be divorced or separated or to be the family-respondent. Our results suggest that having more than one question on the same concept may minimize measurement error of complex concepts (e.g., linked to social and other identities).

Decomposing Data Processing Improvements on Estimates Health Insurance Coverage in the Current Population Survey Annual Social and Economic Supplement (CPS ASEC)

Laryssa Mykyta, *U.S. Census Bureau*
Amy Steinweg, *U.S. Census Bureau*
Katherine Keisler-Starkey, *U.S. Census Bureau*

After a decade of research suggesting that the CPS ASEC captured less health insurance coverage than other federal surveys, the Census Bureau implemented a two-stage redesign. A new questionnaire was introduced in 2014, and a new "processing system" for extracting and imputing data was introduced in 2019. Rising nonresponse makes it critical to evaluate how post-collection survey procedures—not just the questionnaire—contribute to accurate estimates and improved data quality. In previous analyses, we used data from the 2017 CPS ASEC Production and Research Files and the 2018 CPS ASEC Production and Bridge Files, to assess the effect of processing system changes on CPS ASEC health insurance estimates. While results revealed that the updated data processing system improved estimates and addressed previously noted limitations of the CPS ASEC, little is known about how the components of the new system contributed to these improvements.

In this paper, we explore the relative contribution of elements of the updated processing system for health insurance estimates. Specifically, we examine the effects to changes in income imputation and HIU assignment to estimates of coverage. Findings will highlight the contribution of post-collection data processing for improving data quality in surveys measuring health insurance coverage.

Using Insurance Claims Data in the Medical Price Indexes

Brian Parker, *Bureau of Labor Statistics*
John Bieler, *Bureau of Labor Statistics*
Caleb Cho, *Bureau of Labor Statistics*
Brett Matsumoto, *Bureau of Labor Statistics*

The use of medical claims data in the construction of the medical price indexes presents many opportunities and challenges for the Bureau of Labor Statistics. This project seeks to develop a feasible methodology for supplementing manual price collection in the medical CPI using insurance claims data. As part of a feasibility study, we constructed price indexes using nationwide claims data and compared them to the CPI medical indexes. Some of the practical issues we consider are the effect of the time lag involved in using claims data, weighting issues related to the use of both claims data and traditional manually collected data, and the high variability of prices in the claims data at the level of specific provider and service. The results of our preliminary analysis show promise for the use of claims data.

Session D-3
Enhancing Transparency, Reproducibility and Privacy

A Federal Tiered Access Model to Promote Evidence-Based Policymaking
Amy O'Hara, *Georgetown University*
Lahy Amman, *Georgetown University*

In federal statistical agencies, data access for internal and external users is often limited to public-use data files or, after surpassing legal and regulatory barriers, negotiated access to restricted data. However, public use files and this binary access of control decisions for data are not sufficient for performing the necessary research for evidence-based policy. This paper explores existing tiered access models in federal statistical agencies and proposes extensions to facilitate greater data use.
We describe tiered access at the Internal Revenue Service, outlining the history of their access tiers and current projects underway. We then propose extensions to this model, addressing the appropriate physical, technical, and human resource controls required for preserving the confidentiality of government data of varying sensitivity levels. We describe the model in terms of current standards and regulations (FISMA, FIPS, OMB guidance), federal privacy and confidentiality laws, and stakeholder needs, pointing out where privacy preserving technologies, such as secure multiparty computation and synthetic datasets, can improve data access and linkage of datasets.

Sharing Student Data Across Organizational Boundaries Using Secure Multiparty Computation
Stephanie Straus, *Georgetown University*
David Archer, *Galois, Inc.*
Amy O'Hara, *Georgetown University*
Rawane Issa, *Galois, Inc.*

The Federal Data Strategy and the Evidence Act (2018) mandate inter-agency data sharing to promote informed decision making. However, distrust between parties can hamper these efforts. Our research solves this problem through a demonstration use case of cryptographic privacy-preserving technology, secure multiparty computation (MPC), with the National Center for Education Statistics (NCES) at the Department of Education.
Our intra-agency prototype reproduces a portion of the annual 2015-16 National Postsecondary Student Aid Study (NPSAS) report, recreating statistics on average federal Title IV aid received by undergraduates for the 2015-16 academic year. The statistics come from the linkage of two different data sources, NPSAS and the National Student Loan Data System (NSLDS). We simulate the record linkage and joint analysis of data from these two parties, using virtual machines held in distinct "trust zones" to represent NPSAS and NSLDS host computers. These machines cooperate to carry out a strand of MPC called Private Set Intersection with associated computation (Pinkas et al, 2019). We explain the accuracy of our results, resource utilization, and degree of cryptographic privacy assurance, ultimately demonstrating that MPC technology can assure confidentiality of sensitive information while enabling practical, performant analysis of combined data held by diverse organizations.

The Privacy Paradox: How Well do Respondent Attitudes and Concerns about Privacy Predict Privacy-related Behaviors?

Casey Eggleston, *U.S. Census Bureau*
Aleia Clark Fobia, *U.S. Census Bureau*
Jennifer Hunter Childs, *U.S. Census Bureau*

The privacy paradox refers to research that demonstrates discrepancies between privacy attitudes and actual behaviors (Norberg et al. 2007, Barth et al. 2019). Often this research has found that attitudinal measures suggest respondents value privacy, but their behaviors do not seem to support that valuation. Using self-reported attitudes and behaviors from a national survey of respondent privacy concerns, we investigate the relationship between respondent attitudes and concerns about privacy and their privacy-related behaviors. The survey data include measures of privacy concerns, including concern about hacking and reidentification and concern about the confidentiality of individual data items. Self-reported privacy-related behaviors include actions that might be avoided to protect privacy (e.g., posting reviews online, reporting income in a survey) and others that might be taken to protect privacy (such as signing up for the national Do Not Call registry). We explore how well privacy-related behaviors are predicted by respondent privacy concerns and whether there are predictable patterns of respondent behavior related to privacy attitudes that could be used to compose a meaningful privacy-seeking behavioral scale. Much of the literature on the privacy paradox addresses the context of social media and online behavior. While our survey also addresses these behaviors, the context of government data collection and use provides a different focal point by which to understand the potential privacy paradox.

Session D-4
A Call for Federal Statistical Coordination on Climate Change Data Needs

In the decentralized federal statistical system of the United States, thirteen principal statistical agencies function within cabinet level departments under four principles: relevance to policy, credibility among data users, trust among data providers, and independence from political and other external influences. Their work is largely focused on policy and mission areas of their host departments. Cross-cutting issues such as climate change test the readiness of a decentralized system to meet a fundamental statistical operation: to coordinate and collaborate effectively and quickly across agencies. As evidenced by the COVID response, no one agency has the authority or capacity to tackle a broad policy data need alone. Crisis data response requires extraordinary inter-governmental effort. In this session, panelists expert in executive orders, departmental regulations, and federal legislation on climate issues will call out data needs and data strategies that may advance climate awareness, and importantly, public abilities to inform and address vulnerabilities and adaptation. To further resilience in communities, institutions, and the economy, the session will close with considerations for how a modernized, virtually-integrated federal statistical system could meet these challenges.

Panelists:

- Mindy Selman – Senior Analyst, Office of Energy and Environmental Policy, OCE, U.S. Department of Agriculture
- Carla Frisch – Principal Deputy Director, Office of Policy, U.S. Department of Energy
- Richard Allen – Chief Data Officer and Strategist, U.S. Environmental Protection Agency
- Nick Hart – President, Data Foundation

Session D-5
Data Linkage for Government Health and Safety Statistics: Methods and Applications

Assessing Linkage Eligibility Bias in the National Health Interview Survey
Jonathan Aram, *National Center for Health Statistics*
Crescent B. Martin, *National Center for Health Statistics*
Lisa B. Mirel, *National Center for Health Statistics*

Linking survey and administrative data can facilitate richer analyses by augmenting survey data with mortality or other administrative data. However, estimates derived from linked data can include bias when some participants are ineligible for linkage. The Data Linkage Program at the National Center for Health Statistics (NCHS) has quantified linkage eligibility bias and explored ways to reduce this bias in estimates. In this example, we focus on the recent linkage of the National Health Interview Survey (NHIS) data and Centers for Medicare & Medicaid Services (CMS) Medicare records through 2018. This talk will build on previous work that examined changes in linkage consent procedures and their correlates with bias. We will highlight the new linked data resource and examine the effect of linkage eligibility on estimates by assessing bias (e.g., comparing estimates from the full sample to the linkage eligible sample). We will also describe how adjusting sample weights for linkage consent may reduce bias, focusing on key health indicators collected in the survey. We will conclude by discussing implications for analyses and future directions of the NCHS Data Linkage Program.

Using Administrative Data to Supplement and Assess Occupational Health and Safety Statistics
Ellen Galantucci, *Bureau of Labor Statistics*

The Survey of Occupational Injuries and Illnesses (SOII) at the Bureau of Labor Statistics (BLS) collects information on nonfatal workplace injuries and illnesses among private establishments throughout the United States. In 2017, the Occupational Safety and Health Administration (OSHA) began requiring certain employers to submit some of the same information that SOII collects annually through their Injury Tracking Application (ITA). The Office of Management and Budget (OMB) tasked the BLS and OSHA with working together to reduce the burden on employers that needed to report to both programs, which are jointly part of the U.S. Department of Labor. In this paper, I outline the various methods we have attempted to use for linking the data, which include probabilistic linkage and collaboration between the agencies to collect additional variables for linking. I also discuss the ways in which the BLS is currently using the data collected through the ITA which OSHA has been providing to BLS. I conclude with other ways in which the BLS is planning to use the OSHA ITA data in the future.

Social Determinants of Emergency Department Utilization in Utah
David Powers, *U.S. Census Bureau*
Sara Robinson, *U.S. Census Bureau*
Edward Berchick, *U.S. Census Bureau*
J. Alex Branham, *U.S. Census Bureau*
Lucinda Dalzell, *U.S. Census Bureau*
Lorelle Dennis, *U.S. Census Bureau*
Kristi Eckerson, *U.S. Census Bureau*
Alfred Gottschalck, *U.S. Census Bureau*
Joanna Motro, *U.S. Census Bureau*
John Posey, *U.S. Census Bureau*
Andrew Verdon, *U.S. Census Bureau*
Victoria Udalova, *U.S. Census Bureau*

In 2019, Census entered into a partnership agreement with the Utah Department of Health (UDOH). One aim of this partnership is to improve our understanding of the social determinants of emergency department (ED) utilization in Utah. This project emerged from the Enhancing Health Data (EHealth) program at Census which strategically re-uses administrative records to improve the quality and availability of statistical information, which can advance population health. In this project, we demonstrate how existing survey data can be enhanced with administrative records data, and vice versa, to support insights into the social determinants of health (SDOH). We link 2013-2017 UDOH ED encounter-level data with 2013-2017 1-Year American Community Survey (ACS) restricted microdata to study the relationship between SDOH and the likelihood of preventable ED visits. We find that most of the SDOH characteristics we examine are significantly associated with the rate of preventable ED visits. Our findings may inform efforts to reduce costs related to ED visits for non-emergent issues and improve our understanding of the role the broader context of people's lives plays in health outcomes.

Examining Earnings of U.S. Physicians Using Tax Return Information
Victoria Udalova, *U.S. Census Bureau*

This project emerged from the Enhancing Health Data (EHealth) program at Census which strategically re-uses administrative records to improve the quality and availability of statistical information, which can advance population health. In this project, we link the administrative registry of U.S. physicians with the universe of individual federal income tax returns data, Social Security Administration records, and American Community Survey (ACS) responses to provide new estimates on earnings of the U.S. physicians. Existing evidence on physician earnings has relied on survey data and faced measurement challenges, such as top-coding and income underreporting. These linked survey and administrative data overcome many of these issues and allow us to gain a better understanding of earnings in this occupation. We compare tax-based estimates to those from the survey data and describe general patterns in physician earnings by detailed specialty, geography, age, and time.

Session E-1
Strategies for Recruiting and Identifying Subpopulations

Administrative Data Limitations and the Need for Continued Improvement of Face to Face Interviewing for Non-English Speakers in National Surveys
Patricia Goerman, *U.S. Census Bureau*
Alisu Schoua Glusberg, *Research Support Services*
Leticia Fernandez, *U.S. Census Bureau*

In recent years there has been a push towards linking individual-level records across multiple data sources to supplement national surveys; thus, seeking to compensate for decreasing survey response rates and item nonresponse. Hard-to-count populations, such as non-English speakers, often are not well represented in existing data sources and it can be more difficult to create linked data and ensure data quality when using alternative data sources for them. Non-English speakers are often harder to contact in every survey mode, both due to challenges in survey operations and due to a lack of respondent motivation or fears about participating. We will present recent evidence about incomplete inclusion of non-English speakers in administrative and linked data and we will discuss findings from a 2020 Census evaluation where we debriefed bilingual census interviewers who conducted face to face interviews in one of 7 different languages at the doorstep. The bilingual interviewers shared thoughts about why some non-English respondents were reluctant to participate in the 2020 Census and techniques they used at the doorstep to encourage participation. We will end with recommendations about improving coverage of under-represented groups. It is of critical importance that we continue to improve inclusion in multiple modes of data collection so that all groups are represented in national data.

Hard to Get: Understanding Why People Take Our Surveys
Mina Muller, *Ipsos Public Affairs*
Seth Messinger, *Ipsos Public Affairs*
Randall K. Thomas, *Ipsos Public Affairs*

Of increasing concern to many survey researchers is the recruitment and retention of respondents of color. While the existing literature attempts to understand overall respondent rates via reference to choices made by potential respondents, few studies exist that study how various survey features affect participation of these groups specifically. We conducted a study with over 2,000 probability-based sample with large oversamples of Black and Latinx participants to explore the motivational factors of survey participation. We asked questions on prior survey participation and experiences, and focused on factors affecting their survey participation, including what topics were appealing, what types of information are helpful to know before participating, and the types of information concerning a survey would increase or decrease their participation. While some factors like incentives were commonly indicated as important for participation for all participants, other factors like topic were differentially motivating for people of color. We found that incentives and other motivational protocols typically useful for general population studies may not be as useful to attract and retain participants of specific groups. We summarize some of the implications for attracting more diverse participants.

Making Data Collection More Efficient? Optimizing Incentives and Reminder Modes
Jerry Timbrook, *RTI International*
Antje Kirchner, *RTI International*
Emilia Peytcheva, *RTI International*

Leverage saliency theory (Groves et al., 2000) posits that different respondents are persuaded to participate in a survey based on different design attributes. Two such attributes that researchers commonly vary to increase response rates and decrease the potential for nonresponse bias are: 1) the timing and amounts of incentives, and 2) the mode of contacting sample members to notify and remind them about a survey. For

example, offering time-limited incentives to early responders (i.e., "early bird" incentives) can lead to faster responses and increased participation rates (e.g., LeClere et al. 2012), though the effects of early bird incentives on sample representativeness are currently unknown. Additionally, providing advance notification of a survey via text messaging to mobile phones may increase response rates (Callegaro et al. 2011). Furthermore, few studies have explored the effect of text message reminders compared to telephone reminders on response rates or sample representativeness using validation data. In this study, we explore the effect of early bird incentives and text message reminders using data from the 2020/2022 Beginning Postsecondary Students Study (BPS:20/22) Field Test (n=3,703). BPS:20/22 is a mixed-mode web and telephone follow-up study of sample members from the 2019-2020 National Postsecondary Student Aid Study (NPSAS:20). To test the early bird incentive, we randomly assigned half of our sample to receive a $5 early bird incentive offer if they completed the survey within the first three weeks of data collection. The other half did not receive the early bird offer (i.e., the control). To compare text message versus telephone reminders, after about seven weeks of data collection, we randomly assigned half of our nonrespondents to receive only text message reminders, and the other half to only receive telephone reminders (i.e., the control) over a period of three weeks. For both experiments, we compare response rates, sample representativeness (i.e., demographic characteristics), and the timeliness of responses between the control and treatment groups. Preliminary results from the early bird experiment indicate that sample members who received the early bird incentive offer were more likely to respond (RR=49.7%) than those who did not receive the offer (RR=45.9%; diff=3.8%; p<.05). We conclude with implications for collecting data in web and telephone surveys.

Recruiting a Probability Sample of 18 year olds for a Longitudinal Study on Interpersonal Violence
David Cantor, *Westat*
Reanne Townsend, *Westat*

Recruiting a general population sample of young adults in the current survey environment is difficult. Short of doing in-person screening, using contact modes such as mail or the internet all have inherent difficulties. Young people do not readily respond to requests by mail or the internet. This presentation describes a recruitment of 18 year olds into a longitudinal study of inter-personal violence. The goal was to recruit a probability sample that can provide generalizable data on long-term trajectories of risk for and experiences with violence as young adults are transitioning out of the home to independent living. The design uses an address based sample (ABS) that is supplemented with a list of high school seniors to stratify and oversample households that were likely to have an 18 year old. The study has successfully recruited 1,800 young adults by using a combination of postal requests, pushing respondents to the internet, incentives and gamification methods. This presentation will provide an overview of key design features, the response rate and the efficiency of using these methods to recruit this difficult-to-survey group.

Using Crowdsourcing for Survey Administration: A Study of Innovation Activity among Individuals
Audrey Kindlon, *National Center for Science and Engineering Statistics*
Jesse Chandler, *Mathematica*
Rebecca Morrison, *National Center for Science and Engineering Statistics*

Innovation activity is typically studied within firms, but governments and individuals can also undertake innovation activity. Ignoring non-business sectors of the economy leads to an incomplete picture of innovation at the society level. However, there are numerous challenges to understanding non-business innovation. Individual innovation is assumed to be relatively rare in the general population and thus expensive to measure using probability-based samples. The lack of research on individual innovation leads to uncertainty about what topics to prioritize and how to ask about them should a probability-based sample ever be used to estimate individual innovation rates. In 2019, the National Center for Science and Engineering (NCSES) conducted a study of individual innovation using a crowdsourcing tool, Amazon Mechanical Turk (MTurk), for survey administration. The correlates of individual innovation activity observed in this sample were like those observed in other non-probability samples of innovators. Through this study we identify several challenges to efficiently screening for individual innovators and accurately measuring innovation activity that suggest improvements to existing measures of innovation.

Session E-2
Creative Problems in Social Science

Assessing the Drivers of U.S. Food Expenditures
Eliana Zeballos, *Economic Research Service*
Wilson Sinclair, *Economic Research Service*
Timothy Park, *Economic Research Service*

Expenditure on food and beverages in the United States reached $1.8 trillion in 2019. This expenditure included both food-at-home (FAH) establishments and food-away-from-home (FAFH) establishments. While total food expenditures have increased steadily through the decades, the share of expenditures at FAH establishments has decreased from about two-thirds half a century ago to 45 percent in 2019. To better understand changes in food spending, this study utilizes a structural decomposition analysis (SDA) to investigate the roles of income and propensity to spend changes in one framework. Specifically, the SDA framework decomposes food spending into four components: disposable personal income (DPI), personal consumption expenditures (PCE) as a share of DPI, total food spending as a share of PCE, and FAH (FAFH) as a share of total food spending. Using data from 1997 to 2020, this study assesses the relative effects of the drivers of food spending including recession periods. Results show that disposable personal income has a positive relationship with food expenditures, and it is the main driver of changes in food spending in non-recession years. While the decrease in overall food spending played a role, the main driver that contributed to the decrease in FAFH spending during the 2020 Recession was the substitution away from FAFH and towards FAH.

Estimation of Deaths among Health Care Personnel (HCP) with COVID-19 using Capture-Recapture Methods — United States, March 17–April 29, 2020
Jennifer Rammon, *Centers for Disease Control and Prevention*
Kerui Xu, *Centers for Disease Control and Prevention*
Matt Stuckey, *Centers for Disease Control and Prevention*
Michelle Hughes, *Centers for Disease Control and Prevention*
Reid Harvey, *Centers for Disease Control and Prevention*
Sherry Burrer, *Centers for Disease Control and Prevention*
Sophia Chiu, *Centers for Disease Control and Prevention*
Jess Rinsky, *Centers for Disease Control and Prevention*
Matthew Groenewold, *Centers for Disease Control and Prevention*

During the initial period of rapidly increasing transmission of SARS-CoV-2, there was limited information on COVID-19 deaths among health care personnel (HCP). We aimed to address this issue by using data from multiple sources to serve as a guide for conducting mortality surveillance. National case surveillance data on laboratory-confirmed COVID-19 reported from public health departments were matched with data abstracted from media reports via web-scraping techniques. The two data sources reported 316 HCP deaths (159 surveillance only, 133 media report only, and 24 using both sources) during March 17–April 29, 2020. An overall population estimate of 1,161 HCP deaths was calculated using Chapman's estimator. This presentation focuses on evaluating the robustness of Chapman's estimator since violations to the standard assumptions of independence between data sources and homogenous capture probability within sources seem likely. Sensitivity analyses identify negative correlation between data sources, suggesting that the Chapman estimate could be an overestimate. However, compared to methodology that accounts for source dependency, Chapman's estimator appears unbiased and more precise.

Measuring the US Space Economy

Tina Highfill, *Bureau of Economic Analysis*
Annabel Jouard, *Bureau of Economic Analysis*
Connor Franks, *Bureau of Economic Analysis*

Economic activity related to space exploration in the United States dates to the early 1800s with the construction of America's first observatories. Despite the long history of space economic activity in the United States and dominance of U.S. space spending relative to the rest of the world, there is a lack of consistent and comprehensive economic data about the U.S. space economy. To address this, the Bureau of Economic Analysis developed preliminary estimates of space economy gross output, GDP, private employment, and private compensation by industry for 2012-2018. The newly released statistics show in 2018, the U.S. space economy accounted for $177.5 billion of gross output, 0.5 percent ($108.9 billion) of current-dollar GDP, $41.2 billion of private industry compensation, and supported more than 356,000 private sector jobs. The space economy experienced slower growth in all four sets of statistics relative to the overall U.S. economy over the 2012–2018 period. Relatively slow growth was driven mainly by the Information and Manufacturing sectors, with strong growth in the Wholesale Trade sector partially offsetting these declines. These statistics are the first to shed light on the contribution of space-related goods and services to the U.S. economy using a framework consistent with how the overall U.S. economy is measured. However, additional research and resources are needed to develop an official time series of the entire U.S. space economy.

Session E-3
Advances in Disclosure Limitation and Publication Standards

Comparative Analysis of Differential Privacy and Swapping Methods in the Context of the U.S. Census
Miranda Christ, *Columbia University*
Sarah Radway, *Columbia University*

This work examines the data de-identification methods of swapping and differential privacy in the context of the privacy-utility tradeoff, studying how the manifestation of this tradeoff varies across subpopulations. We attempt to de-identify a dataset of personal information, first using a differentially private mechanism, similar to one that will be used by the U.S. Census Bureau in 2020, and second, using a swapping-based algorithm, similar to one likely used by the U.S. Census Bureau in 2010. We evaluate the accuracy performance of both mechanisms from the lens of a policy maker via aggregate statistical analysis, and we analyze claims regarding differential privacy's potential to adversely affect minority groups. We then evaluate the privacy guarantees of both mechanisms, both theoretically, by proving properties of our mechanisms, and empirically, by simulating linkage and reconstruction attacks. We discuss the strengths and weaknesses of differential privacy as a method of census data de-identification, focusing on its performance in minority subpopulations. Our results justify the U.S. Census Bureau's switch to differential privacy, but identify unavoidable limitations.

Evaluating Publication Rules for the County Agricultural Production Survey Using Simulation
Andrew Dau, *National Agricultural Statistics Service*
Nathan Cruze, *National Agricultural Statistics Service*
Joe Parsons, *National Agricultural Statistics Service*
Linda Young, *National Agricultural Statistics Service*

Annually, the National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA) publishes county level estimates of acreage and yield for several principal crops in the United States. NASS makes publication decisions based on a combination of factors including the number of reports and coverage. This paper evaluates the current NASS publication rules for county estimates by conducting a simulation study using administrative USDA data. Through the simulation study we are able to compare a variety of potential publication rules looking at metrics of coverage, sample size, coefficient of variation, and accuracy to evaluate both our current publication rule and future potential publication rules.

Posterior Risk and Utility from Private Synthetic Weighted Survey Data
Quentin Brummet, *NORC at the University of Chicago*
Jeremy Seeman, *Pennsylvania State University*

Differentially private (DP) synthetic data methods offer survey administrators the ability to share synthetic data while limiting the probabilistic risk of disclosing information about individuals. However, applying these methods can be difficult given the desire to be transparent about survey methodology such as sampling design and weighting scheme. This yields additional potential disclosure risk and makes it difficult to interpret tuning parameters in DP such as the privacy budget. To address this issue, we present methods for joint synthesis of survey responses and weights that accommodate methodological disclosure, allowing for sample size adjustments based on whether sample design and weighting methodology are public knowledge. We apply these methods to generate DP synthetic data from a survey containing information about the status of food allergies in adults, where the final inference is adjusted for privacy-preserving measurement error.

Private Tabular Survey Data Products through Synthetic Microdata Generation

Terrance Savitsky, *Bureau of Labor Statistics*
Matthew Williams, *National Center for Science and Engineering Statistics*
Jingchen Hu, *Vassar College*

We propose three synthetic microdata approaches to generate private tabular survey data products for public release. We adapt a disclosure risk based-weighted pseudo posterior mechanism to survey data with a focus on producing tabular products under a formal privacy guarantee. Two of our approaches synthesize the observed sample distribution of the outcome and survey weights, jointly, such that both quantities together possess a probabilistic differential privacy guarantee. The privacy-protected outcome and sampling weights are used to construct tabular cell estimates and associated standard errors to correct for survey sampling bias. The third approach synthesizes the population distribution from the observed sample under a pseudo posterior construction that treats survey sampling weights as fixed to correct the sample likelihood to approximate that for the population. Each by-record sampling weight in the pseudo posterior is, in turn, multiplied by the associated privacy, risk-based weight for that record to create a composite pseudo posterior mechanism that both corrects for survey bias and provides a privacy guarantee for the observed sample. Through a simulation study and a real data application to the Survey of Doctorate Recipients public use file, we demonstrate that our three microdata synthesis approaches to construct tabular products provide superior utility preservation as compared to the additive-noise approach of the Laplace Mechanism. Moreover, all our approaches allow the release of microdata to the public, enabling additional analyses at no extra privacy cost.

Concurrent Session E-4
Communicating Fitness for Use

Using Data Visualizations, Short Articles, and Social Media to Communicate Complex Data Simply
Jay Meisenheimer, *Bureau of Labor Statistics*

Social media and other technology make it easy for people to find information, but the information they find isn't always accurate and unbiased. I discuss our approach at the U.S. Bureau of Labor Statistics of using data visualizations, short articles, blogs, and social media to help people understand what's going on in the labor market and economy. Using these short- form communications includes explaining not just what we know but what we don't know, while being transparent about the strengths and limitations of the data.

Journalists Communicating Statistical Information
Regina Nuzzo, *American Statistical Association*

Journalists are experts at conveying complex information in short missives (and for audiences with short attention spans). Can we learn anything from them? In this session I will present examples of successful journalistic communication of statistical information and explore their commonalities for insights that could be applied to statistical communication in the government.

Short Communication as a Medium: Is Engagement a Substitute for Efficacy?
Travis Hoppe, *National Center for Health Statistics*

We provide a case-study of various official Twitter accounts from Federal and State agencies that communicate both statistical information to the larger public. Viral tweets may reach a larger audience, but do these tweets come at the expense of accurate or timely information? We examine communication strategies that been successful and ways to help official tweets reach the intended audience. We also examine how the absence of an authoritative source, such as a statistical agency, may be supplanted by other actors via unintentional misinformation or malicious propaganda.

Session E-5
Collecting Spending Data during a Pandemic – an Evaluation of Quality and Response in the Consumer Expenditure Surveys

An Examination of Nonresponse Bias in the Consumer Expenditures Survey during the COVID-19 Period
Stephen Ash, *Bureau of Labor Statistics*
Brian Nix, *Bureau of Labor Statistics*
Barry Steinberg, *Bureau of Labor Statistics*
David Swanson, *Bureau of Labor Statistics*

Like many household surveys, the Consumer Expenditure Interview and Diary Surveys' response rates decreased last year due to COVID-19. The decrease naturally raised questions about differences between the surveys' respondents and nonrespondents, and how much nonresponse bias, if any, those differences generated in survey estimates during the COVID-19 period. In this presentation we describe a method of estimating the amount of nonresponse bias in the Consumer Expenditure Survey by developing a nonresponse adjustment process tailored to the COVID-19 period, and then comparing the survey estimates generated from the tailored process to the estimates generated from the normal production process. A second analysis comparing the distribution of demographic characteristics between the Consumer Expenditure Surveys and the American Community Survey is also presented.

Consumer Expenditure Interview Survey: Data Quality Assessment Pre vs. Post COVID-19
Yezzi Angi Lee, *Bureau of Labor Statistics*
David Biagas, *Bureau of Labor Statistics*

In response to the COVID-19 pandemic, the Consumer Expenditure Interview Survey shifted to all telephone interviewing for the health and safety of both interviewers and respondents in late March 2020. The purpose of this study is to evaluate whether the data quality of the Interview Survey was affected by the emergence of the COVID-19 pandemic and the changes in data collection method. In this research, we attempt to isolate data quality changes due to the change in protocol to find the real effect of the COVID-19 on Interview Survey. We explored household characteristics and data quality metrics (e.g. total expenditures, survey time, number of rounded items, number of entries, etc.) by changes in mode of the interview before and after the onset of the COVID-19 pandemic. We also examined the effect of the COVID-19 pandemic on key outcomes using the discontinuous growth curve models, controlling for several covariates related to data collection and household characteristics.

Evaluating Diary Collection Mode Changes in the Context of the COVID-19 Pandemic
Brett McBride, *Bureau of Labor Statistics*
Nikki Graf, *Bureau of Labor Statistics*

The onset of the COVID-19 pandemic prompted changes to the survey collection methods of the Consumer Expenditure Diary Survey due to restrictions on personal visits, which limited the use of paper diaries to collect expenditures from respondents. This led to an accelerated roll-out of an online diary in June of 2020. English-speaking households that had regular internet access were provided a link to a diary, where a respondent could enter the household expenses for two weeks via a computer or mobile device. For those disinclined or not eligible to use the online diary, interviewers collected their expenses via telephone. Data resulting from the modified collection methods will be compared to those from standard collection methods. We compare diary entries, reported expenditure amounts, and completeness of reporting, while controlling for changes in demographic characteristics of sampled households. We also examine respondent adherence to instructions about diary keeping (e.g., itemizing grocery expenses). This presentation can inform how collection modes impact data quality, at a time when the pandemic led to significant changes in household spending patterns.

COVID-19's Effect on the Consumer Expenditure Surveys' Estimates

Scott Curtin, *Bureau of Labor Statistics*
Bryan Rigg, *Bureau of Labor Statistics*
Brett Creech, *Bureau of Labor Statistics*

As a result of the COVID-19 pandemic, the Consumer Expenditure Surveys (CE) endured change, both in terms of data collection and processing, and economic change as it pertains to spending patterns among consumers. The U.S. Census Bureau, who collects data for the CE Surveys on behalf of the Bureau of Labor Statistics, ceased in-person data collection for both the CE Interview and the Diary Surveys. Interviews were shifted to telephone, and an online diary was added as an alternative collection means. This year's expenditure data will use all of the expenditures collected throughout the first 3 quarters of the COVID-19 pandemic and will help us understand the true impact on consumers. This study aims to gauge the true impact of this pandemic on the economy as a whole, as well as looking at the effects of this pandemic on several metrics to gauge the impact of decreasing response rates, as well as changes in data collection. Similar to how we plan to present the changes in spending, we will also look at estimates among different demographic groups to see if specific groups were impacted differently by these conditions."

Session F-1
Supplementing Data Collection and Research at the National Center for Health Statistics using the Research and Development Survey

Introduction to the Research and Development Survey and calibration approaches for panel surveys
Katherine Irimata, *National Center for Health Statistics*

The National Center for Health Statistics (NCHS) monitors the health of the U.S. through various data systems including household surveys, establishment surveys, physical health assessments, and vital records. NCHS also conducts the Research and Development Survey (RANDS), a series of commercial panel surveys, which collects data primarily through web administration for methodological research purposes. The RANDS program started in 2015 and has been used to collect information on several health-related topics including chronic conditions, physical activity, disability, and opioids. The flexibility of RANDS makes it a valuable tool for survey methodology and questionnaire design research. Due to potential bias in web-based panel surveys compared to traditional household surveys, NCHS has been performing research on calibration methods to adjust the RANDS panel weights for estimation using national household surveys. This talk introduces RANDS and discusses considerations for calibrating panel survey weights.

An Overview of the 2019 Research and Development Survey (RANDS)
Li-Yen Rebecca Hu, *National Center for Health Statistics*
Paul Scanlon, *National Center for Health Statistics*
Kristen Miller, *National Center for Health Statistics*
Yulei He, *National Center for Health Statistics*
Katherine Irimata, *National Center for Health Statistics*
Guangyu Zhang, *National Center for Health Statistics*
Kristen Cibelli Hibben, *National Center for Health Statistics*

The National Center for Health Statistics (NCHS) launched the Research and Development Survey (RANDS) series in 2015 to investigate the use of commercial probability panels for evaluating questionnaire design and developing statistical methodology. In 2019, NCHS contracted NORC at the University of Chicago (NORC) to conduct the third round of the RANDS (RANDS 3) as a cross-sectional survey administered in web-mode only. Probe questions and four sets of experiments were embedded in RANDS 3 to assess question-response patterns, and participants were randomized into two groups for each set of experiments. Among the 4,255 sampled individuals, 2,646 completed RANDS 3, resulting in a completion rate of 62% and a weighted cumulative response rate of 18%. In this presentation, we give an overview of RANDS 3 on questionnaire design and randomized experiments. Participants' characteristics and key outcome measurements are summarized and compared between the two randomized groups involved in one set of experiments, encompassing topics on affect, self-rated health, pain frequency and electronic cigarettes.

Comparison of Mental Health Estimates by Sociodemographic Characteristics in the Research and Development Survey and National Health Interview Survey, 2019

Leanna Moron, *National Center for Health Statistics*
Katherine Irimata, *National Center for Health Statistics*
Jennifer Parker, *National Center for Health Statistics*

Due to the prevalence of mental health disorders in the United States, it is important for national surveys to be able to accurately measure and report estimates of major depressive disorder (depression) and generalized anxiety disorder (anxiety). This research compares national and subgroup estimates of depression and anxiety among the civilian, noninstitutionalized adult U.S. population from two data sources, the 2019 National Health Interview Survey (NHIS) and the third round of the Research and Development Survey (RANDS 3). The mental health subgroup estimates were compared by the following sociodemographic characteristics: age, sex, race and Hispanic origin, education, and Census region. The eight-item Patient Health Questionnaire (PHQ-8) was used to measure the severity of depressive symptomology. The seven-item Generalized Disorder Scale (GAD-7) was used to measure the severity of symptoms pertaining to generalized anxiety disorder. The results indicate that national estimates of depression were comparable across the two data sources, though there were significant differences for select subgroup estimates. For estimates of generalized anxiety disorder, there were significant differences for both national and subgroup-level estimates between the NHIS and RANDS 3. Potential factors for the observed differences in subgroup estimates may be due to differences in sample sizes and response rates by survey data source. Alternative methods such as weight adjustments may be helpful for obtaining more comparable subgroup estimates.

Estimates from Selected Variables in Three Rounds of RANDS during COVID-19 Pandemic

Rong Wei, *National Center for Health Statistics*
Yulei He, *National Center for Health Statistics*
Van Parsons, *National Center for Health Statistics*
Paul Scanlon, *National Center for Health Statistics*

The Research and Development Survey (RANDS) is based on a web-panel platform designed for conducting questionnaire evaluation and statistical research. Conducted by the National Center for Health Statistics (NCHS), RANDS has been an ongoing series of surveys since 2015. Availability of the RANDS platform allows NCHS to explore the capability of producing more timely focused data releases than are possible when using traditional data systems. During the COVID-19 pandemic in 2020, the survey added a component, RANDS during COVID-19, to focus on the general population's health experience. This survey was conducted in three closely spaced time periods. The first two were based on a longitudinal design structure, while the third period selected different subjects, but from the same geographical clusters as the first period. We propose statistical methods to assess differences of RANDS estimates collected at different time periods, accounting for the longitudinal structure and panel design of RANDS during COVID-19. These methods are demonstrated using variables such as self-rated health status, health insurance coverage, and outcomes of anxiety and depression.

Session F-2
From Data Collection to Estimation: Highlights of Survey Lifecycle Issues from FoodAPS

Is "Proof of Purchase" Really Proof?
Adam Kaderabek, *University of Michigan*
Brady T. West, *University of Michigan*
John A. Kirlin, *Kirlin Analytic Services*
Elina T. Page, *Economic Research Service*
Jeffrey M. Gonzalez, *Economic Research Service*

The National Household Food Acquisition and Purchase Survey (FoodAPS) fielded a subsequent Alternative Data Collection Method (ADCM) study requesting that respondents submit images of their food-purchase receipts. The ADCM aimed to investigate the likelihood respondents would provide images of receipts and to what end that data could be leveraged to reduce the overall reporting burden and improve data quality. Review of the receipt images found multiple sources of error that reduced the efficacy of the receipts as a source of reliable data, but within events where a receipt could be expected, we found that approximately 40% of events maintained an itemized and legible receipt. We employ logistic regression to study the relationships between respondent characteristics and the likelihood of submitting an itemized receipt. Our findings indicate significant influences related to respondent, household, and event characteristics, and these differences indicate that targeted interventions could possibly increase the frequency of submission as well as reduce errors associated with receipt quality. We explore possible intervention protocols and the implications for future data collection.

Usability Evaluation of Smartphone-based Data Collection Instrument
Lin Wang, *U.S. Census Bureau*
Anthony Schulzetenberg, *U.S. Census Bureau*
Alda G. Rivas, *U.S. Census Bureau*
Heather Ridolfo, *Energy Information Administration*
Shelley Feuer, *U.S. Census Bureau*

The second National Household Food Acquisition and Purchase Survey (FoodAPS-2) will leverage latest technology to use a smartphone app for data collection. The app is code-named FoodLogger. In order to ensure the viability of using FoodLogger and the quality and completeness of data collected via FoodLogger, an independent usability evaluation of FoodLogger is required. This paper presents an overview of the usability evaluation. The usability evaluation involves multiple iterations through the instrument development lifecycle. Each iteration consists of a training session, 7-day field data collection, laboratory-based usability testing, and debriefing sessions. A sample of households, representing different socio-economic status and age groups, will participate in the evaluation. Twenty-two critical tasks were identified for data entry to FoodLogger. The laboratory-based usability testing was designed such that all critical tasks are to be tested in three use cases: Food-at-home event, Food-away-from-home event, and School-meal event. Qualitative and quantitative data will be collected and analyzed to assess participants' performance of food information entry to FoodLogger. Important findings will be reported in this presentation.

Examination of the Data Quality Properties of USDA and Proprietary Databases with Information on Food Item and Food Retailers to Reduce Nonsampling Errors in FoodAPS-2

Clare Milburn, *The George Washington University School of Public Health*
Jeffrey M. Gonzalez, *Economic Research Services*
Linda Kantor, *Economic Research Service*
Elina T. Page, *Economic Research Service*

The National Food Acquisition and Purchase Survey (FoodAPS) collects comprehensive data about household food acquisitions. The ability to link these data to USDA and proprietary databases containing information on food item and food place attributes allows for a richer set of outputs and addresses a broader set of research questions than the FoodAPS could alone. Linkage is also anticipated to reduce respondent burden and improve data quality as the data are being collected. Although linkage is an important feature of the overall data collection and processing strategy planned for FoodAPS-2, there are still barriers to leveraging the full capacity of the linked databases. And while crosswalks among the databases exists, there are operational issues (e.g., prioritization when a data element links to multiple records, handling missing attributes) associated with their integration within a dynamic data collection instrument. This presentation synthesizes the data quality properties of several external databases containing information on food item and food places and necessary developing crosswalks among them and makes recommendations for their implementation and use in FoodAPS-2.

Using the Weighted Finite Population Bayesian Bootstrap to Account for Complex Sample Design Features when Estimating State- and Sub-state-level Food Insecurity Prevalence

Katherine Li, *University of Michigan*
Yajuan Si, *University of Michigan*
Brady T. West, *University of Michigan*
John A. Kirlin, *Kirlin Analytic Services*
Xingyou Zhang, *Bureau of Labor Statistics*

Model-based small area estimation (SAE) approaches fit models to survey data and use external data from auxiliary or population records for prediction. We demonstrate this process by estimating food insecurity prevalence at the tract, county, and state levels in the U.S., modeling with data from the National Food Acquisition and Purchase Survey (FoodAPS) and predicting with external data integrated from the American Community Survey (ACS). The methodological challenge is accounting for the multi-stage complex sampling design features and survey weights of the FoodAPS. To solve this problem, we implement the weighted finite population Bayesian bootstrap (WFPBB), which propagates the sampling uncertainty and model estimation in a single workflow and streamlines estimation of SAEs and their standard errors. We compare the accuracy of standard error estimates from the WFPBB with alternative methods such as jackknife replication and pseudo-likelihood approaches using simulation studies. We contrast the food insecurity prevalence estimates based on these methods against those from the CPS as a proxy for external validation.

Session F-3
Privacy and Sample Surveys

Adapting Surveys for Formal Privacy: Where We Are, Where We Are Heading
Aref Dajani, *U.S. Census Bureau*

The U.S. Census Bureau implemented formal privacy in 2008 and is implementing formal privacy for the full release of the 2020 Census. This presentation begins with an overview of legacy disclosure avoidance methods currently used for Census Bureau surveys, then transitions to where the Census Bureau is today. The Census Bureau is committed to moving towards formal privacy for its surveys, with the understanding that there remains science to develop through its collaboration with international experts in formal privacy. This science includes the consideration of survey weights, longitudinality, and aggregate statistics such as medians, totals, and model parameters. The Census Bureau is well aware of the greater risks involved in implementing formal privacy. The Census Bureau will use the rule for Geographic Areas with Small Populations as a stepping-stone to the end goal of implementing formal privacy for its surveys.

Controlling Privacy Loss in Survey Sampling
Mark Bun, *Boston University*
Joerg Drechsler, *University of Maryland*
Marco Gaboardi, *Boston University*
Audra McMillan, *Apple*
Jayshree Sarathy, *Harvard University*

Social science and economics research is often based on data collected in surveys. Due to time and budgetary constraints, this data is often collected using complex sampling schemes designed to increase accuracy while reducing the costs of data collection. A commonly held belief is that the sampling process affords the data subjects some additional privacy, since their data may, or may not, be in the final sample.
This intuition has been formalized in the differential privacy literature for simple random sampling: a differentially private mechanism run on a simple random subsample of a population provides higher privacy guarantees than when run on the entire population.
In this work we initiate the study of the privacy implications of more complicated sampling schemes including cluster sampling and stratified sampling. We find that not only do these schemes often not amplify privacy, but that they can result in privacy degradation.

Leveraging Public Data for Practical Private Query Release
Terrance Liu, *Carnegie Mellon University*
Giuseppe Vietri, *University of Minnesota*
Thomas Steinke, *Google*
Jonathan Ullman, *Northeastern University*
Steven Wu, *Carnegie Mellon University*

In many statistical problems, incorporating priors can significantly improve performance. However, the use of prior knowledge in differentially private query release has remained underexplored, despite such priors commonly being available in the form of public datasets, such as previous U.S. Census releases. With the goal of releasing statistics about a private dataset, we present PMWfPub, which -- unlike existing baselines -- leverages public data drawn from a related distribution as prior information. We provide a theoretical analysis and an empirical evaluation on the American Community Survey (ACS), which shows that our method outperforms state-of-the-art methods.

Session F-4
Considerations for Calculating and Communicating the Value of Federal Statistics

Panel Discussion: Reflections on Census Quality Indicators Taskforce

- Constance F. Citro - Senior Scholar, Committee on National Statistics
- Julia Lane, Professor - NYU and Cofounder, Coleridge Initiative
- Joseph Salvo – Institute Fellow, Social and Data Analytics Division, University of Virginia Biocomplexity Institute; Senior Advisor, National Conference on Citizenship

Discussant: Andrew Reamer – Research Professor, George Washington Institute of Public Policy, George Washington University

Concurrent Session F-5
Web Scraping

Detecting and Measuring Product Innovation in News Articles Using Natural Language Processing Methods

Gizem Korkmaz, *University of Virginia*
Neil Alexander Kattampallil, *University of Virginia*
Gary Anderson, *National Center for Science and Engineering Statistics*

Innovation, the availability and usage of novel ideas, products and business practices is central to the improvement of living standards. Policy makers in part rely on survey-based measures of innovation to design, develop, and implement policies to promote innovation. In the U.S., innovation is measured through nationally representative surveys of businesses such as the Annual Business Survey. To reduce respondent fatigue and to provide more timely information, statistical organizations are interested in exploring non-traditional methods for measuring innovation. In this paper, our goal is to show how large corpus of opportunity data, in particular news articles, and advanced natural language processing methods can be used to identify and to measure innovation in various sectors (food and beverage, pharmaceutical, and computer software). We present a novel approach utilizing Bidirectional Encoder Representation from Transformers (BERT) developed by Google. Our methods include (i) text classification to identify news articles that mention innovation, (ii) named-entity recognition (NER), and (iii) question answering (QA) to extract company names and (iv) developing yearly innovation indicators for companies in these sectors.

Machine-Learning Based Identification of Emerging Research Topics Using Research & Development Administrative Data

Eric J. Oh, *University of Virginia*
Kathryn Linehan, *University of Virginia*
Joel Thurston, *University of Virginia*

Nearly 22% of all Research and Development (R&D) funding in the United States is provided by the Federal Government (Table 3, https://ncses.nsf.gov/pubs/nsb20203, 2020), yet there is little information describing the content areas that this public funding supports. To address this issue, we use Federal RePORTER, a repository of publicly available administrative data on R&D grants, to identify the range of R&D topics funded by federal science and technology agencies, including trends in the pattern of funding across time (e.g., identifying emerging research areas). We use natural language processing and machine learning algorithms to automatically extract information from the large amount of text-based data. Specifically, we build topic models to identify latent research topics within the corpus of R&D grants. In addition, we build an information retrieval pipeline to identify relevant grants for specific research topics of interest (e.g. coronavirus, artificial intelligence), which are then used for generating topic models within a specific subject area.

Using Web Scraping and Network Analysis to Study International Collaboration in Open Source Software

Brandon Kramer, *University of Virginia*
Gizem Korkmaz, *University of Virginia*
José Bayoán Santiago Calderón, *University of Virginia*
Carol Robbins, *National Center for Science and Engineering Statistics*

Over the past two decades, international collaboration has more than doubled in academic research. At the same time, the open source software community has burgeoned from a collection of small, dispersed communities to a multi-billion dollar industry spanning several prominent industrial sectors around the world. To date, few studies have examined the structure of open source software development as a transnational collaboration system. In this paper, we study international collaboration networks in the open source community using data scraped from GitHub - the world's largest remote-hosting repository platform. After collecting data from roughly 740,000 GitHub users from 214 different countries, we analyze longitudinal trends for both contributor- and country-level network data from 2008-2019. Our findings demonstrate that the contributor-level networks have grown exponentially while simultaneously becoming less dense, less centralized, and less transitive over time. In this network, GitHub users from the US have a disproportionately higher impact on collaborative efforts, as indexed by the fraction of contributions from other countries and various centrality measures. This influence carries over to the country-level networks where most nations around the world are more likely to collaborate with the US than they are to collaborate with any other country, including their own. More generally, we find that the country-level network has become more structurally integrated over time, translating to some countries, like China and India, gaining more influence in the open source community. In addition to offering novel insights about the history of open source collaboration tendencies, this paper also raises a number of important questions for future research to address.

Progress in the Use of Web-Scraped List Frames and Capture-Recapture Methods: Insights from a National Farmers Markets Managers Survey
Michael Jacobsen, *National Agricultural Statistics Service*
Linda Young, *National Agricultural Statistics Service*

Surveys are often based on a sample drawn from a list frame. In recent years, the percentage of target population units on the list frames has been decreasing, making it important to adjust for this undercoverage in the estimation process. In 2020, NASS conducted the National Farmers Market Mangers (NFMM) Survey. Because NASS does not include farmers markets on its list frame, the USDA Agricultural Marketing Service (AMS) business register of farmers markets was the only list frame initially available. To assess its undercoverage, a web-scraped list frame was developed, and capture-recapture methods provided the foundation for estimation. In this paper, the two advances in the use of capture-recapture methods when conducting a survey with two list frames are discussed: (1) the sample design incorporated information identifying records on only the AMS business register, on only the web-scraped list frame, or on both frames and (2) a composite estimator for this overlap design allowed full use of all sample information to produce survey estimates. Directions for future research are highlighted.

Using a Web-scraped List Frame for an Agricultural Survey
Habtamu Benecha, *National Agricultural Statistics Service*
Bruce A. Craig, *National Agricultural Statistics Service*
Grace Yoon, *National Agricultural Statistics Service*
Zachary Terner, *National Agricultural Statistics Service*
Denise A. Abreu, *National Agricultural Statistics Service*
Linda J. Young, *National Agricultural Statistics Service*

USDA's National Agricultural Statistics Service (NASS) uses the annual June Area Survey (JAS) to produce comprehensive estimates of land uses and agricultural activities across the US. Because conducting the JAS costs NASS a significant proportion of its annual budget, the agency is exploring ways to lower costs by leveraging new statistical methods and technologies. As a part of this effort, NASS designed a pilot project aimed at assessing the viability of replacing the JAS area-frame for some or all of the U.S. with a web-scraped frame. The pilot study involved four states and data were collected from both the web-scraped list frame and NASS's main list frame in two-phases. In this paper, a capture-recapture methodology, based on the data collected from these two phases, is used to estimate land-use activities. These estimation methods are described, and the accuracy of the estimates and the measures of list-frame undercoverage are compared to results from the JAS. Furthermore, challenges and strengths are discussed, and potential improvements are proposed based on results from simulation studies.

Session G-1
Reply YES to Innovate: Text Messages for Federal Surveys

Texting Interviewers to Encourage Proper Protocols in the Survey of Income and Program Participation
Kevin Tolliver, *U.S. Census Bureau*

The Survey of Income and Program Participation (SIPP) is a face-to-face longitudinal household survey, conducted by the U.S. Census Bureau, with about 53,000 households. SIPP interviewers have a four- to five-month span to work about 35 cases. Although there is plenty of requisite training on proper protocols that has to be completed before any interviewer is allowed to work their cases, these protocols are not enforced and how interviewers choose to work their caseloads is left to their discretion. Since the 2019 data collection, SIPP has experimented with sending SMS text messages to their interviewers to encourage proper protocols.

In the 2019 SIPP, the content and timing of these messages were partially randomized each week. In the following data collection, content and timing were fully randomized monthly, with each month designed to answer a different research question. In the most recent data collection, additional text message content options were added based on activity in the prior two weeks, and their timing was based on prior years' information. This research summarizes the results of the experiments and whether they can be used to tailor future interventions.

How Should We Text You? Designing and Testing Text Messages for the 2021-22 Teacher Follow-Up Survey (TFS) and Principal Follow-Up Survey (PFS)
Jonathan Katz, *U.S. Census Bureau*
Kathleen Kephart, *U.S. Census Bureau*
Jasmine Luck, *U.S. Census Bureau*
Jessica Holzberg, *U.S. Census Bureau*

There is interest in adding text messaging as a contact and/or a response mode to many surveys. However, good mobile phone numbers are not always available, and there are constraints around texting people who have not opted into receiving text messages. As a result, there are unanswered questions about how to implement text messaging into a data collection strategy. The National Center for Education Statistics and Census Bureau have an opportunity to add a text messaging mode to the data collection strategy for the 2021-22 Teacher Follow-Up Survey (TFS) and Principal Follow-Up Survey (PFS). These are follow-up surveys administered to teachers and principals who complete the 2020-21 National Teacher and Principal Survey (NTPS). In the NTPS, respondents can consent to receive text messages for follow-up. In the upcoming TFS/PFS a sample of respondents will be assigned to participate in the survey by navigating to a web link embedded in the text messages or by answering the questions via two-way SMS. Prior to the fielding of the 2021-22 TFS/PFS, we conducted remote cognitive and usability testing of the messages, including using Qualtrics text messaging capabilities and mobile screensharing. We will discuss our methodology for testing and review participants' impressions of the text messages. Findings will help inform best practices as more surveys use text messaging in data collection.

Yes, I Consent to Receive Text Messages: Conducting Follow-Up Text Surveys with Principals and Teachers
Maura Spiegelman, *National Center for Education Statistics*
Allison Zotti, *U.S. Census Bureau*

The U.S. Department of Education's National Teacher and Principal Survey (NTPS) collects data from schools, principals, and teachers. Select administrations are followed by the Principal Follow-up Survey (PFS) and Teacher Follow-up Survey (TFS) to measure staff attrition, that is, whether a principal or teacher is a stayer (same job at the same school), mover (at a different school), or leaver (no longer in the profession) during the following school year. The TFS also includes a longer survey for both current and former teachers. The NTPS asks responding principals and teachers to provide contact information, including cellphone numbers, and the 2020-21 NTPS asked respondents to check a box indicating "I consent to receive text messages for follow-up purposes only." For the upcoming PFS and TFS, consenting principals and teachers may be contacted to complete a text message survey, and teachers may receive a link to complete their longer web surveys. This presentation discusses who consents to receive text messages, the experimental design of this texting operation, and evaluation metrics to determine whether employment status can be successfully collected by text message

The Future of SMS and Email in Federal Surveys
Jennifer Hunter Childs, *U.S. Census Bureau*

In March 2020 the Census Bureau launched the Household Pulse Survey to measure social and economic impacts of COVID-19. Due to the COVID-19 pandemic, the Bureau's mail and phone centers were closed and field visits were suspended, therefore the Household Pulse Survey was conducted using only SMS and email invitations. This represented the first large scale survey conducted by the Bureau (maybe by the Government) using only email and SMS as contact modes. The viability of this innovation prompted the Census Bureau to start investigating the use of SMS for the future. This talk will discuss the success rates of using SMS and email invitations in the Household Pulse Survey and then will discuss lessons learned from the use of SMS thus far and applications for the future. As we move forward with this endeavor, it will be necessary to establish an infrastructure for using SMS across the enterprise. This presentation will also discuss developments along these lines.

Session G-3
Statistical Methods for Improving Small Domain and Key Survey Estimates

A Practical Framework for Area-level Small Area Estimation
Stephanie Zimmer, *RTI International*
Dan Liao, *RTI International*
Rachel Harter, *RTI International*

The process of developing models and obtaining administrative data for small area estimation (SAE) requires careful consideration in both model selection and data sourcing. Few examples exist in literature of detailed approaches to all steps of the process from administrative data acquisition to model building to variance estimation. In this review paper, we discuss the practical framework to SAE and give examples of two projects and how they align with the framework. This framework will enable researchers new to SAE to see the entire picture, not only the theoretical underpinnings of the models. The first survey is the National Crime Victimization Survey, a national longitudinal survey of the civilian, noninstitutionalized population aged 12 or older with a rotating panel sampling design. The second is the Ohio Medicaid Assessment Survey, a survey of the child and adult population in residential households in Ohio. The designs, modes, models, auxiliary data sources, and areas of analysis differed between the two SAE implementations, but we will discuss how they follow a common framework that can be implemented on other surveys.

Using American Community Survey Data to Improve Estimates from Smaller U. S. Surveys through Bivariate Small Area Estimation Models
William R. Bell, *U.S. Census Bureau*
Carolina Franco, *National Opinion Research Center*

We demonstrate the potential for borrowing strength from estimates from the American Community Survey (ACS), the largest U.S. household survey, to improve estimates from smaller U.S. household surveys. We do this using simple bivariate area-level models to exploit strong relationships between population characteristics estimated by the smaller surveys and ACS estimates of the same, or closely related, quantities. Applications presented show impressive variance reductions from this approach for state estimates of health insurance coverage from the National Health Interview Survey, for state estimates of disability from the Survey of Income and Program Participation, and for modeling ACS five-year estimates of poverty of school-age children jointly with corresponding ACS one-year estimates. The models achieve large variance reductions without using regression covariates drawn from auxiliary data sources.

Using Bayesian Regression Models in Small Sample Size Contexts to Support System-level Education Decision-making
Bradley Rentz, *REL Pacific at McREL International*
Christina Tydeman, *REL Pacific at McREL International*

This presentation discusses how the Regional Educational Laboratory of the Pacific (REL Pacific), funded by the Institute of Education Sciences at the U.S. Department of Education, has used Bayesian regression models to inform education decision making around college readiness. In the Pacific region, small sample sizes are a common limitation for quantitative research. The solution to small samples sizes is often to simplify regression models to avoid computational or estimation issues. In contrast to frequentist models, Bayesian regression models provide many advantages for researchers working with small sample sizes, including potentially more meaningful results with actionable next steps for practitioners. However, while common in other fields, Bayesian models are used infrequently in education research. In this paper, we discuss how using Bayesian models helped inform system-level education decision-making around college readiness on Pohnpei in the Federated States of Micronesia and how the methods used can inform education research practices more broadly.

Utilizing Occupational Employment and Wage Statistics (OEWS) Survey to Improve Small Domain Estimation (SDE) in the Occupational Requirements Survey (ORS)

Xingyou Zhang, *Bureau of Labor Statistics*
Erin McNulty, *Bureau of Labor Statistics*
Ellen Galantucci, *Bureau of Labor Statistics*
Patrick Kim, *Bureau of Labor Statistics*
Joan Coleman, *Bureau of Labor Statistics*
Tom Kelly, *Bureau of Labor Statistics*

The Occupational Requirements Survey (ORS) is an establishment-based survey and provides job-related information about the physical demands; environmental conditions; education, training, and experience; as well as cognitive and mental requirements in the U.S. economy. However, more than 60% of estimates for 843 SOCs (6-digit 2018 Standard Occupational Classification (SOC) system) under the ORS sampling frame are not publishable because of lack of sample or small sample sizes. In order to improve ORS estimates, we have explored and developed a multilevel Small Domain Estimation (SDE) approach that utilizes the Occupational Employment and Wage Statistics (OEWS) survey, a semiannual survey designed to produce estimates of employment and wages for specific occupations, with an annual sample size of nearly 400,000. This Multilevel Regression and Poststratification (MRP) approach includes two basic steps: 1) multilevel statistical models are constructed with ORS data; and 2) the fitted multilevel models are applied to OWES data to produce estimates for detailed occupations. It has produced reliable estimates for 840 out of 843 SOCs.

Statistical Data Integration Using Multilevel Models to Predict Employee Compensation

Andreea L. Erciulescu, *Westat*
Jean D. Opsomer, *Westat*
Benjamin J. Schneider, *Westat*

Considered in this paper is the case where two surveys collect data on a common variable, with one survey being much smaller than the other. The smaller survey collects data on an additional variable of interest, related to the common variable collected in the two surveys, and out-of-scope with respect to the larger survey. Estimation of the two related variables is of interest at domains defined at a granular level. We propose a multilevel model for integrating data from the two surveys, by reconciling survey estimates available for the common variable, accounting for the relationship between the two variables, and expanding estimation for the other variable, for all the domains of interest. The model is specified as a hierarchical Bayes model for domain-level survey data and posterior distributions are constructed for the two variables of interest. A synthetic estimation approach is considered as an alternative to the hierarchical modeling approach. The methodology is applied to wage and benefits estimation using data from the National Compensation Survey and the Occupational Employment Statistics Survey, available from the Bureau of Labor Statistics, Department of Labor, United States.

Session G-4
Data Quality – Nonresponse Bias

2017 Census of Agriculture Non-Response Sample
Mark Apodaca, *National Agricultural Statistics Service*
Peter Quan, *National Agricultural Statistics Service*
Franklin Duan, *National Agricultural Statistics Service*

The 2017 Census of Agriculture (COA) conducted by the National Agricultural Statistics Service (NASS) was administered through the internet, mail and phone. Operations with an email address on NASS's Census Sampling Frame received instructions to complete the COA form on-line. All others received a COA form in the mail. Operations that did not respond by a predetermined time were contacted by phone to complete the COA form. The nonresponse group was categorized into high and low priority data collection groups. Every operation in the high priority group was contacted by phone. Under normal circumstances, operations in the low priority group would have been sorted based on a hierarchal scheme, and an attempt made to contact all operations by phone. In March 2018, the COA non-response rate was much higher than that of previous COA's during this same time frame; hence, the number of operations to contact via phone was relatively high. To address the issue arising from the higher non-response rate, a sample design was developed. The FCSM presentation will discuss the 2017 COA non-response issues and the non-response sample design. Additionally, the 2022 COA non-response preliminary sample design will also be presented.

Subsampling to Reduce Nonresponse Bias in FoodAPS: A Simulation Study
Jeffrey Gonzalez, *National Agricultural Statistics Service*
Joseph Rodhouse, *National Agricultural Statistics Service*
Darcy Miller, *National Agricultural Statistics Service*

Subsampling is effective at combating nonresponse and reducing the potential for nonresponse bias while controlling survey costs. Diary surveys, like the National Household Food Acquisition and Purchase Survey (FoodAPS), are prone to daily nonresponse because responding each day of the data collection period may seem burdensome. This may increase the likelihood of underreporting or failing to report for some or all days during the data collection period. Given these concerns, researchers at ERS and NASS are exploring subsampling strategies to reduce nonresponse in the second round of FoodAPS. Specifically, we identify core survey items that should be asked of sample units and a subset of items that need only be collected from a subsample in order to reduce burden and decrease the likelihood of nonresponse. We present results from a simulation conducted on prior FoodAPS data that explored relationships among sample characteristics, respondent behaviors, and response tasks and considered different subsampling strategies based on these relationships. We conclude by identifying possible collection strategies to increase data quality, accuracy, and completeness in FoodAPS.

Using Process Data to Study Interviewer Effects on Measurement Error and Nonresponse in the Consumer Expenditure Survey
John Dixon, *Bureau of Labor Statistics*
Erica Yu, *Bureau of Labor Statistics*

This study investigates the relationship between survey, household, and interviewer characteristics to estimate interviewer effects in the Consumer Expenditure Interview Survey (CE). The focus of this study is on understanding the impact of interviewer effects on two data quality indicators: measurement error, approximated using item missing data, edit indicators during post-interview processing, and changes to recorded answers during the interview; and unit nonresponse bias, explored through details of contact attempt outcomes and expressions of respondent concerns, as recorded by interviewers in the Contact History Instrument (CHI). Multilevel models use indicators such as interviewer characteristics (e.g., tenure, workload) and household-level interview process characteristics (e.g., interview length, collection mode) to model measurement error and a separate model for noncontact and refusal. Findings from this study will strengthen our understanding of the role that interviewers play in nonresponse bias, with possible implications for how survey programs and field offices manage interviewer training, workload, and contact procedures.

Who's Left Out?: Nonresponse Bias Assessment for an Online Probability-based Panel Recruitment
Frances Barlas, *Ipsos*
Randall K. Thomas, *Ipsos*

Low response rates have not been found to be a useful indicator of nonresponse bias in survey estimates. However, there is still cause for concern that differential nonresponse or systematic exclusion of segments of the population are greater threats to representativeness. This paper focuses on a study we conducted using a mailed nonresponse follow-up study to investigate bias that might occur in recruitment to an online probability-based panel using address-based sample. Shortly after completing a wave of recruitment for KnowledgePanel®, we initiated the non-response survey, sending a survey to households we had successfully recruited in that wave of recruitment as well as to nonresponding households. We had responses from 866 nonrespondents (a 26.2% completion rate) and 673 responses from recruited households (a 67.4% completion rate). We examined primary demographics, secondary demographics, and several external benchmarks to compare the differences between responders and non-responders. We found that across most of the variables we evaluated, households recruited to the online panel were almost identical to those of households not recruited. Differences between recruited and nonrecruited households were rarely statistically significant. We review some of differences in terms of how recruitment can be improved to bolster the representativeness of the panel.

Session G-5
Science and Engineering Indicators: Measuring S&E Education, Workforce, and Research Output

Publications Output: U.S. Trends and International Comparisons
Karen White, *National Center for Science and Engineering Statistics*

Publication output is a measure of new scientific knowledge, diffusion, and impact of S&E research in the United States and across the globe. This presentation presents findings from the National Science Board's Science and Engineering Indicators report on Publications Output: U.S. Trends and International Comparisons (PBS). The 2020 PBS used data from Elsevier's Scopus database to highlight trends in the production, international collaboration, and impact of peer reviewed academic journals and conference proceedings. demographics of U.S. authors, and publication response to Covid-19.

Science and Engineering Indicators Academic Research and Development
Josh Trapani, *National Center for Science Engineering Statistics*

The Science and Engineering Indicators thematic report on Academic R&D contains information on sources of support for R&D performed by colleges and universities. It provides data on R&D performance across different types of academic institutions (e.g., public and private, medical schools, minority serving institutions). The report looks at funding across science and engineering fields, and also provides information on research equipment and facilities. For the first time, this report also contains data on financial support for graduate students (master's and doctoral) and postdocs. Graduate students and postdocs are vital to the academic R&D enterprise, and education and training often go hand-in-hand with R&D performance. The largest component of academic R&D direct costs are salaries, wages, and benefits and, for example, more than 80% of federally funded postdocs were paid through research grants. However, combining multiple sources of data to integrate the dollars and the people is difficult. In this presentation, I will provide highlights from the latest report. I will also discuss some of the challenges in presenting a clear and complete picture of academic R&D.

Science and Engineering Indicators: Trends in U.S. Elementary and Secondary STEM Education
Susan Rotermund, *RTI International*
Amy Burke, *National Center for Science and Engineering Statistics*

Elementary and secondary education in mathematics and science is the foundation for student entry into postsecondary STEM majors as well as a wide variety of STEM-related occupations. This presentation is based on findings from the National Science Board's Science and Engineering Indicators report on Elementary and Secondary STEM Education (K12). The report analyzes national trends in K–12 student achievement in mathematics and science and compares U.S. student performance with that of other nations. The report also includes information about secondary school mathematics and science teachers and how students' beliefs about their mathematics and science ability and identity in high school are associated with their choice of a STEM major in college. Finally, the report explores how COVID-19 affected student learning and access to educational resources.

The STEM Workforce of Today: Scientists, Engineers and Skilled Technical Workers
Abigail Okrent, *National Center for Science and Engineering Statistics*
Amy Burke, *National Center for Science and Engineering Statistics*

Individuals in the STEM workforce fuel a nation's innovative capacity through their work in research and development (R&D), and other technologically advanced activities. Studies show that lower participation among particular demographic groups in STEM signals a lack of diversity, which may negatively impact productivity, innovation and entrepreneurship. The Science and Engineering Indicators (Indicators) thematic report on the STEM workforce analyzes trends in the composition of the STEM workforce in terms of gender, race or ethnicity, nativity and citizenship of workers in STEM occupations by educational attainment (i.e., less than bachelor's degree and at least a bachelor's degree). This analysis deviates from other federal statistical agencies and from past Indicators reports by broadly interpreting the STEM workforce to include workers at all educational levels with appreciable levels of STEM skills and technical expertise. Within this broad framework and combining data from several federal statistical agencies, we find representation of demographic groups to vary by STEM sub-workforces, which are characterized by occupational group and educational attainment.

Discussant: Megan Fasules, *Micronomics*

Session H-1
Data Collection in the Post-Pandemic Era

Adapting Data Collection in a Pandemic – What Adaptations will Endure?
Barbara R. Rater, *National Agricultural Statistics Service*

USDA's National Agricultural Statistics Service (NASS) conducts over 500 state and national agricultural surveys a year and a census of the nation's 2.0 million farmers once every five years. Data users including farmers, lenders, market analysts, researchers, government agencies and others rely on timely and accurate statistics provided by NASS. Anywhere from 10-to-15 surveys are in the field on a given day, many conducted by mail, web, telephone and/or personal interview. The Covid-19 pandemic required the Agency to move away from centralized data collection systems to a decentralized operation model. 1,700 contract interviewers could no longer conduct in-person farmer surveys. Consequently, business models needed to be quickly evaluated and the impact of operating in a virtual environment had to be assessed. This presentation describes the measures taken to collect data during the pandemic and which changes and adaptations became a regular part of the organization's operation going forward.

Beyond the Pandemic: Building New Programs at BTS
Rolf Schmitt, *Bureau of Transportation Statistics*

COVID-19 inspired BTS to launch new data programs and revise existing programs to measure dramatic changes in transportation during the pandemic. Recovery of transportation activity from the pandemic creates new demands for information. BTS is building on lessons from the pandemic to find a new normal in transportation and in how to measure transportation.

COVID-19 Effects on the Health of Education Data and Implications for Future Education Data Collection
Chris Chapman, *National Center for Education Statistics*

During the first year of the COVID-19 pandemic, data collection operations at the U.S. Department of Education faced many challenges. Many of our collections required physically entering elementary and secondary schools for direct interactions with students and staff. The challenges there are obvious. Less obvious were challenges collecting data through surveys that are traditionally remote by design such as surveys of college students and educational institutions at all levels. We had already been developing new collection approaches for situations where in-person data collections might not be possible, and the pandemic is accelerating those. However, capturing information from increasingly stressed establishments remains a challenge, though we are working on new strategies on this that have yet to be tested. The presentation includes information on some of the innovations we have implemented or are designing, and background on challenges for quick collection and release of data.

Session H-2
Using Administrative Data to Examine Food Assistance Program Effectiveness

Estimating SNAP Eligibility and Access Using Linked Survey and Administrative Records
Renuka Bhaskar, *U.S. Census Bureau*
Brad Foster, *U.S. Census Bureau*
Brian Knop, *U.S. Census Bureau*
Maria Perez-Patron, *U.S. Census Bureau*

The Supplemental Nutrition Assistance Program (SNAP) is the nation's largest federal effort to reduce hunger, reaching 38 million people in fiscal year 2019. However the U.S. Department of Agriculture's Food and Nutrition Service (FNS) estimates that about one in six of those eligible for SNAP did not participate in the program. SNAP eligibility rules are complex and research has shown persistent underreporting of SNAP participation in survey data, making estimation of eligibility and access difficult. We use state administrative records on SNAP participants linked with American Community Survey (ACS) data to estimate SNAP eligibility and access rates by demographic, socioeconomic, and household characteristics at the state and county levels for selected states. We will present our methodology and demonstrate our results through data visualizations which aim to increase understanding of access to SNAP and inform future outreach.

Assessing SNAP Unit Simulations with Linked Survey and Administrative Data
Karen Cunnyngham, *Mathematica Policy Research*
John Czajka, *Mathematica Policy Research*

To provide critical information to the Food and Nutrition Service (FNS) about the performance of the Supplemental Nutrition Assistance Program (SNAP), Mathematica uses microsimulation to estimate the number of persons eligible for the program and combines these with counts of participants from SNAP administrative data to estimate SNAP participation rates. For some subgroups, the number of participants exceeds the number of simulated eligible persons, resulting in participation rates above 100 percent. To explore such anomalies, Mathematica linked SNAP administrative data from Illinois, Mississippi, and Tennessee to multiple years of data from the Current Population Survey Annual Social and Economic Supplement (CPS ASEC). To facilitate the linkage and subsequent analysis, the administrative data from each State and year were standardized. SNAP participant records were then linked to individual survey household members, and a dozen subgroups were identified among the linked individuals. Our findings regarding subgroup estimates address the role of household composition, multiple SNAP units within the same household, survey household roster omissions, and differences in the identification of subgroups between survey and administrative data.

Investigating the Factors Behind High SNAP Participation Rate Estimates using Linked SNAP Administrative Data, CPS ASEC Data, and TRIM3 Microsimulation Model Estimates
Laura Wheaton, *Urban Institute*
Nancy Wemmerus, *Decision Demographics*
Tom Godfrey, *Decision Demographics*

We investigate factors that may cause microsimulation model Supplemental Nutrition Assistance Program (SNAP) eligibility estimates to be underestimated for some subgroups, resulting in unrealistically high participation rate estimates. We use SNAP administrative record data for three states that have been linked with the CPS ASEC and with TRIM3 microsimulation model SNAP unit identifiers and eligibility flags. We find that SNAP cases simulated as ineligible by TRIM3 are much more likely than those simulated as eligible to have mismatches in administrative case and TRIM3 unit composition and to have whole ASEC imputation or income item imputation. SNAP cases with one adult and one or more children appear less likely than other types of SNAP cases to respond to the CPS interview and appear underrepresented in the final CPS ASEC, while those with multiple adults and children appear overrepresented. Finally, a substantial share of cases with one adult and children according to the administrative data have multiple adults or no children in the TRIM3 units formed using ASEC household membership and reported relationships.

Using Administrative Data to Examine Cross-program Participation in SNAP and WIC
Leslie Hodges, *Economic Research Service*
Laura Tiehen, *Economic Research Service*

Despite the benefits of participation in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), many eligible households leave the program when a participating child turns 1 year old, and participation continues to decline as children age. This research uses administrative data to examine WIC participation of SNAP households. This method avoids the problem of underreporting of participation in household surveys and allows us to identify an important group of eligible nonparticipants, since all SNAP participants are income-eligible for WIC and have demonstrated willingness to participate in food assistance programs. We use data from three states and focus on the measurement challenges that arise when linking between the two programs. For example, SNAP participants are identified in units (like households) while WIC identifies each participant separately. Also, it can be difficult to identify whether there are members of a SNAP unit (such as newborns) who belong to a demographic group that is eligible for WIC. Findings will provide information on how well WIC is reaching eligible children and provide a framework for future work using linked SNAP-WIC data.

Discussant: Constance Newman, *USDA*

Session H-3
Modernizing Data Dissemination

Moving to a More User Centered Design for Data Dissemination at the US Department of Agriculture's National Agricultural Statistics Service
Bryan Combs, *National Agricultural Statistics Service*
Jackie Ross, *National Agricultural Statistics Service*
Elvera Gleaton, *National Agricultural Statistics Service*
King Whetstone, *National Agricultural Statistics Service*

The National Agricultural Statistics Service (NASS) of the U.S. Department of Agriculture is moving toward a more modern data dissemination model. The first phase of a new project called Data Repository and User Interface Design (DRUID) produced an initial framework and prototype to demonstrate a solution for an improved data repository that included user-centric API driven methods to access the data and information in published NASS releases. This prototype used a layered data architecture and new database design approach within USDA's Enterprise Data Analytics Platform Toolset. Human-Centered Design methods were used to develop and showcase a functional user interface and visualization of trending historical data, geospatial data, and other commodity-specific views. The vision for DRUID is to modernize the existing NASS Quick Stats database by designing and developing a more robust, visually appealing, and user-friendly solution. This new solution will provide a well-designed user experience for data collection, data dissemination, and interactive analysis capability.

Modernizing Data Dissemination at the U.S. Bureau of Labor Statistics
Clayton Waring, *Bureau of Labor Statistics*

As part of the Bureau of Labor Statistics (BLS) ongoing efforts to replace its current LABSTAT data dissemination tools, it has adopted what is referred to as a definitional approach for storing and organizing statistical data. Under the new BLS approach, statistical data is stored and organized around these definitional elements of measures; people, places and things; and time. One of the main advantages of this approach, besides its simplicity, is that it is specifically designed to be both scalable and flexible. The basic method for data location and retrieval is based on the notion of three query entry points in order to minimize the paths to the data. Any statistical value in the system may be accessed as the intersection of these three query paths. These query paths also explain how all the data in the system is interconnected. Building the new data dissemination system will likely be challenging, however, at each step the goal should be to eliminate as much complexity as possible from the overall system. This begins with discovering easy and intuitive concepts and principles for storing and organizing statistical data.

Developing an Enterprise-wide Dissemination Program at the U.S. Census Bureau
Zachary Whitman, *U.S. Census Bureau*

The U.S. Census Bureau is undertaking an enterprise-wide consolidation of its data dissemination services through the instantiation of the Center for Enterprise Dissemination Services and Consumer Innovation (CEDSCI) program. This program is developing a consolidated data platform, a suite of dissemination open API services, and several web-based applications that streamline the public discovery, access, and consumption of Census Bureau data products. It is also responsible for developing a centralized governance and management model that provides standards and goals for the enterprise's dissemination activities, products, and metadata. These governance activities are aimed to build the metadata standards required to improve data interoperability and improved data product design based on user needs. The program is structured so that the collection and analysis of user feedback serve as a core component in prioritizing development and governance activities for the program as well as providing these insights back to the data providers. Ultimately, this program will serve as a key advocate for the data user community by continually improving its data dissemination services and products for broader consumption.

Discussant: Cynthia Parr, *USDA*

Session H-5
Assessing and Improving the Quality of Prescription Drug Data from Surveys

Improving Self-Reported Prescription Medicine Data Quality with a Commercial Database Lookup Tool and Claims Matching

Kali Defever, *NORC at the University of Chicago*
Becky Reimer, *NORC at the University of Chicago*
Michael Trierweiler, *NORC at the University of Chicago*
Elise Comperchio, *NORC at the University of Chicago*

The Medicare Current Beneficiary Survey (MCBS) is a continuous survey of a nationally representative sample of the Medicare population, conducted by the Centers for Medicare & Medicaid Services (CMS) through a contract with NORC at the University of Chicago. This survey collects data from beneficiaries about their prescription medicine use, including medicine name, strength, form, and quantity. In 2017, the prescription medicine lookup tool in the MCBS questionnaire was revised to integrate a high-quality commercial medicine name database. Medicines reported in the survey are linked to CMS administrative prescription medicine claims during data processing to create a complete dataset. In this study, we evaluate the proportions of reported medicines that are matched to the commercial database and that include medicine quantity during data collection, and how these metrics are improved after claims matching. To assess whether regularly prescribed medications are recalled more readily than other medicines, we compare commercial database match rates and claims match rates for widely prescribed medicines vs. medicines overall.

Exploring Potential Benefits of Enumerating All Prescribed Medicines as a Tool for Estimating Opioid Use in the Medicare Current Beneficiary Survey (MCBS)

Becky Reimer, *NORC at the University of Chicago*
Elise Comperchio, *NORC at the University of Chicago*
Andrea Mayfield, *NORC at the University of Chicago*
Jennifer Titus, *NORC at the University of Chicago*

Researchers seeking to obtain accurate population estimates of opioid use via self-reported measures often focus on ideal wording and administration of opioid use questions. This study assesses the feasibility of an alternate approach: estimating opioid usage by enumerating all prescribed medicines and determining which, if any, are opioids in data processing rather than data collection. We use data collected from the 2018 and 2019 Medicare Current Beneficiary Survey (MCBS), a nationally representative survey of the Medicare population, conducted by the Centers for Medicare & Medicaid Services (CMS) through a contract with NORC at the University of Chicago. The MCBS collects details about prescription medicines filled by beneficiaries. We link these data to an administrative list of opioids to categorize medicines as opioids vs. non-opioids. We estimate the proportion of beneficiaries who filled prescriptions for at least one opioid within one year's time in 2018/2019 and compare reported opioid usage based on beneficiary characteristics. We discuss potential benefits and challenges of enumerating all prescription medicines as an approach for estimating opioid usage.

Evaluating Alternative Benchmarks to Improve Identification of Outlier Drug Prices for MEPS Prescribed Medicines Data Editing

Yao Ding, *Agency for Healthcare Research and Quality*
Steven C. Hill, *Agency for Healthcare Research and Quality*

The Medical Expenditure Panel Survey (MEPS) is a nationally representative household survey supplemented with data collected from pharmacies and other providers that serve the sample members. The MEPS Prescribed Medicines (PMED) file provides detailed information on drug expenditures. As the MEPS is widely used for descriptive, behavioral, and simulation analyses to inform health care policy, we evaluate potential improvements to PMED data editing. A quality control check in the PMED editing process compares pharmacy-reported unit prices with widely available benchmark unit prices. This study first compares MEPS data with private claims from IBM MarketScan Databases in terms of unit price (price per pill) relative to the benchmark unit price. Second, we compare unit prices relative to alternative benchmark prices that became available since the quality check was last refined. Third, we compare prices per fill or refill between the two datasets to assess the potential importance of implementing additional quality checks. The comparisons are stratified by the patent status of the drug (generic, originator, or single source brand name) and separately for biologics and orphan drugs.

Discussant: Geoffrey Paulin, *Bureau of Labor Statistics*

Session I-1
Survey Design and Measurement

Blind to the Consequences of Measurement?: Response Format Effects on Self-Reported Disability
Megan A. Hendrich, *Ipsos Public Affairs*
Randall K. Thomas, *Ipsos Public Affairs*
Frances M. Barlas, *Ipsos Public Affairs*

Many governmental programs require accurate estimates of the prevalence of disabilities for planning and funding purposes. To assess prevalence, we often present disabilities as a series of items and ask respondents if each is true for them. The response format is typically a yes-no dichotomous format (DF) or a multiple response format (MRF; select all that apply). Smyth et al. (2006) and Thomas and Klein (2006) found higher endorsement rates using a DF compared to a MRF. Thomas and Klein (2006) hypothesized that the DF requires more cognitive effort to respond, leading to the recall of less salient events than the MRF. To test the salience hypothesis, this study used two experiments comparing a yes-no DF with a MRF to assess the prevalence of four types of disability. In a web-based survey, respondents were randomly assigned to either the DF or MRF. For each item endorsed, we asked about the severity of their disability, expecting more severe disabilities to be more salient. Replicating earlier work, those assigned the DF yielded a higher prevalence of disability than those assigned the MRF. However, in the follow-up question, average disability severity was lower with the DF than the MRF. This supports the hypothesis that less salient responses are less likely to be selected with a MRF compared to a traditional yes-no DF. As a result, MRF may underestimate the prevalence of disabilities.

Evaluating the National Agricultural Statistics Service's Grain Stocks Program: Results from Behavior Coding
Ashley Thompson, *National Agricultural Statistics Service*
Heather Ridolfo, *National Agricultural Statistics Service*

Engaging with respondents and the farmers who produce agricultural commodities is an important component to completing the many surveys the National Agricultural Statistical Services (NASS) administers yearly. Quarterly, NASS conducts the Crops Acreage and Production Survey (Crops APS) for all states. Respondents provide information on crop acreage estimates, yields and production, and total quantities of grain and oilseeds that are stored on farms. To address several possible sources of survey error, NASS is currently performing a technical review of the Grain Stocks program. As part of this review, behavior coding was conducted on Crops APS Computer Assisted Telephone Interviews (CATI). This research presents the results of the behavior coding of the Crops APS survey and an evaluation of the findings. The evaluation will focus on the crop storage and total storage capacity of the Crops APS survey and the many challenges associated with storage capacity. Storage of commodities for farmers has changed dramatically over the past 10 years to include new technology for quick storage methods that need to be accounted for in NASS's survey questionnaire and methods. The findings will highlight possible improvements that can be made to the storage capacity section of the Crops APS Survey, ultimately improving NASS's published estimates for the data users.

Examining Proxy Response Bias in a Large-Scale Survey of People with Disabilities
Eric Grau, *Mathematica, Inc.*
Jason Markesich, *Mathematica, Inc.*

Surveys that collect health-related data from people with disabilities often allow proxy respondents to answer questions for sample members who, because of their disabilities, cannot respond for themselves. The rationale for using proxy respondents is to minimize the potential for bias associated with nonresponse. But bias could still be present if the responses provided by the proxies are systematically different from those provided by self-respondents with disabilities. In this study, we will examine this bias, called proxy response bias, in a large-scale survey of people with disabilities—the National Beneficiary Survey (NBS). Our study has three objectives: (1) to determine if the use of proxy respondents in the NBS is related to the demographic characteristics of the sample members; (2) to examine the size and direction of the differences between proxy and self-reported responses to questions on health-related questions; and (3) to assess whether the proxy-sample member relationship affects the differences between proxy and self-reported responses.

Integrating Previously Reported Data into The Census of Agriculture
Gavin Corral, *National Agricultural Statistics Service*
Greg Lemmons, *National Agricultural Statistics Service*
Linda J. Young, *National Agricultural Statistics Service*

The USDA National Agricultural Statistics Service (NASS) is incorporating the use of Previously Reported Data (PRD) into its Census and survey programs. One major goal is to have PRD integrated into its 2022 Census of Agriculture. In preparation for the Census, several tests of the potential impact of PRD on data quality and the PRD delivery system being developed are being conducted. In this paper, the results from studies embedded in (1) the 2020 Census of Agriculture Content Test and (2) the 2020 September Crops Agricultural Production Survey are discussed. Furthermore, the lessons learned from the first operational use of the PRD delivery system, which will be for the June APS Survey, are presented. Areas of future research are considered.

Session I-3
Price Indices and Consumer Demand

Consumer Prices During a Stay-In-Place Policy: Theoretical Inflation for Unavailable Products
Rachel Soloveichik, *Bureau of Economic Analysis*

Major product categories like in-person restaurant meals, live entertainment, and nonessential personal services are unavailable during a stay-in-place policy. As a result, their inflation rates cannot be measured directly. This paper uses previous research on the value of tourist amenities (Florida 2018a) and a newly developed model of tourist behavior to calculate a theoretical price for unavailable products. In this paper, the word "theoretical" designates an imputed price which is consistent with price measurement theory (Diewert and Fox 2020) (Diewert et al. 2019) (Diewert 2003). It does not imply any data problems or computational mistakes with either the consumer price index (CPI) published by the Bureau of Labor Statistics (BLS 2018) or with any other cost-of-living indexes. This analysis estimates monthly product unavailability in every region of the United States. Based on those regional estimates, the paper calculates that the average U.S. consumer experienced theoretical inflation at least 1.4 percentage points higher than the published CPI in the first quarter of 2020, at least 6.0 percentage points higher than the published CPI in the second quarter, and at least 2.8 percentage points lower than the published CPI in the third quarter. The faster inflation rates in the first two quarters of 2020 reinforces the measured declines in real consumption during those quarters and the slower inflation rate in the third quarter of 2020 reinforces the measured recovery in real consumption during the third quarter. In other words, at least one third of the theoretical decline and recovery in real consumption is not reflected in published economic statistics.

Democratic Aggregation: Issues and Implications for Consumer Price Indexes
Robert Martin, *Bureau of Labor Statistics*

The Bureau of Labor Statistics' current Consumer Price Index (CPI) methodology is so-called "plutocratic," implicitly weighting households by their total expenditure. If inflation varies systematically with expenditure, then the CPI may differ from an aggregate that is equally-weighted across households, so-called "democratic." Equally-weighted aggregates are sometimes considered to better summarize household inflation experiences. I calculate democratic counterparts to the BLS's CPI and Chained CPI products in order to estimate the impact of aggregation method. The building blocks are household-level Lowe and Tornqvist indexes I construct using matched Consumer Expenditure Survey (CE) diary and interview microdata, along with CPI elementary indexes. I estimate that over 2002-2018, democratic aggregation would have increased the CPI-U and C-CPI-U, on average, by 0.14 and 0.27 percentage points per year, respectively. I also document how aggregation impacts differ by time horizon and subpopulation.

Session I-4
A Discussion of the Final Report of the Interagency Technical Working Group on Evaluating Alternative Measures of Poverty

After more than two years of work, the Interagency Technical Working Group on Evaluating Alternative Measures of Poverty issued its final consensus report in January 2021. The group was convened by then Chief Statistician of the United States, Nancy Potok, and included members of 11 federal agencies.
The final report recommends that the federal government produce an income poverty measure blending administrative and survey data and a consumption poverty measure. These new poverty measures would not replace existing poverty measures including the Official Poverty Measure and the Supplemental Poverty Measure, but would be in addition to them.

For an extended income resource measure, the Working Group recommends expanding beyond pre-tax cash income to include at least some in-kind transfers and accounting for taxes and tax credits, much like the SPM resource measure. In addition, the Working Group recommends producing resource measures with and without a value for health insurance with some direction provided on how to do so. The Working Group discussed how to value implicit flows from non-financial assets (e.g., vehicles, owner occupied housing, and other properties), and how to value net flows from financial assets for inclusion in both income- and consumption-based resources. An extended income resource measure would also integrate administrative data with household survey income information when appropriate, taking advantage of recent research on the use and the increased availability of potentially more accurate administrative data. The Working Group considered other approaches for adjusting survey data for misreporting as well.

A consumption-based resource measure may more directly capture the resources available to a family if it records the consumption that was actually achieved. These measures begin by summing most categories of expenditures on goods and services. The Working Group recommends beginning in this way, but recommends excluding certain categories of expenditures that are often thought of as enhancing future consumption, such as pension contributions and education expenses. As with the extended income-based resource measure, the Working Group recommends that any consumption-based resource measure be produced with and without a value of health insurance. The Working Group also recommends that a flow of consumption resources be attributed to owned vehicles and owner-occupied homes and that current expenditures on these items be excluded from consumption.

The Working Group also identifies other areas worthy of future research by the Federal Statistical System.

This session will include three working group members – Bruce Meyer, Thesia Garner and Liana Fox. Commentary on the report and its recommendations will be provided by two experts on poverty measurement – Gary Burtless and Laura Wheaton.

Session I-5
Machine Learning Applications

Machine Learning Assisted Complex Survey Weights
Stas Kolenikov, *Abt Associates*

We propose a workflow to create complex survey, nonresponse adjusted and calibrated weights that utilize machine learning methods (1) to support estimation of response propensities, and (2) to augment the population control totals for the model calibration step with outcome propensities. In a "traditional" workflow, the sampling statistician would fit a logistic regression of the unit response indicator on the main effects of the frame/baseline variables and calibrate the weights to the population characteristics and/or the frame/baseline demographic variables. We replace the nonresponse adjustment step with a more flexible estimation of response propensities using machine learning methods, and we also use the predictions from the machine learning models for the primary outcomes of the survey to create synthetic variables whose control totals can be used for weight calibration. Improvements are found in both the accuracy of the nonresponse adjustments, and in calibration to the augmented variables. The process is demonstrated with a list sample of participants in a training program where rich baseline data provide the input features for the machine learning training models.

Using Machine-Learning Algorithms to Improve Imputation in the Medical Expenditure Panel Survey
Chandler McClellan, *Agency for Healthcare Research and Quality*
Emily Mitchell, *Agency for Healthcare Research and Quality*
Jerrod Anderson, *Agency for Healthcare Research and Quality*
Samuel Zuvekas, *Agency for Healthcare Research and Quality*

Survey data collection often results in incomplete or missing responses. Imputation methods are widely used to complete survey data, and improvements to those methods are an active area of research. We study the feasibility and performance of applying machine learning methods to imputation of expenditures in the Medical Expenditure Panel Survey (MEPS). Expenditure imputation in the MEPS is critically important as expenditures are both the primary purpose of the survey and have high rates of missing data. Currently, MEPS expenditure data are imputed with a predictive mean matching (PMM) algorithm in which a linear regression model is used to predict expenditures for recipients and donors. Recipients and donors are then matched based on the smallest distance between predicted expenditures. We assess whether machine-learning (ML) algorithms can replace linear regression in the PMM framework to predict expenditures and generate superior matches. We find that ML algorithms offer substantial improvements in both expenditure prediction and imputation matching performance. Better imputations have implications beyond just improving the MEPS expenditure data. A demonstrable improvement in predictive accuracy may be of value in other national surveys that currently rely on methods similar to PMM for imputation.

An Overview of Business Establishment Automated Classification of NAICS (BEACON) for the Economic Census
Daniel Whitehead, *U.S. Census Bureau*
Brian Dumbacher, *U.S. Census Bureau*

BEACON is a machine learning tool developed to assist Economic Census respondents in the selection of their establishment's North American Industry Classification System (NAICS) code. BEACON uses the text provided by the respondent, in real time, to estimate the respondent's most likely NAICS code. BEACON utilizes past Economic Census responses in conjunction with Internal Revenue Service data and NAICS manual descriptions to create a data dictionary for training and testing purposes. Through an ensemble method, BEACON hierarchically predicts a respondent's NAICS code, first at the 2-digit level and then at the 6-digit level. This report details the modeling methods, disclosure avoidance procedures, and evaluation metrics used to create a production-ready version of BEACON.

Supplementing Cognitive Interviewing with Natural Language Processing Approaches from Data Science

Katherine Blackburn, *RTI International*
Peter Baumgartner, *RTI International*
Stephanie Eckman, *RTI International*
David Henderson, *RTI International*
Y. Patrick Hsieh, *RTI International*
Patricia Green, *RTI International*
Kelly Kang, *National Center for Science and Engineering Statistics*

Cognitive interviews are a valuable tool to improve questionnaires, but they are time and labor intensive. This research explores whether natural language processing techniques can generate supplementary insights about question performance. In 2020, the NCSES Survey of Earned Doctorates fielded 6 new closed-ended questions to collect data about the impact of the Coronavirus pandemic on students. Because the situation was novel, open-ended follow-up questions were used to obtain more information, which is rare for a survey with about 55,000 respondents per year. To identify groups of thematically similar responses, we used natural language processing and cluster analysis to support decisions about additional questions to be added to the survey in the following cycle. Simultaneously, methodologists developed closed-ended questions from the responses, which were then evaluated with cognitive interviews. Similar themes emerged from the cluster analysis and the cognitive interviews. These results suggest that data science techniques could supplement cognitive interviews for questionnaire testing. This presentation will be of interest to researchers who are developing new survey items.

For What It's Worth: Measuring Land Value in the Era of Big Data and Machine Learning

Scott Wentland, *Bureau of Economic Analysis*
Gary Cornwall, *Bureau of Economic Analysis*
Jeremy Moulton, *University of North Carolina*

We develop methods to estimate the value of privately owned land the continental United States using a novel machine learning approach paired with "big data" from Zillow. Because this data includes detailed information from hundreds of millions of property transactions covering much of the US, the heterogeneous nature of this data serves as fertile ground for highlighting some of the practical limitations of linear hedonic regression techniques in the context of "big data." We first construct hedonic estimates of land value at the parcel-level for most of the US in order to build up from small geographies to aggregate national values. For comparison, we then modify the hedonic method using a machine learning approach, generating new land value estimates using a random forest procedure. We show how a machine learning approach can systematically address issues of spatial controls and thin cells at smaller levels of geography and population density (with fewer corresponding market transactions), along with addressing other shortcomings associated with linear hedonic regression approaches to valuation. Our initial estimates using the traditional hedonic approach show land values fell about $10 trillion (or 28%) from the boom to bust periods in the 2000s and experienced a nearly full recovery by the middle of the second decade. These proof-of-concept estimates show that private land in the contiguous 48 states to be worth approximately $35.5 trillion in 2016.

Session J-1
Modernizing Federal Survey Data Collections through Modularized Design

NCSES' BAA Program and Emphasis on Modular Design Research
Jennifer Sinibaldi, *National Center for Science and Engineering Statistics*

In 2020, NCSES established several research partnerships focused on specific needs of the Center through a new contract mechanism called a Broad Agency Announcement (BAA). A couple of the research areas for which NCSES was seeking collaboration were focused on modernizing the delivery of NCSES' surveys of the Science and Engineering workforce in ways that would streamline data collection and reduce respondent burden. One method to achieve this goal is through modular questionnaire design – thoughtfully dividing the questionnaire into segments and distributing the administration of those segments across the sample. "Thoughtfully dividing" into subsets of questions and determining the delivery of those questions can involve a few different methods. This introductory presentation will provide an overview of the BAA program and NCSES' goals for this research area, providing background for the rest of the speakers in this session.

Investigating Modular Designs for the Survey of Doctorate Recipients
Ai Rene Ong, *University of Michigan*

To investigate the potential for modular design in the Survey of Doctorate Recipients (SDR), we evaluated which questions were most important to the SDR user community. We used three main sources to represent the user community and assess the usage and importance of SDR survey items: 1) Congressional mandates to the sponsoring agency (the National Center for Science and Engineering Statistics) on reporting, 2) a literature review of academic research using the SDR, and 3) frequency counts of data table requests made via the SDR website. Our research team coded the SDR bibliography for SDR variable usage and citation count. Putting this information together, we were able to identify a number of items that were not used by any of these sources or cited infrequently. These items could be dropped, collapsed, or used in modular design where the question is not asked of everyone or asked every year. We suggest that this methodology could be applied more generally to make decisions about which questions to ask on federal surveys.

Developing and Evaluating Methodology for Split Questionnaire Design in the National Survey of College Graduates
Andy Peytchev, *RTI International*

This research evaluated the potential for NCSES to reduce the delivery time of the National Survey of College Graduates (NSCG) but still generate a complete dataset for all respondents. The project is two-fold, involving input from questionnaire design experts and survey statisticians. First, the team divided the questionnaire into essential questions ("core") that must be asked of everyone and those that could be modularized. Modules were optimized considering both necessary question context and essential correlations necessary to impute the missing data for the modules not asked of each respondent. Following the questionnaire design, a simulation study assessed the feasibility of the design and the quality of the final data.

Co-Designing a Smartphone App with Target Audience Members
Chris Antoun, *University of Maryland*

We conducted three participatory design (PD) workshops with individuals who match our target audience for the Survey of Doctorate Recipients (SDR) conducted by NCSES. The objective was to co-design a potential data collection app that could administer short modularized surveys, as opposed to one long instrument. In each workshop, participants were asked to identify the design features or functionality that would be most important to them, develop a plan for encouraging respondents to download and use an app, discuss their ideas about the optimal length and timing of survey modules, and work in small groups (via Zoom's "Breakout Rooms") to design their own prototypes of an app and then have them evaluated by other participants. At the conclusion of the workshops, our team will propose a modular design with several experimental components for delivery of the SDR via smartphone app. The design will include decisions on: optimal module length and question types appropriate for this mode of data collection, recruitment and retention contact protocols, and frequency of module delivery.

Session J-2
Perseverance and Resilience: The Medical Expenditure Panel Survey in the Pandemic

Managing Survey Change during the Pandemic
Brad Edwards, *Westat*
Rick Dulaney, *Westat*

Successful continuing surveys may ask the same questions year after year to produce data on trends and trajectories. But how do you manage dramatic survey changes without disrupting trend lines in the face of earth-shaking events like the coronavirus pandemic? Using the Medical Expenditure Panel Survey (MEPS) as a case study, we review the challenges faced in 2020 to keep the project afloat with continuous data collection and an unbroken data series on American's use of health care its cost. Solutions included switching from in-person to telephone interviewing, introducing a web mode, extending the longitudinal data collection from 2 years to 4, and exploring more resilient study designs. Switching modes and introducing design and technology changes while managing resources, maintaining operations continuity, and measuring changes in behavior in a standardized way can be challenging to say the least. We spotlight design, schedule, and cost tradeoffs, show how we used paradata to inform results of changes, and illustrate how we monitored effects on an array of key survey statistics such as number of medical provider visits and health insurance coverage. The effort is justified by the critical importance of measuring the enormous changes in health care from pre-pandemic, through the past year, and forward to post-pandemic.

Assessment of the Effects of COVID-19 on Data Collected by the Medical Expenditure Panel Survey
Alisha Creel, *Westat*
Ralph DiGaetano, *Westat*
Hanyu Sun, *Westat*
Alexis Kokoska, *Westat*
David Cantor, *Westat*

The Medical Expenditure Panel Survey (MEPS) is the nation's primary source of medical expenditures, utilization and insurance coverage. On March 17, 2020 the MEPS switched from face-to-face to telephone interviewing because of COVID-19. This presentation describes analyses assessing the effect of this switch on the data collected by the survey. One set of analyses examines the effect on survey processes that affect data quality, including the use of records during the interview, respondent use of show cards and interviewer-respondent interaction. The second set of analyses assesses whether the change in methodology affected key survey estimates, such as utilization and health insurance status. To separate out the changes in methodology from real change in health care outcomes, the analyses take advantage of the longitudinal design of the MEPS by developing comparison groups that isolate the effects of the methodology. For survey process measures it is possible to compare before and after the change in methodology occurred. For key survey estimates, the analyses compare interviews that occurred after the switch but reference data prior to the COVID period with comparable groups in prior years.

From one, many: Hatching a Multimode, Multiple-respondent Supplement via a Household Interview
Darby Steiger, *Westat*
Angie Kistler, *Westat*
Sandra Decker, *Agency for Healthcare Research and Quality*

The Agency for Healthcare Research and Quality's (AHRQ's) Medical Expenditure Panel Survey (MEPS) provides nationally-representative data on health care expenditures, health care usage, and household characteristics. For the Household Component (HC), Westat interviews annual panels of about 10,000 households 5 times over 2.5 years. MEPS-HC data are collected (via CAPI or CATI) with one individual on behalf of the household, as well as a yearly SAQ for adult householders. Recognizing that social and economic conditions of a household are important predictors of health status and health care usage, AHRQ sought to add another SAQ to measure social and behavioral determinants of health (SDOH). In 2020-21, Westat designed and launched this SAQ as a multimode instrument (web and paper) that aimed to protect privacy due to the sensitive nature of some of the items. The MEPS SDOH SAQ is currently in the field. This paper will review the policy goals behind the SDOH, the protocol that was designed to encourage web response, data collection challenges during the coronavirus pandemic, and our experiences with both modes of data collection, including overall and item non-response rates.

Interviewer Training for Classroom versus Distance Learning: Initial Skill Gains and Measures of Drift
Hanyu Sun, *Westat*
Angie Kistler, *Westat*
Ryan Hubbard, *Westat*
Brad Edwards, *Westat*
Marcia Swinson-Vick, *Westat*

There is abundant literature about interviewer effects on the survey process but studies of interviewer training are quite limited (e.g. Groves & McGonagle 2001). We conducted two training experiments with a group of 250 experienced field interviewers working on the Medical Expenditure Panel Survey's Household Component (MEPS-HC) to assess training effectiveness. (1)The interviewers were stratified by performance on gaining cooperation and data quality. We selected 40% of the interviewers at random to attend a 2.5 day "refresher" training in a classroom setting. Peer learning was an important feature of the curriculum. We compare the performance of each interviewer before and after training, and compare the treatment group's performance with the remaining interviewers who serve as a control group. (2)We selected two classroom modules and developed two distance learning: one for delivery in a virtual classroom setting, and the other for delivery in PowerPoint slides. Interviewers who were not invited to the classroom training were assigned to one of the distance learning modes. We compare performance post-training on specific skills to explore the effectiveness of distance learning.

Discussant: Joel Cohen, *Agency for Healthcare Research and Quality*

Session J-3
After the Fact – Data Revision, Imputation, and Analysis Methods

A Kriging Approach for Representing Crop Progress and Condition at Small Domains
Arthur Rosales, *National Agricultural Statistics Service*

The USDA National Agricultural Statistics Service (NASS) provides crop progress and condition estimates for select crops on a weekly basis during the crop-specific growing seasons. Data are collected from respondents at the county level, but it is aggregated to the state-level to follow disclosure guidelines for protecting privacy, as well as to increase stability of estimates. However, this aggregation results in a loss of information about spatial trends at the county level. In recent years there has been a push for NASS estimates to be provided for smaller spatial domains, which coincides with the increased availability of high-resolution geospatial datasets containing information about vegetation, soil, weather, and climate. In this paper novel NASS datasets that provide estimates of crop progress and condition at the sub-state level in a geospatial, gridded format are presented. These fully synthetic datasets were created using a geostatistical approach called kriging, which relies on spatial autocorrelation to fit a model that is then applied to the existing data to create a prediction surface. Additionally, this paper explores the reliability of county estimates derived from these new gridded datasets and draws comparisons to existing state-level crop progress and condition data.

Coming Clean: Does Data Cleaning Reduce or Increase Bias in Sub-groups?
Randall K. Thomas, *Ipsos Public Affairs*
Frances M. Barlas, *Ipsos Public Affairs*
Megan Hendrich, *Ipsos Public Affairs*

To improve accuracy of results, many researchers exclude cases from analyses when participants have demonstrated sub-optimal or inattentive behaviors, such as speeding or straightlining. We have seen some go so far as to clean out up to 20 percent of participants. This raises questions about the validity of the survey results, but also has cost implications as replacement sample is often required, and those being cleaned are often less likely to respond in the first place (e.g., younger, people of color, etc.). We analyzed the impact of data cleaning in a web-based study that had over 3,300 participants from a probability-based sample and 5,800 from non-probability online samples. We explored the effects of data cleaning at different rates (from 2.5% to 50%) for effects on average error from 24 different benchmarks, both overall and for specific subgroups (e.g., male or female; Black, Hispanic, or White). While we found that average error for probability-based sample was lower than opt-in sample, we also found that that cleaning does not reduce error at an overall level, nor does it increase error. At the subgroup level, there is some evidence that error is somewhat higher for some groups. Looking at covariance measures, we found that excessive cleaning protocols may actually increase bias, especially for subgroups.

Growing a Modern Edit and Imputation System
Darcy Miller, *National Agricultural Statistics Service*
Vito Wagner, *National Agricultural Statistics Service*
Travis Smith, *National Agricultural Statistics Service*

The USDA National Agricultural Statistics Service (NASS) conducts hundreds of surveys each year and the Census of Agriculture (COA) every five years (in years ending in 2 and 7) and prepares reports covering virtually every facet of U.S. agriculture - production and supplies of food and fiber, prices paid and received by farmers, farm labor and wages. A few survey programs with large surveys and the COA have an automated editing and imputation processing system. However, for most small and medium-sized surveys conducted by NASS, the corrections of edit failures and item imputations include a manual, interactive process. In some cases, especially for large producers, the adjustments for nonresponse are manual, unit imputations based on previously reported data. In 2016, NASS worked with Westat to conduct a full review of its editing and imputation processes, and a full report was delivered in 2017. Over the past 3 years, NASS has taken steps to modernize and generalize its processes to include automation as well as improved imputation methodology to increase overall efficiency and data quality. In this paper, we discuss the technical, administrative, and cultural growth moving NASS toward a more modern editing and imputation system.

How Large are Long-run Revisions to U.S. Labor Productivity?
Peter B. Meyer, *Bureau of Labor Statistics*
John Glaser, *Bureau of Labor Statistics*
Kendra Asher, *Bureau of Labor Statistics*
Jay Stewart, *Bureau of Labor Statistics*
Jerin Varghese, *Bureau of Labor Statistics*

This paper examines revisions to official estimates of quarterly US productivity growth for 1994-2015 and to the underlying data on output and hours worked. These figures are revised substantially in the first months after the reference quarter. The magnitudes of revisions decline to near zero within five years. Revisions come from additional microdata, benchmarking, seasonal adjustment, and changes to definitions and procedures. Revisions to output are larger than revisions to labor. Revisions are larger for first quarters and recessions. Later revisions are approximately normally distributed but early ones are not. Periodic comprehensive revisions to GDP slightly raise measured productivity. Revisions are not very predictable. We test for trends over time in the size of revisions and examine the outliers in the data series.

Session J-5
The Reporting, Analysis, and Mitigation of Nonresponse Bias in Federal Surveys

A Systematic Review of Nonresponse Bias Studies in Federally Sponsored Surveys
Tala H. Fakhouri, *Food & Drug Administration*
Jennifer Madans, *National Center for Health Statistics*
Peter Miller, *U.S. Census Bureau*
Morgan Earp, *National Center for Health Statistics*
Kathryn Downey Piscopo, *Substance Abuse and Mental Health Services Administration*
Steven M. Frenk, *National Center for Health Statistics*
Elise Christopher, *National Center for Education Statistics*

In 2006, the Office of Management and Budget (OMB) published Standards and Guidelines for Statistical Surveys requiring that all Federal surveys with a unit response rate of less than 80 percent conduct an analysis of nonresponse bias (NRB). Since 2006, Federal surveys have increased activities involving analyses of nonresponse bias; however, the quality and style of the reports have greatly varied in terms of methodology, rigor, and reporting. The FCSM chartered the Nonresponse Bias Subcommittee to help assess the overall impact of OMB's Standards and Guidelines for Statistical Surveys. The NRB subcommittee has finalized a report that summarizes more than 160 NRB studies and describes the general characteristics of nonresponse bias studies, including agency sponsorship, response rates, type of survey, and mode of data collection; the types of NRB assessment method(s) used; the target of the NRB analyses (i.e., sample composition, survey estimates, or both); and whether post-survey nonresponse adjustments appeared to reduce bias in final estimates.

FCSM Best Practices for Nonresponse Bias Reporting
Morgan Earp, *National Center for Health Statistics*
Jennifer Madans, *National Center for Health Statistics*
Elise Christopher, *National Center for Education Statistics*
Jenny Thompson, *U.S. Census Bureau*
Tala Fakhouri, *Food and Drug Administration*
Robert Sivinski, *Office of Management and Budget*
Kathryn Downey Piscopo, *Substance Abuse and Mental Health Services Administration*
Joseph Schafer, *U.S. Census Bureau*
Stephen Blumberg, *National Center for Health Statistics*

Since the US Office of Management and Budget published the *Standards and Guidelines for Statistical Surveys* (2006), requiring all Federal surveys with a unit response rate of less than 80 percent to conduct an analysis of nonresponse bias, US Federal agencies, contractors, and data users have written almost 200 reports on assessments of nonresponse bias. The analytical and reporting practices for NRB have varied widely within and across agencies, as documented in the previous presentation of the FCSM NRB Subcommittee 2020 report, "A Systematic Review of Nonresponse Bias Studies in Federally Sponsored Surveys". In an effort to address nonresponse bias reporting inconsistencies, the FCSM Nonresponse Bias Subcommittee drafted a report of best practices and guidelines for reporting on evaluations of nonresponse bias. This presentation will provide a high-level overview of the best practices and guidelines outlined in the report.

Nonresponse Bias Analysis Methods: A Taxonomy and Summary
James Wagner, *University of Michigan*

There are several methods available for conducting nonresponse bias analyses. These include comparing responders with nonresponders on known characteristics, comparing early and late respondents, comparing respondents to known characteristics of the population, and evaluating the impact of nonresponse weighting adjustments on estimates. This presentation will focus on how these broad classes of methods can be used as well as describing recent innovations. Each of these approaches provides a different viewpoint from which to assess the risk and likely impact of nonresponse. The presentation suggests that a combination of methods are likely to be the best approach. Further, within each method, sensitivity analyses should be conducted so that best- and worst-case scenarios can be outlined.

Nonresponse Bias Mitigation Strategies
Andy Peytchev, *RTI International*

Declining response rates have increased the threat of nonresponse bias. There are different strategies to mitigate this risk, which can be classified into three broad groups: (1) overall study design features, such as modes of data collection, (2) strategies during data collection, such as responsive and adaptive survey design, and (3) postsurvey processing and estimation, including the weighting methodology. This presentation will provide an overview of these strategies, with examples. Particular attention will be devoted to more recent developments in the field.

Discussant: Robert Sivinski, *Office of Management and Budget*

Session K-1
Previously Reported Data and Dependent Interviewing in Official Statistics: Existing Practices and New Findings

Displaying Previously Reported Data to Respondents in the Quarterly Census of Employment and Wages Program at the Bureau of Labor Statistics
Emily Thomas, *Bureau of Labor Statistics*

The Quarterly Census of Employment and Wages (QCEW) is a federal/state cooperative that publishes monthly employment and quarterly wages covering 98 percent of U.S. jobs, available at the county, MSA, state and national levels by industry. The Multiple Worksite Report (MWR) asks multi-location employers for employment and wage data for all establishments covered under one Unemployment Insurance (UI) account. In the MWR Web system, respondents are shown the employment and wage data that they submitted in the prior quarter alongside the current quarter data entry area. This assists respondents with what they should be reporting, levels reported previously, and ensures that respondents report data for the correct establishments. The Annual Refiling Survey (ARS) ask employers to verify or update geography and industry on a three-year cycle. In the ARS Web system, respondents view the address and industry currently on file for their company. The respondent either verifies that the displayed information is accurate or provides an update. Respondents have an industry baseline to verify, rather than newly selecting a new industry. This reduces churn and respondent burden.

Testing Dependent Interviewing on a Self-Administered Survey
Jennifer Sinibaldi, *National Center for Science and Engineering Statistics*

The Survey of Doctorate Recipients (SDR), collected every two years, recently changed its design to include both cross-sectional and longitudinal samples. For many respondents, the main variables of interest (e.g., job title and employer) will not change biannually. If the web survey was pre-filled with the responses from the last cycle and respondents were asked if the information is still accurate (a method called dependent interviewing), it could potentially reduce respondent burden and the administration time of the survey, not to mention provide consistency of responses. We tested the feasibility of using dependent interviewing (DI) with respondents (and their data) to the 2019 SDR cycle. The pilot tested two versions of the SDR web questionnaire that differed in how the DI information was presented, alongside the standard (not pre-filled) SDR web questionnaire to serve as a control. In addition, all 3 versions included new questions about the respondent's experience, addressing topics like their sensitivity to and enjoyment of DI. Research questions of interest included: real and perceived time to complete the survey, underreporting of job changes, usability issues, evidence of satisficing, and respondents' impressions of the experience. The three versions of the DI questionnaire were compared for significant differences in data quality and respondent preference.

Evaluating Previously Reported Data in the Census of Agriculture: Results from Usability Testing
Heather Ridolfo, *National Agricultural Statistics Service*
Kenneth Pick, *National Agricultural Statistics Service*

The National Agricultural Statistics Service (NASS) conducts hundreds of surveys a year covering all aspects of U.S. agriculture. Farm and ranch operations that are large or produce key commodities are often surveyed multiple times per year. These respondents often provide constructive feedback to NASS, stating that they are asked to report the same information in different surveys throughout the calendar year. To reduce response burden, NASS is currently evaluating the use of previously reported data (PRD) in survey data collection. In 2019, NASS explored the use of PRD in the Census of Agriculture Content Test. The Census of Agriculture (COA) is conducted every five years (in years ending in 2 and 7) and the Content Test serves as a dress rehearsal for the upcoming 2022 COA. PRD was added to five sections of the web version of the questionnaire using respondents' answers to the 2017 COA. Usability testing was conducted to evaluate if the use of PRD assisted respondents in completing the COA, improved response efficiency, and impacted data quality. This paper will present the results of this research.

Utilizing Respondents' Previously Reported Data in a Census of Establishments: Results from an Experiment in the Census of Agriculture's 2020 Content Test
Joseph Rodhouse, *National Agricultural Statistics Service*
Kathy Ott, *National Agricultural Statistics Service*
Zachary Terner, *National Institute of Statistical Sciences*

In preparation for the 2022 Census of Agriculture (COA), an experiment was conducted in the 2020 COA Content Test. The study goals were to assess whether informing sampled establishments of the availability of previously reported data (PRD) in the COA web form would increase response rates and whether the presence of PRD reduced respondents' perceived burden of completing the form. Two versions of the survey contact materials were created: i) a control version with standard contact language and no mention of PRD, and ii) an alternate version similar to the control, but with an additional sentence stating that PRD would be provided to aid in form completion. Sample units were randomly assigned to one of three groups: 1) a control with no PRD, 2) a PRD group receiving the control contact materials, and 3) a PRD group receiving the alternate contact materials. In this paper, the study results evaluating the impact of PRD on web response rates and whether that impact differs with the pre-notification of its use are discussed. Respondent perceptions of the ease of use, or burden, based on responses to debriefing questions administered at the end of the test forms are presented. Potential impacts that could help inform policy discussions of PRD in economic censuses are highlighted.

Discussant: Matt Jans, *ICF International*

Session K-4
Academic Statistical Agency Collaboratives to Create Data Infrastructure for Evidence-Building

Re-Engineering Statistics using Economic Transactions (RESET)
Matthew D. Shapiro, *University of Michigan*
Gabriel Ehrlich, *University of Michigan*
David Johnson, *University of Michigan*
John Haltiwanger, *University of Michigan*
Ron Jarmin, *U.S. Census Bureau*

Re-Engineering Statistics using Economic Transactions (RESET) aims to provide the architecture for re-engineering official economic statistics—literally to build key measurements such as GDP and consumer inflation from the ground up. Our new measurement architecture offers internally consistent real expenditure and inflation measures that adjust for product turnover and product quality change at scale. It builds measures of inflation and spending from granular, item-level transactions data. It therefore engineers statistics directly from the information systems of firms rather than superimposing a measurement system based on surveys and enumerations implemented by statistical agencies. To implement the architecture, the project, and ultimately the statistical agencies, will need to partner with firms. We have conducted successful pilot projects with multiple partner firms and plan in the new project to scale up the project to cover the entire Retail Trade sector. We are collaborating with leading computer scientists to develop techniques that perform quality adjustment under a range of circumstances—including cases where firms can provide refined labels of goods and cases where it is necessary to use machine learning techniques because existing product descriptions are not conveniently pre-coded.

Criminal Justice Administrative Records System: A Collaborative Approach to Building Next Generation Criminal Justice Data Infrastructure
Michael Mueller-Smith, *University of Michigan*
Keith Finlay, *U.S. Census Bureau*

The Criminal Justice Administrative Records System (CJARS) is a joint Census Bureau-University of Michigan project started in 2016 to create a national, integrated, harmonized collection of criminal justice microdata at the Census Bureau. CJARS links records from of tens of millions of criminal justice events, largely from state court and correctional administrative records, to earnings, employment, and other records. The project will both provide valuable aggregate statistical information to criminal justice agencies and increase the quality and quantity of criminal justice research by making the data available through the Federal Statistical Research Data Centers. This project highlights the opportunities enabled by collaboration between federal statistical agencies and academia. We discuss how responsibilities are distributed or shared across the institutions, as well as some of the governance decisions we have made to keep the project operational and efficient. Finally, we explore some of the long-term risks the project faces as it matures and becomes more integrated into federal statistical programs.

Decennial Census Digitization and Linkage Project
Katherine Genadek, *U.S. Census Bureau*
J. Trent Alexander, *University of Michigan*

The Decennial Census Digitization and Linkage project (DCDL) is digitizing and linking individual records across the 1960-1990 censuses and creating tools to improve the dissemination of these data. When combined with already-available linkages between the censuses of 1940, 2000, 2010, and soon-to-be 2020, DCDL will complete a massive longitudinal data infrastructure covering almost the entire U.S. population since 1940. The resulting data resource will provide transformational opportunities for research, education, and evidence-building across the social, behavioral, and economic sciences. We describe the collaborative

project's innovative methods of data rescue, record linkage, and data access.

Agricultural and Food Data Systems for 2021 and Beyond
Brent Hueth, *Economic Research Service*
Mark Denbaly, *Economic Research Service*

The National Household Food Acquisition and Purchase Survey (FoodAPS) is an established data product produced through a complex collaboration across multiple federal agencies and proprietary data providers. Designed initially as a collaboration between USDA ERS and Food and Nutrition Service (FNS), the product was conceived of in response to recognized limitations of U.S. consumption and expenditure surveys at the time. The Agricultural and Food Business Data series is a new product under development at USDA ERS in collaboration with academic partners, the Coleridge Initiative, several proprietary data providers, and the Census Bureau. The project aims to build a longitudinal sub-economy data frame for all businesses engaged in the agricultural and food supply chain. These products each are responding to recent CNSTAT consensus-report recommendations regarding modernization of data collection and statistical reporting on food demand and human nutrition, and agricultural production, processing, and distribution. Both reports emphasize the need to complement traditional survey products with administration and commercial data sources.

Discussant: Daniel Goroff, *Sloan Foundation*

Session K-5
Leveraging Data Science to Improve Survey Operations

Machine Learning and the Commodity Flow Survey
Christian Moscardi, *U.S. Census Bureau*

The Commodity Flow Survey (CFS), a joint effort between the Bureau of Transportation Statistics and the Census Bureau, has implemented a machine learning (ML) process to improve data quality as well as reduce operational costs and respondent burden. This has improved production data quality and the resulting estimates, while simultaneously reducing respondent burden and costs. While this work was implemented into 2017 CFS production, we are working on ways to make a bigger impact with data science approaches in future CFS collections. First, we are using Amazon's MTurk to develop a better training dataset to improve the model's performance. Second, we are deploying and further integrating this model into future (2022 and beyond) CFS collections, to replace a burdensome survey question and improve data quality. This subsequently unlocks our ability to collect significantly more data from respondents, enabling more granular and higher-quality estimates without significantly increasing respondent burden. During this talk, we will share experiences using Amazon MTurk to improve data quality, and provide a high-level overview of the processes we have put in place to integrate ML into production.

Parsing the Code of Federal Regulations for the Commodity Flow Survey's Hazardous Materials Supplement
Krista Chan, *U.S. Census Bureau*
Christian Moscardi, *U.S. Census Bureau*

In partnership with the Pipeline Hazardous Materials and Safety Administration (PHMSA), the Census Bureau is implementing a hazardous materials (hazmat) supplement to the Commodity Flow Survey (CFS) called the Expanded Hazardous Materials Supplement. We are asking hazmat shippers to provide information about materials shipped and the packaging used to protect those shipments in transit. Hazmat packaging is federally regulated - the Code of Federal Regulations (CFR) specifies packaging for each hazardous material that can be shipped. However, these specifications are not in a structured data format - they are written as English natural-language text. We have used Natural Language Processing (NLP) techniques to categorize the packaging regulations into a structured data format. We can now use this information to augment survey response data, meaning that we can produce richer data about hazmat packaging for PHMSA without needing to further burden respondents by asking them look through opaque regulatory text themselves. Last, PHMSA will use this structured data to enable easier and more streamlined searching through the CFR, e.g. for companies that need to comply with hazmat packaging regulations.

Using PDF Extraction and Web Scraping Tools to Collect Government Health Insurance Plan Information
Virginia Gwengi, *U.S. Census Bureau*

The Census Bureau serves as the data collection agent for the Agency for Healthcare Research and Quality (AHRQ) for the Medical Expenditure Panel Survey- Insurance Component (MEPS-IC). The survey collects data on health insurance form private and public sector employers. Although the survey has a sample size of 45,000, in this work, we focus on approximately 1,000 government units that are sampled annually. Because these units are sampled every year, to reduce the burden on these respondents, they are only asked to respond to respondent unit-level questions and some insurance plan-level questions. These respondents are then asked to upload plan brochures to the data collection instrument or to provide websites where Census analysts can search for the plan forms and manually extract the remaining plan-level information from these forms. This transfers burden of response from respondents to Census analysts. We seek to lessen this burden and to increase efficiency using data extraction and web scraping tools. In particular, we use search API and crawl respondent websites to collect Summary and Benefits and Coverage (SBC) forms, which have a standardized format that was mandated by the Affordable Care Act. Using Camelot, a python library to extract tables from text-based PDFs, we create a PDF extraction tool. We also employ regular expression as well as named entity recognition (NER) to extract response items from the SBC forms. The extraction tool collects the information faster than the manual process while the web scraping tool allows us to crawl each webpage and search for the SBC forms more efficiently.

Using Open Source Tools to Build a Custom Data Entry Application for Creating Truth Data
Cecile Murray, *U.S. Census Bureau*
Katherine Genadek, *U.S. Census Bureau*

The Decennial Census Digitization and Linkage project (DCDL) is producing linked restricted microdata files from the decennial censuses of 1960 through 1990. The information necessary to link these censuses is currently preserved on microfilm, and therefore cannot be linked at the individual level over time or to other data. In order to digitize this massive set of records, automation is required to reduce costs and ensure data quality. Using open-source Python libraries, we developed a custom user interface to enable staff to manually enter data from a subset of forms to serve as a training dataset for Optical Character Recognition (OCR) algorithms. These algorithms will enable us to digitize the majority of images in an automated way, so the quality of training data is paramount. Our streamlined data entry system allows us to efficiently obtain such high-quality training data. This presentation will discuss the conceptual and practical hurdles we encountered in developing and deploying the app, as well as how the code could be re-purposed for other projects.

Discussant: Kevin Deardorff, *U.S. Census Bureau*