



# Federal Committee on Statistical Methodology

## 2022 Research and Policy Conference

October 25-27, 2022

Walter E. Washington Convention Center  
Washington, DC

**Full Program**

*Updated: October 20, 2022*

**Sponsored By:**



**Hosted By:**



## Federal Sponsors

The 2022 FCSM Research and Policy Conference received generous support from the following federal agencies.



# **FCSM 2022 Research and Policy Conference Program Committee**

Jessica Graber, Co-Chair, *National Center of Health Statistics*

Jeffery Gonzalez, Co-Chair, *Bureau of Labor Statistics*

Darius Singpurwalla, Co-Chair, *National Center for Science and Engineering Statistics*

David Kashihara, Co-Chair, *Agency for Healthcare Research and Quality*

Paul Schroeder, Co-Chair, *Council of Professional Associations on Federal Statistics*

Scott Boggess, *U.S. Census Bureau*

Richard Chard, *Social Security Administration*

Yang Cheng, *National Agricultural Statistics Service*

Stephanie Coffey, *U.S. Census Bureau*

Joseph Conklin, *Bureau of Justice Statistics*

Ben Franco, *Social Security Administration*

Ellen Galantucci, *Federal Maritime Commission*

Jessica Holzberg, *U.S. Census Bureau*

Kyle Hood, *Bureau of Economic Analysis*

Tiandong Li, *Department of Housing and Urban Development*

Janine McFadden, *Bureau of Transportation Statistics*

Ann Miller, *Department of the Interior*

Chris Moriarity, *National Center for Health Statistics*

Xuan Pham, *Economic Research Service*

Jennifer Rhorer, *National Agricultural Statistics Service*

Maura Spiegelman, *National Center for Education Statistics*

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*

Morgan Earp, *National Center for Health Statistics*

John Eltinge, *U.S. Census Bureau*

John Finamore, *National Center for Science and Engineering Statistics*

Dennis Fixler, *Bureau of Economic Analysis*

Ellen Galantucci, *Federal Maritime Commission*

Michael Hawes, *U.S. Census Bureau*

Travis Hoppe, *National Center for Health Statistics*

Wendy Martinez, *U.S. Census Bureau*

Lisa Mirel, *National Center for Health Statistics*

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*

Mark Prell, *Economic Research Service*

Michael Ratcliffe, *U.S. Census Bureau*

Rolf R. Schmitt, *Bureau of Transportation Statistics*

Darius Singpurwalla, *National Center for Science and Engineering Statistics*

Robert Siviniski, *Office of Management and Budget*

Matt Spence, *U.S. Census Bureau*

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

Linda J. Young, *National Agricultural Statistics Service*



*Council of Professional Associations on Federal Statistics*

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.

# 2022 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; and
- Provide a mechanism for statisticians in different federal agencies to meet and exchange ideas.

The theme for the 2022 FCSM Research and Policy Conference is: Great Expectations: New Directions and Innovations for Sustainable Federal Statistics. 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 Speakers

**Karin Orvis, Office of Management and Budget**  
**Robert Santos, U.S. Census Bureau**  
**Xiao-Li Meng, Harvard University**

### **Tuesday, October 25**

**8:30AM**

**Dr. Karin Orvis** is the new Chief Statistician of the U.S. On April 29, 2022, Dr. Orvis joined the Office of Management and Budget (OMB) to serve in this role. As CSOTUS, Orvis will oversee the federal statistical system, serve as chair of the Interagency Council on Statistical Policy, and provide governmentwide leadership in the continuing implementation of the Foundations for Evidence-Based Policymaking Act. At the conference, Dr. Orvis will share her vision and priorities for the statistical system. She previously served as the director of the Defense Suicide Prevention Office within the Department of Defense. She also held several other positions within the Department of Defense including acting principal director of Military Community and Family Policy and director of the Transition to Veterans Program Office. She holds a doctorate and a master of arts in industrial/organizational psychology from George Mason University.

**Robert Santos** is the 26th director of the U.S. Census Bureau. Santos' career spans more than 40 years in survey research, statistical design and analysis, and executive-level management. Santos served as the 2021 president of the American Statistical Association (ASA) and is an ASA Fellow and recipient of the ASA Founder's Award in 2006. He was the 2014 president of the American Association for Public Opinion Research (AAPOR) and received the 2021 AAPOR Award for Exceptionally Distinguished Achievement. He earned a B.A. in mathematics from Trinity University in San Antonio and an M.A. in statistics from the University of Michigan at Ann Arbor. Mr. Santos's talk is entitled, "The Role of Data Quality in a 21st Century Federal Statistical System".

### **Wednesday, October 26**

**8:30AM**

**Xiao-Li Meng** is the Whipple V. N. Jones Professor of Statistics at Harvard University and the Founding Editor-in-Chief of Harvard Data Science Review. Meng was named the best statistician under the age of 40 by Committee of Presidents of Statistical Societies (COPSS) in 2001, and he is the recipient of numerous awards and honors for his more than 150 publications in at least a dozen theoretical and methodological areas, as well as in areas of pedagogy and professional development. In 2020, he was elected to the American Academy of Arts and Sciences. Meng received his BS in mathematics from Fudan University in 1982 and his PhD in statistics from Harvard in 1990. He was on the faculty of the University of Chicago from 1991 to 2001 before returning to Harvard, where he served as the Chair of the Department of Statistics (2004–2012) and the Dean of Graduate School of Arts and Sciences (2012–2017). Dr. Meng's talk is entitled, "Building Representative Miniatures out of Non-representative Big Data: An Interplay of Data Quantity, Quality, and Privacy".

## Schedule Overview

### **Tuesday, October 25**

7:30 am – 5:00 pm **Registration Open**

7:30 – 8:30 am **Continental Breakfast**, *Sponsored by Westat*

8:30 - 9:30 am **PLENARY SESSION I (146ABC)**

- Welcoming Remarks
- Karin Orvis, *Chief Statistician of the United States*
- Robert Santos, *U.S. Census Bureau*

9:30 - 10:00 am **Coffee Break**

10:00 - 11:45 am **CONCURRENT SESSIONS A**

**A-1.** Economic Statistics at the Cutting Edge (146A)

**A-2.** Data Linkage, Privacy, and Modernization (146B)

**A-3.** Statistical Innovation and Measuring the Innovative (146C)

**A-4.** Best Practices for Monitoring and Evaluating the ARP, IIA and Other Programs:  
Report of the Department of Commerce Data Governance Working Group (140A)

**A-6.** Sampling and Participation (144ABC)

11:45 am - 1:15 pm **Lunch on Your Own**

1:15 – 3:00 pm **CONCURRENT SESSIONS B**

**B-1.** Assessing Quality in Non-Traditional Data Sources for National Statistics: Report from the Task Team on Big Data for the Sustainable Development Goals (146A)

**B-2.** Using Linked Data to Understand Mortality Patterns for Vulnerable Populations (146B)

**B-3.** Considerations and Advancements in Sampling (146C)

**B-4.** Pandemic and Related Fiscal Policies (140A)

**B-5.** Information Quality Standards and Processes (140B)

**B-6.** Leveraging Probability Panels for Federally Sponsored Statistical Data Collections (144ABC)

3:00 – 3:15 pm **Break**

3:15 – 5:00 pm **CONCURRENT SESSIONS C**

**C-1.** Methods for Refining and Supplementing Food Security Statistics and Importance to Policy Research (146A)

**C-2.** Furthering Research through Data Linkages (146B)

**C-3.** Advances in Small Area Estimation (146C)

**C-4.** Strengthening the Autonomy of the Federal Statistical Agencies (140A)

**C-5.** Advancements in Sexual Orientation and Gender Identity Measurement (140B)

**C-6.** Lifting the Fog: Editing and Imputation (144ABC)

## **Wednesday, October 26**

7:30 am – 5:00 pm **Registration Open**

8:30 - 9:30 am **PLENARY SESSION II (146ABC)**

- Xiao-Li Meng, *Harvard University*

9:30 - 10:00 am **Coffee Break, Sponsored by the American Statistical Association**

10:00 - 11:45 am **CONCURRENT SESSIONS D**

**D-1.** Advancing Efficiency and Accuracy with Data Science Techniques (146A)

**D-2.** Dying for Better Data: Using Administrative Records to Improve the Quality and Utility of Health Survey Data (146B)

**D-3.** Evaluation of EEOC Collection of Pay Data: Quality for Intended Use and Recommendations for the Future (146C)

**D-4.** Statistical Issues in Research to Inform Evidence-based Nutrition Security Policymaking (140A)

**D-5.** Reimagining NCES and the Education Data Ecosystem for the 21st Century (140B)

**D-6.** Evaluating Methods for Evaluating Questionnaires (144ABC)

11:45 am – 1:15 pm **Lunch on Your Own**

1:15 – 3:00 pm **CONCURRENT SESSIONS E**

**E-1.** Statistical Modeling (146A)

**E-2.** Mosaic Data Linkage (146B)

**E-3.** Modernizing Bias Reduction Techniques (146C)

**E-4.** Topics in Income Volatility Measurement (140A)

**E-5.** Implementing Title III of the Evidence Act: A Framework to Expanding Data Access for Evidence Building (140B)

**E-6.** Measuring Sex, Gender Identity, and Sexual Orientation: Findings of the National Academies (144ABC)

3:00 - 3:15 pm **Break**

3:15 – 5:00 pm **CONCURRENT SESSIONS F**

**F-1.** Applications of Statistical Methods in Economic Areas (146A)

**F-2.** Innovations from the Criminal Justice Administrative Records System (CJARS) Project (146B)

**F-3.** Drawing Optimal Statistical Inferences from Nonprobability or Low Response-rate Probability Surveys: Some Recent Advances and Applications (146C)

**F-4.** The Need for a National Study of Post-Secondary Faculty (140A)

**F-5.** Classifying Settlement Form and Function: Urban, Rural, Metropolitan and Micropolitan Area Definitions and Classifications (140B)

**F-6.** Reporting on the Pulse of America – Rapid Data and Collection and Information Sharing Initiatives in the Federal Government (144ABC)

## **Thursday, October 27**

7:30 am – 5:00 pm    **Registration Open**

### **8:30 - 10:15 am        CONCURRENT SESSIONS G**

**G-1.** Insights from the Congressional Watchdog on Equity in Evaluation Design *(140A)*

**G-2.** Highlights of CNSTAT Report on Transparency in Statistical Information *(140B)*

**G-3.** Using Spatial Data to Inform Official Statistics *(143A)*

**G-4.** Connecting the Dots: Leveraging Extant Data for Research *(143BC)*

**G-5.** Building a Sampling System for the Annual Integrated Economic Survey *(144ABC)*

10:15 - 10:30 am    **Break**

### **10:30 am - 12:15 pm CONCURRENT SESSIONS H**

**H-1.** Data Science Improving Statistical Products at the National Center for Education Statistics *(140A)*

**H-2.** Responsible Creation, Handling, and Use of Sexual Orientation and Gender Identity Data *(140B)*

**H-3.** Practical Applications of Web-Scraping *(143A)*

**H-4.** If You Can't Beat Them, Join Them: Leveraging Administrative Data to Evaluate Survey Responses and Enhance the Value of Survey Data *(143BC)*

**H-5.** Web Survey Design Standards for the Enterprise Data Collection System at the U.S. Census Bureau *(144ABC)*

12:15 - 1:30 pm     **Lunch on Your Own**

### **1:30 – 3:15 pm        CONCURRENT SESSIONS I**

**I-1.** Foreign-Born Scientists and Engineers and the U.S. Workforce Evidence Building *(140A)*

**I-2.** A Marshall Plan for Reconstructing National Statistical Offices After Conflict: Practical Guidance from International Principles *(140B)*

**I-4.** Developing the Infrastructure to Build Data Infrastructure: The Decennial Census Digitization and Linkage Project *(143BC)*

**I-5.** Innovations in Health Care Survey Data Collection *(144ABC)*

3:15 – 3:30 pm     **Break**

### **3:30 – 5:15 pm        CONCURRENT SESSIONS J**

**J-1.** Evolution of the U.S. Economic Census for 2022 – An Overview of Changes, Testing, and Innovations Employed to Support Continued Improvement *(140A)*

**J-5.** Analyzing Respondent Behavior and Data Quality in Web Surveys *(144ABC)*

---

## ***Abstract Listings for Tuesday, October 25***

---

- |                         |                  |
|-------------------------|------------------|
| • Plenary Session I     | 8:30 – 9:30 am   |
| • Concurrent Sessions A | 10:00 – 11:45 am |
| • Concurrent Sessions B | 1:15 – 3:00 pm   |
| • Concurrent Sessions C | 3:15 – 5:00 pm   |

**Tuesday, October 25**

**8:30AM**

---

***Plenary Session I***

---

Organizer: Shelly Martinez, *Office of Management and Budget*  
Room 146ABC

**Welcoming Remarks**

Shelly Martinez, *Office of Management and Budget*

**Remarks from the Chief Statistician of the United States**

Karin Orvis, *Chief Statistician of the United States*

**The Role of Data Quality in a 21st Century Federal Statistical System**

Rob Santos, *U.S. Census Bureau*

As we advance into the 21st century, we see a new statistical universe, one that features new and complex data user needs, data collection challenges, pressures to do more with less. We also see the need for improved collaborations with stakeholders, partners, policy makers and the public. We see stronger computing power, proliferation of alternative unofficial data products, and emergence of new technologies. Censuses and surveys alone, while still critical, can no longer answer the questions completely or quickly enough to satisfy the modern need for data. The U.S. Census Bureau is committed to producing data that reflects an accurate portrait of America both in its population and its economy. To effectively meet our mission, we need a strategy that blends new and traditional data sources and aligns our talent and resources with a data-centric vision of a federal statistical agency. But in doing so, we must be cognizant of data quality, what it means, and how different levels offer different levels of insight and inference. Specifically, this presentation reveals an underlying assumption that is seldom asked: how we can ensure the data we collect is relevant to users and timely for their needs? The principles of diversity, equity and inclusion are used to illustrate the merit of questioning the quality of the data that are currently being collected.

---

**Session A-1: Economic Statistics at the Cutting Edge**

---

Organizer: Kyle Hood, *Bureau of Economic Analysis*

Chair: Kyle Hood, *Bureau of Economic Analysis*

Room 146A

**Do Workers have a Firm Sense of Industry Codes? Comparing Worker and Employer Descriptions of Business Activity**

Melissa Chow, *U.S. Census Bureau*

Martha Stinson, *U.S. Census Bureau*

Many government agencies produce statistics stratified by firm industrial classification to identify economic trends in the economy. Data collectors can gather information about industry directly from the firm or they can use household surveys to ask workers about the activities of their employers. At the Census Bureau, both types of industry codes are collected and used for publication. However, little is known about the differences between household and employer industry reports and how these might impact official statistics and research that rely on industry codes.

In this paper, we link the American Community Survey to the Longitudinal Business Database and compare industry codes reported by people to those reported by their employers. We calculate sector-specific agreement rates and show how these are related to respondent demographics and business characteristics. We also examine how published statistics based on household surveys change when employer-reported industry is used and how this change impacts the measurement of important workforce characteristics and trends such as the number of essential workers or the shift of workers to “factory-less” firms. Finally, we measure how the role of industry changes in explaining earnings differentials by gender, race, and nativity when we use different data sources to measure industry of work.

**Learning a Low-Dimensional Representation of Job History for Economic Adjustment**

Keyon Vafa, *Columbia University*

David Blei, *Columbia University*

Susan Athey, *Stanford University*

The relationship between individual employment history and labor market outcomes such as wages or employment is central in the field of labor economics. Empirical studies typically attempt to compare groups of workers who are similar in terms of observable characteristics but are different in terms of an intervention or characteristic of interest. These studies produce estimates that are said to be “adjusted” for the covariates of interest. Although an individual's complete job history may explain labor market outcomes, it is too high-dimensional to adjust for in full. Instead, economists typically summarize an individual's career by including hand-constructed summary statistics about the past, such as years of experience and current occupation. However, the interpretation of adjusted estimates often depends on whether the researcher has adequately adjusted for worker characteristics. In this work, we propose estimating adjusted quantities by conditioning on a low-dimensional representation of an individual's complete job history. Specifically, we develop CAREER, a machine learning method that uses neural networks to learn low-dimensional representations of job sequences. CAREER takes inspiration from methods developed for natural language processing; instead of modeling sequences of words, it models sequences of jobs. This method uses large-scale auxiliary data to aid the estimation of adjusted quantities on smaller datasets. To demonstrate, we estimate the gender wage gap after adjusting for individuals' full job history on PSID. We show that adjusting for job history can reduce the unexplained gender wage gap.

## **Estimating Women's Eligibility for WIC: Postpartum and Breastfeeding Women**

Linden McBride, *U.S. Census Bureau*

Maria Perez-Patron, *U.S. Census Bureau*

Renuka Bhaskar, *U.S. Census Bureau*

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) provides benefits to low-income, nutritionally at-risk infants, children, and women. As part of an ongoing research effort, Census Bureau researchers estimate WIC eligibility among infants and children for several states using data from administrative records and the American Community Survey (ACS). These eligibility estimates, combined with data on participation, provide WIC administrative offices with important information on the reach and uptake of local WIC programs in so far as infants and children are concerned. However, Census Bureau researchers do not provide estimates of WIC eligibility among women because the ACS does not include sufficient data on whether women are pregnant, postpartum, or breastfeeding, criteria that determine women's categorical eligibility for WIC. We propose a strategy to resolve one part of this problem: identifying postpartum and breastfeeding women. We identify postpartum women in the ACS by using birth date and relationship data for children in the ACS household. To estimate the probability of breastfeeding among postpartum WIC-eligible women, we generate a predictive model based on observable characteristics in the Center for Disease Control's National Immunization Surveys and apply that model to the same characteristics in the ACS.

## **Twenty-One Years of Adjustments for Quality Change in the U.S. Consumer Price Index**

Brendan Williams, *Bureau of Labor Statistics*

Quality change is frequently cited as a source of bias in the CPI. We provide a systematic accounting of the impact of quality adjustments and other item replacement imputations on the U.S. CPI over the past 21 years. We produce counterfactual indexes varying the use of quality adjustment, direct comparisons, and class-mean imputation for all commodities and services in the CPI. We find that hedonic adjustment has a large impact on a limited number of categories in the CPI. Differences due to the treatment of item replacements can be substantial. Our results provide guidance on how item replacement methodologies should be used in the CPI to address biases.

**Tuesday, October 25**

**10:00AM**

---

### ***Session A-2: Data Linkage, Privacy, and Modernization***

---

Organizer: Janine McFadden, *Bureau of Transportation Statistics*

Chair: Janine McFadden, *Bureau of Transportation Statistics*

Room 146B

## **Modernization of NCHS' National Health Care Surveys to Meet Continually Evolving Data Needs for the 21st Century**

Brian W. Ward, *National Center for Health Statistics*

Carol J. DeFrances, *National Center for Health Statistics*

Manisha Sengupta, *National Center for Health Statistics*

For many decades, the National Center for Health Statistics (NCHS) has conducted the National Health Care Surveys – a family of nationally representative surveys that are provider-based, and cover a broad spectrum of ambulatory, hospital, and long-term care settings. To continue their responsiveness to the ever-changing health care delivery system and the related decline of survey participation response rates, the National Health Care Surveys have engaged in ongoing modernization and evaluation to improve the quality and efficiency of existing and proposed survey operations. This presentation provides an overview of the improvements and modernization strategies implemented for the National Health Care Surveys to collect and disseminate ambulatory, hospital, and long-term care data. This includes the use of electronic health record (EHR) data and development of new EHR standards, development of cloud-based CIPSEA-compliant

infrastructure to support collection of these electronic data, use of new statistical methodologies and model-based estimation, and finally the push for more timely dissemination through preliminary data estimates via user-friendly dashboards.

## **The NCHS Data Presentation Standards for Rates and Counts**

Jennifer Parker, *National Center for Health Statistics*

The National Center for Health Statistics (NCHS) disseminates information on a broad range of health topics through diverse publications. These publications rely on transparent standards that can be broadly and efficiently applied. Standards are particularly important when indicators of precision cannot be included with the estimates, such as for short communications on social media. Updated NCHS Data Presentation Standards for Proportions were implemented in 2017.

This presentation provides an overview of the new NCHS Data Presentation Standards for Rates and Counts and initial implications for data dissemination.

As with the 2017 Standards for Proportions, the multistep NCHS Data Presentation Standards for Rates and Counts are based on a minimum sample size, on the relative width of a confidence interval, and on the degrees of freedom for estimates for population health surveys. Specific criteria differ between vital statistics and population health surveys, differ between rates with denominators assumed to be constant and those assumed to be non-constant, and differ from those used in the standards for proportions.

## **Fulfilling Statistical Policies with Data Curation Practices**

Leighton L Christiansen, *Bureau of Transportation Statistics*

Jesse Long, *Bureau of Transportation Statistics*

The Foundations for Evidence-Based Policy Making Act calls on statistical agencies to adopt “best practices and appropriate procedures” that will enable them to produce relevant statistical information, conduct credible, accurate, and object statistical activities, and project the trust of information providers. This means statistical agencies need to engage with data assets over their entire lifecycle, from planning and collection, to re-use and storage. Data curation is a suite of best practices and procedures that greatly enhances statistical data processing by enabling greater data discovery and retrieval, maintaining data quality, adding value, and providing for re-use over time. By adopting data curation best practices, agencies can set themselves for long-term success and policy fulfilment. The speakers, staff at the Bureau of Transportation Statistics’ (BTS) National Transportation Library (NTL), will explore how data curation practices can benefit U.S. statistical agencies.

## **Comparison of Poisson-gamma and Laplace Mechanisms for Differential Privacy**

Harrison Quick, *Drexel University*

Motivated by the lack of formal privacy protections in place on CDC WONDER, recent work has proposed the creation of a Synthetic CDC WONDER based on a differentially private Poisson-gamma modeling framework, which samples values from the posterior predictive distribution associated with modeling event count data with a Poisson-likelihood and assuming a gamma prior on the underlying event rate. The Poisson-gamma framework incorporates (and relies on) publicly available information such as estimates of the underlying population sizes and event rates to improve its utility and protects the sensitive data by increasing the informativeness of the prior distribution. The goal of this work is to present a comparison of the Poisson-gamma framework and the Laplace mechanism for the purpose of generating a synthetic dataset comprised of the 26,000 cancer-related deaths in Pennsylvania counties from 1980. We show that while the Poisson-gamma framework preserves inference on quantities such as urban/rural and black/white disparities in death rates, the Laplace mechanism – when forced to produce non-negative values – can fail to preserve both the magnitude and direction of such disparities.

---

**Session A-3: Statistical Innovation and Measuring the Innovative**

---

Organizer: Ben Franco, *Social Security Administration*

Chair: Ben Franco, *Social Security Administration*

Room 146C

**The Census Bureau's New Annual Total Business Demographics Program: A Hybrid-Data Approach**

Adela Luque, *U.S. Census Bureau*

Aneta Erdie, *U.S. Census Bureau*

In response to declining survey response, increasing imputation rates and costs, and a higher demand for reliable, frequent and timely data, the Census Bureau initiated a consolidation of business surveys and the committed leveraging of administrative records and Census data to provide firm information by business owners' demographics. The result is the new Census Bureau's Annual Total Business Demographics Program, which provides demographic characteristics for US businesses through a hybrid-data approach that combines data from the Annual Business Survey for employer firms, and the new annual (administrative-records based) Nonemployers Statistics by Demographics (NES-D). Among other things, these data make possible the examination of business ownership trends by owner demographics, geography and industry, and of how business cycles and shocks play out across different demographic groups.

We will offer an overview of the Annual Total Business Demographics Program, provide high-level estimates for both employer and nonemployer firms, and the timeline on employer topics such as the adoption of technology and innovation. We will also discuss the renewed interest in nonemployers, the rise in the share of minority nonemployer ownership, NES-D's value-added over its predecessor (Survey of Business Owners), and our plans to enhance NES-D with characteristics relevant to nonemployers.

**Are Some Innovation Self-Reports in the Annual Business Survey Biased? A Regression Discontinuity Test**

Luyi Han, *Pennsylvania State University*

Zheng Tian, *Pennsylvania State University*

Timothy R. Wojan, *National Center for Science and Engineering Statistics*

Stephan Goetz, *Pennsylvania State University*

An ongoing debate within innovation survey methodology concerns possible underreporting of innovation in joint innovation-R&D surveys relative to innovation-only surveys. Since the definition used to derive a positive measure of innovation is somewhat nebulous, the inclusion of R&D questions related to the number of PhDs employed, or the size of formal R&D budgets may cause respondents to apply a more stringent standard to the "new or significantly improved" innovation criterion. This potential source of bias is of particular concern in the Annual Business Survey where microbusinesses (fewer than 10 employees) answer both an R&D and innovation module but where businesses with 10 or more employees do not receive an R&D module. This analysis tests whether microbusinesses just under the sharp threshold are less likely to report innovation relative to small businesses just over the threshold using a regression discontinuity design. The analysis also demonstrates the importance of effect size considerations and statistical power for making sense of null hypothesis statistical testing applied to very large Federal datasets.

**Ownership Team Diversity and Innovation: Using Split Sample Design to Generate More Credible Findings**

Timothy R. Wojan, *National Center for Science and Engineering Statistics*

Splitting datasets into an exploratory sample for specification testing and a confirmatory sample for hypothesis testing has been suggested to increase the replicability of study results. Replicability is a central concern when a large set of reasonable hypotheses are to be tested since the probability of false discovery

increases with the number of tests. The hypothesis that diversity among ownership teams increases (decreases) the probability of innovation poses this problem of false discovery as there are many types of diversity that could be tested (e.g., based on race/ethnicity, gender, educational background, age, foreign born status, and composites) using many alternative measures (e.g., ethnic-linguistic fractionalization; Shannon, Simpson, or Brillouin diversity indices; etc.). Numerous specification tests might produce a subset of reported findings that appear to be highly statistically significant and support the author's priors but that fail to capture true phenomena in the population. This paper will discuss protocols for splitting the sample, establishing thresholds for passing exploratory tests through for confirmation, publishing a pre-registration plan that documents the exact set of confirmatory tests to be performed including the corresponding Bonferroni correction, and performing confirmatory hypothesis tests.

## **Modeling Approaches to Estimate Community Annoyance Due to Sonic Booms using Data from Repeated Surveys**

Robyn Ferg, *Westat*

Jean Opsomer, *Westat*

In an on-going project for NASA, we are developing a modeling approach to analyze the relationship between the noise level of sonic booms from an experimental supersonic plane and the level of annoyance measured through community response surveys. The goal of the project is to obtain a quantitative relationship between noise level and annoyance that is representative for the affected population. Particular modeling challenges include multiple annoyance measurements per survey respondents and very few occurrences of annoyance overall. To address these challenges, we propose a two-stage model for the presence of high annoyance, with the first stage modeling the probability that a respondent is ever highly annoyed and the second stage a multilevel logistic regression model for high annoyance based on noise level and demographic characteristics. We use a variation of Multilevel Regression and Poststratification (Gelman and Little, 1997) to obtain an overall representative noise-annoyance curve for the population. The approach is applied to data from a NASA pilot study.

## **The Power of Pandemic Nearcasting to Harness Bigdata and Nudge Evidence-Based Public Policymaking 2.0**

Asaph Young Chun, *Seoul National University and ISR Foundation*

David Cho, *ISR Foundation*

KyuBin Moon, *ISR Foundation*

Benjamin YoungJin Tae, *ISR Foundation*

Today, nowcasting the course of COVID-19 pandemic faces a Herculean task as challenging as forecasting the 1918 Spanish flu that claimed over 55 million lives across continents. In the case of Korea, science-and-data-based modeling has been a cornerstone of policymaking since the COVID-19 infected the Korean population in February, 2019.

The purpose of this paper is to unravel the coevolution of forecasting science and public policy throughout the course of COVID-19 pandemic in Korea. We initially used the model of Incidence Decay with Exponential Adjustment (IDEA) based on a well-established model that classifies population into the three groups: Susceptibles, Infected, and Recovered (SIR). The IDEA model outcomes have guided Korea to develop and adjust measures of non-pharmaceutical interventions (NPI), such as masking, social distancing, and business and school closures. We will show the extent to which model outcomes have been aligned with people's compliance with social distancing as well as COVID-19 transmission speed and scope in a number of crucial junctures, and point out, when necessary, the nature of gaps between prediction science and policy implementation.

Since the middle of August 2021 when Korea was inflicted by the 2nd wave of COVID-19, we turned to developing and integrating multiple prediction models that are grounded on forecasting methodology and bigdata science in a transdisciplinary manner. The data used in modeling included official organic data and commercial big data, both of which enhanced the rigor of prediction modeling. The outcomes of big data science-based modeling in Korea include nearcasting models quantifying the impact of NPI and predicting infection cases and mortality to harness the impact of COVID-19. The paper points out conditions for further innovation in prediction science and issues for evidence-based policymaking 2.0.

**Tuesday, October 25**

**10:00AM**

---

***Session A-4: Best Practices for Monitoring and Evaluating the ARP, IJA and Other Programs: Report of the Department of Commerce Data Governance Working Group***

---

Organizer: Carla Medalia, *U.S. Census Bureau*

Chair: Ben Page, *Department of Commerce*

Discussant: Oliver Wise, *Department of Commerce*

Room 140A

**Framework for Enabling Program Evaluation in the Federal Government**

Ron Jarmin, *U.S. Census Bureau*

The Department of Commerce (DOC) Data Governance Working Group (DGWG) issued a report in May 2022 to provide federal agencies with a useful framework for enabling program evaluation. The working group included representation from across DOC, including the office of the Chief Data Officer, agencies which implement federal programs, and statistical agencies with data and evaluation expertise. This presentation will highlight the main findings of the report and describe its recommendations. One chief finding is that agencies should consider a three-dimensional framework when developing program evaluation plans considering the phases of program implementation, direct versus indirect program impacts, and projected versus observed program impacts. Another central recommendation is that agencies should consider implementing large-scale observational studies that link program administrative data to previously collected data from censuses, surveys, administrative records, commercial vendors, and aggregated indices. The Census Bureau's Data Linkage Infrastructure may be particularly useful in enabling this type of analysis.

**Census Bureau's Data Linkage Infrastructure for Program Evaluation**

Carla Medalia, *U.S. Census Bureau*

The U.S. Census Bureau maintains a Data Linkage Infrastructure of census, survey, administrative, and third-party data that supports high-quality research and evaluation, advancing the Census Bureau's mission of providing timely and unbiased data to support evidence-based decision making. The Data Linkage Infrastructure includes data on households and businesses which come from census and survey collections as well as administrative records. All data, whether census, survey, or administrative records, are assigned unique identifiers at the individual, address, or organizational level either probabilistically or deterministically, depending on the information available. This presentation will describe the data maintained in the Census Bureau's Data Linkage Infrastructure and discuss how these data can be leveraged for program evaluation. It will also highlight a prototype living library of program impact metrics that agencies may consider examining as they develop their program evaluation plans.

**Data Requirements for Program Evaluation: Perspectives from an Implementing Bureau**

Ryan Smith, *Economic Development Administration*

The Economic Development Administration (EDA) was allocated \$3 billion in supplemental funding under the American Rescue Plan (ARP) to assist communities nationwide in their efforts to build back better by accelerating the economic recovery from the pandemic and building local economies that will be resilient to future economic shocks. The investments EDA makes under the American Rescue Plan will place strong emphasis on Equity and will directly benefit underserved communities impacted by the pandemic. EDA's program performance goals fall into two categories: promoting private enterprise and job creation in economically distressed communities and regions through critical infrastructure and revolving loan fund investments; and creating the conditions for long term economic development through non-infrastructure investments that build community capacity to achieve and sustain regional competitiveness and economic

growth. This talk will describe EDA's newly funded programs and discuss in detail the data needs to enable the evaluation of the programs funded under the ARP.

**Tuesday, October 25**

**10:00AM**

---

### ***Session A-6: Sampling and Participation***

---

Organizer: Jennifer Rhorer, *National Agricultural Statistics Service*

Chair: Danny Yang, *Bureau of Labor Statistics*

Room 144ABC

#### **Predicting Eligibility of a Very Rare Household Population to Improve Sample Design**

Jamie Ridenhour, *RTI International*

Joseph McMichael, *RTI International*

Rob Andrews, *National Oceanic and Atmospheric Administration*

Probability surveys of rare populations tend to need extremely large starting sample to yield enough data for analyzing the rare target population. In this paper, we discuss a modeling approach that utilized American Community Survey data, address-based sampling frame data, and historical survey data to create rare population propensity strata (i.e., predicted probability of being a member of the rare population) in a survey conducted by RTI on behalf of the National Oceanic and Atmospheric Administration in three U.S. states. We explain how the model was trained and tested, the resulting stratification, and how the sample was allocated to optimize precision given fixed costs. We also discuss how we dealt with sparseness of the eligibility outcome for the rare population in the training data and the lack of independent predictors. Our discussion includes alternative approaches and tradeoffs considered along with an evaluation of how closely the results obtained from data collection match what we were expecting based on the model prediction.

#### **Hitting the Target?: The Use of Targeted Samples in Probability-based Samples.**

Randall K. Thomas, *Ipsos Public Affairs*

Frances M. Barlas, *Ipsos Public Affairs*

Megan A. Hendrich, *Ipsos Public Affairs*

Kathleen Santos, *Ipsos Public Affairs*

Address-based (ABS) studies often attempt to obtain people with specific characteristics (e.g., black people, 18 to 24 year olds) at higher rates for more precision of group estimates. Households in the sample frame can have commercial-based information appended useful for targeted sample selection. Targeted samples can improve study efficiency and lower study costs. Though many researchers treat individuals with the desired characteristics from targeted sample as equivalent to those from non-targeted sample, individuals from targeted samples might differ. In the study we report, we were interested in obtaining higher rates of cigarette smokers. We first selected a non-targeted general population sample before selecting a targeted smoking sample. We compared the smokers obtained for each sample type and found that targeting was indeed associated with higher the prevalence of smokers, but also found differences in smoker characteristics – smokers from targeted households were older, smoked more, started smoking younger, and were less likely to use other tobacco products. We examined how weighting could offset the effect of targeting, but such adjustments can come at a cost in the unequal weighting effect. Our results show caution is recommended when using targeted samples.

#### **The Contrafactual Degrees of Freedom**

Phillip S. Kott, *RTI International*

Quite often in statistics, observations are assumed to be independent and identically distributed (iid). Not so, with observations coming from a complex sample survey featuring stratification, clustering, and unequal weighting. Nearly unbiased estimates of population and domain means computed from a complex survey sample and their standard errors are often quite different from what they would have been had the sample been iid. Although survey statisticians have developed reasonable methods for estimating the variance and

standard error of an estimated mean computed from a multistage stratified sample, they have been less successful in developing a reasonable measure of the degrees of freedom for such an estimated variance. The popular method of subtracting the number of strata from the number of PSUs is nothing short of scandalous for many applications. In its place, this author has suggested computing a contrafactual degrees of freedom based on (contrafactually) treating the sample elements as iid. Although far from perfect, an empirical investigation showed this approach to be often superior to the usual nominal degrees of freedom. In this presentation, the method of computing contrafactual degrees of freedom is extended from estimated population and domain means to estimated regression coefficients with the estimated difference of domain means serving as a special case.

### **Predicting Future Panel Participation in AmeriSpeak panel**

Stas Kolenikov, *NORC at the University of Chicago*

Ipek Bilgen, *NORC at the University of Chicago*

David Dutwin, *NORC at the University of Chicago*

Poulami Maitra, *NORC at the University of Chicago*

Attrition and nonresponse are important concerns in probability-based panels such as AmeriSpeak, given the importance of achieving a representative panel with active panelists and the cost of keeping the panel highly representative. Each panelist is thought of following a latent trajectory model with person-by-survey response propensity featuring person-specific intercepts (their propensity to respond to the very first survey they are invited) and person-specific slopes (the extent to which they lose interest with each subsequent survey). We extend this idea to build a multilevel crossed two-way model to estimate panelists' survey-specific propensity of unit response, incorporating predictors at appropriate levels, and estimate it using Bayesian methods (RStan). We find that slower attrition is associated with higher education, age, and being female. Racial and ethnic minorities have much lower initial propensities to respond than non-Hispanic whites, but attrite at slower rates. The past values of the self-reported willingness to participate in future surveys are predictive of unit response, validating it as a lead indicator of attrition. Predicting future response allows the panel operations implement interventions before attrition occurs.

### **Is A Picture Worth a Thousand Words?: Impact of Infographics on Response Rates in a Federal Survey**

Julia Merlin, *National Center for Education Statistics*

Allison Zotti, *U.S. Census Bureau*

The decline in federal survey response rates in recent years is well documented and challenges agencies' ability to report reliable estimates on subpopulations of interest. Previous research has indicated that providing potential respondents with personalized data during recruitment efforts can increase response rates, though with mixed effects. However, most experiments only measured high level response rates and lacked detail on the effect of personalization on the recruitment of respondents in key subgroups. The 2021-22 Teacher Follow-up Survey (TFS), a follow-up study to the 2020-21 National Teacher and Principal Survey (NTPS), conducted an infographic experiment to measure the effect of mailed general and personalized infographics on TFS response rates. Each infographic presents findings from prior collections with plain language explanations about the importance and utility of the survey data. Response rates amongst some subgroups show mixed effects, with both increased and decreased response rates for different types of respondents. This presentation will investigate the effect of each infographic condition by selected teacher or school characteristics to determine whether the infographics were more persuasive for some subgroups than others. These results could have practical implications for survey organizations seeking to use personalized outreach materials as a recruitment strategy.

**Tuesday, October 25**

**1:15PM**

---

***Session B-1: Assessing Quality in Non-Traditional Data Sources for National Statistics: Report from the Task Team on Big Data for the Sustainable Development Goals***

---

Organizer: Jennifer Park, *Committee on National Statistics, World Bank*

Chair: Maciej Truszczyński, *Statistics Denmark*

Discussant: Melissa Chiu, *Committee on National Statistics*

Room 146A

Sustainable Development Goals with their broad coverage of societal, environmental and economic fields pose challenges to a comprehensive statistical follow-up. Possible data sources towards this aim are widely discussed and data from non-traditional sources are frequently mentioned in this regard. In this context, data from non-traditional sources can include big data, geo data or citizen generated data. The introductory remarks to the session will give a short outline of the ongoing work on monitoring the SDGs and, in parallel, on how national statistical offices are working on methods to complement existing statistics outside of the framework of the SDGs with non-traditional data sources, both in order to address the increasing demand for data and decreasing resources. Different initiatives will be presented and challenges will be outlined. WHO will present the current aspirations, trials and successes in the use of non-traditional data in improving health statistics. Sweden will present the application of scanner data in calculation of household consumption expenditure and Poland, on the application of maritime data in the calculation of emissions. Finally, UK will present how to assess quality of non-official statistical sources for SDG reporting and more regular statistics.

**Assessing Quality in Non-Traditional Data Sources**

Maciej Truszczyński, *Statistics Denmark*

**Emissions and Ship Traffic**

Dominik Rozkrut, *Statistics Poland*

**Creation of a Composite Quality Indicator for Estimates Based on Administrative Data Using Clustering**

Roxanne Gagnon, *Statistics Canada*

**Nowcasting Global Poverty**

Daniel Gerszon Mahler, *World Bank*

---

**Session B-2: Using Linked Data to Understand Mortality Patterns for Vulnerable Populations**

---

Organizer: Sonya R. Porter, *U.S. Census Bureau*

Chair: Eric English, *U.S. Census Bureau*

Discussant: Lauren Rossen, *National Center for Health Statistics*

Room 146B

**Mortality in a Multi-State Cohort of Former State Prisoners, 2010-2015**

Sonya R. Porter, *U.S. Census Bureau*

Leticia Fernandez, *U.S. Census Bureau*

Eric English, *U.S. Census Bureau*

Sharon Ennis, *Department of Veteran Affairs*

E. Ann Carson, *Bureau of Justice Statistics*

In this study we explore the role that race/Hispanic origin, other demographic characteristics, and custodial/criminal history factors have on post-release mortality, including on the timing of deaths. We also assess whether conditional release to community supervision or reimprisonment may explain higher post-release mortality found among non-Hispanic whites. In the second part of the analysis, we estimate standardized mortality ratios (SMRs) by sex, age group, and race/Hispanic origin using as reference the U.S. general population. The data come from state prison releases from the Bureau of Justice Statistics' (BJS) National Corrections Reporting Program (NCRP). The NCRP records were linked to the Census Numident to identify deaths occurring within five years from prison release. We found that non-Hispanic white former prisoners were more likely to die within five years after prison release and more likely to die in the initial weeks after release compared to racial minorities and Hispanics. Reimprisonment, age at release, and a history of multiple prison terms had a similar influence on the odds of dying across all race/Hispanic origin groups.

**Age, Sex, and Racial/Ethnic Disparities and Temporal-Spatial Variation in Excess All-Cause Mortality During the COVID-19 Pandemic: Evidence from Linked Administrative and Census Bureau Data**

Thomas B. Foster, *U.S. Census Bureau*

Leticia Fernandez, *U.S. Census Bureau*

Sonya R. Porter, *U.S. Census Bureau*

Nikolas Pharris-Ciurej, *U.S. Census Bureau*

This research takes both a broad view of the first full year of the pandemic in the U.S. and, at the same time, examines granular variation in excess mortality. We link death records for 2010 through 2021 from the Social Security Administration, covering the universe of individuals ever issued a Social Security Number, to decennial census and American Community Survey race/ethnicity responses. We use these data to estimate excess all-cause mortality for age-, sex-, and race/ethnicity-specific subgroups and examine racial/ethnic variation in excess mortality across states and over the course of the pandemic. Results show that non-Hispanic American Indian/Alaskan Natives experienced the highest excess mortality of any racial/ethnic group in the first year of the pandemic, followed by Hispanics and non-Hispanic Blacks. Temporal-spatial trends suggest that the social and economic determinants known to drive health disparities in the pre-pandemic era were amplified in the pandemic's first year to disproportionately concentrate excess mortality among racial/ethnic minorities.

## **The Impacts of Rent Burden and Eviction on All-Cause Mortality in the United States**

Nick Graetz, *Princeton University*

Emily Lemmerman, *Princeton University*

Carl Gershenson, *Princeton University*

Danielle H. Sandler, *U.S. Census Bureau*

Sonya R. Porter, *U.S. Census Bureau*

Matthew Desmond, *Princeton University*

We link formal eviction records to all-cause mortality data from the Census Numident file and renter characteristics from the Census 2000, resulting in 38 million linked records from 2000-2016. We first describe substantial mortality disparities between homeowners and renters within race, gender, and cohort across the income distribution. We next document cumulative risk of eviction, which is highest for Black renters, especially for low-income Black women where we estimate 17-48% are evicted between ages 30 and 50. Across all race-gender-cohort groups, we estimate that eviction leads to a 5-67% increase in the risk of all-cause mortality (up to 67% for white men ages 30-50), finding acute and chronic impacts across the life course conditional on a large set of baseline renter characteristics. Based on counterfactual estimates where no evictions occur, we show that eviction explains between 10-46% of homeowner-renter mortality gaps depending on age, race, and gender (46% for white men ages 70-90). Across all race-cohort-gender groups, the impact of eviction on mortality is generally greatest for Black renters because the incidence of evictions is highest for these groups. We discuss implications for health and housing policy during the COVID-19 pandemic while contextualizing these debates within the broader, durable ecology of eviction and death that has been entrenched for decades.

**Tuesday, October 25**

**1:15PM**

---

### ***Session B-3: Considerations and Advancements in Sampling***

---

Organizer: Stephanie Coffey, *U.S. Census Bureau*

Chair: Stephanie Coffey, *U.S. Census Bureau*

Room 146C

## **Demographic Surveys Redesign Program: Evidence-based, Participatory Decision-making**

Richard Levy, *U.S. Census Bureau*

The Demographic Surveys Redesign Program is a research and development program managed by the US Census Bureau. The program's main purpose is to select updated samples for major demographic surveys based on shifting demography. These surveys provide detailed, continuously updated information on all aspects of American life. The program also sponsors research on topics ranging from measuring nonsampling error, developing customized frames, and assessing the business environment. The selection of new samples is coordinated for efficiency and to reduce respondent burden. Yet amidst this coordination, the selection process is customized to meet the specific requirements and methodologies of individual surveys. Major decisions require buy-in from multiple stakeholders. This presentation will discuss major decision points in the Redesign (such as the need to draw new first-stage samples in our multi-stage design), and how they were researched, documented, and vetted across diverse stakeholders. This could be informative to others who manage projects with key decision points requiring interagency cooperation and buy-in.

## **Bayesian Stratified Sampling for Establishment Surveys with Uncertain Design Parameters**

Jonathan Mendelson, *Bureau of Labor Statistics*

Michael Elliott, *University of Michigan*

For many survey sample designs, there are analytical results indicating the optimal sample allocations, but which assume knowledge of certain survey design parameters (e.g., strata variances). Practitioners commonly use these results by substituting survey-based estimates of design parameters in place of their true (unknown) values. Typically, little attention is given to the potential effects of using imperfect information at the sample design stage, even though this could harm sample efficiency. We consider sample allocation for a univariate regression model with heteroscedastic errors, using Bayesian decision theory to accommodate uncertain design parameters. By allowing for heteroscedasticity, our assumptions may be more realistic in an establishment context than some previous Bayesian design work, which did not. We compare performance of the proposed Bayesian sampling strategy with that of key design-based and model-assisted alternatives across several settings, finding that the proposed methods do as well or better than the alternatives under the scenarios considered. We apply our methods in analyzing revenues of public charities, using publicly available IRS Form 990 data.

## **Matrix Sampling in General Population Web Surveys**

Z. Tuba Suzer-Gurtekin, *University of Michigan*

Zhen Sun, *University of Michigan*

Yuxuan Chen, *University of Michigan*

Xinran Wang, *University of Michigan*

James Wagner, *University of Michigan*

Joanne Hsu, *University of Michigan*

Matrix sampling is considered as one of the methods that can reduce the survey length, which could potentially reduce nonresponse and measurement errors as a result lowered burden. While there are known applications in consumer satisfaction studies, it is not widely applied for the other official statistics as it increases the complexity of the questionnaire design and reduces the sample sizes for the estimates. To the best of our knowledge, there is no one fit for all for the design optimization in terms of block and sample randomization. In this study, we would be exploring matrix sampling properties in the context of consumer expectations for an Address Based Sample (ABS) web survey that has bi-monthly reporting deadlines. In the first phase of the study, the simulations with varying subsets of question blocks and sample sizes will be run. The simulation evaluations will include comparison of the quarterly change estimates against those from the parallel cell Randomized Digit Dialing (RDD) telephone surveys. The results from the simulation will be used to design the field experiments, as the simulations cannot model the possible context effects due to different questionnaire orders

---

**Session B-4: Pandemic and Related Fiscal Policies**

---

Organizer: Derek Wu, *University of Virginia*

Chair: William Gale, *Brookings Institution*

Discussant: Stan Veuger, *American Enterprise Institute*

Room 140A

**The Anti-Poverty, Targeting, and Labor Supply Effects of Replacing a Child Tax Credit with a Child Allowance**

Bruce D. Meyer, *University of Chicago*

Kevin Corinth, *University of Chicago*

Matthew Stadnicki, *University of Chicago*

Derek Wu, *University of Virginia*

The replacement of the Child Tax Credit (CTC) with a child allowance has been advocated by numerous policymakers and researchers. We estimate the anti-poverty, targeting, and labor supply effects of such a change by linking survey data with administrative tax and government program data which form part of the Comprehensive Income Dataset (CID). We focus on the provisions of the 2021 Build Back Better Act, which would have increased maximum benefit amounts to \$3,000 or \$3,600 per child (up from \$2,000 per child) and made the full credit available to all low and middle-income families regardless of earnings or income. Initially ignoring any behavioral responses, we estimate that the replacement of the CTC would reduce child poverty by 34% and deep child poverty by 39%. The change to a child allowance would have a larger anti-poverty effect on children than any existing government program, though at a higher cost per child raised above the poverty line than any other means-tested program. Relatedly, the child allowance would allocate a smaller share of its total dollars to families at the bottom of the income distribution—as well as families with the lowest levels of long-term income, education, or health—than any existing means-tested program with the exception of housing assistance. We then simulate anti-poverty effects accounting for labor supply responses. By replacing the CTC (which contained substantial work incentives akin to the EITC) with a child allowance, the policy change would reduce the return to working at all by at least \$2,000 per child for most workers with children. Relying on elasticity estimates consistent with mainstream simulation models and the academic literature, we estimate that this change in policy would lead 1.5 million workers (constituting 2.6% of all working parents) to exit the labor force. The decline in employment and the consequent earnings loss would mean that child poverty would only fall by at most 22% and deep child poverty would not fall at all with the policy change.

**The Effects of the Child Tax Credit on Labor Supply**

Kye Lippold, *Department of Treasury*

The Child Tax Credit (CTC) is a major earnings subsidy in the US tax and transfer system, but its effects have received little research attention compared to the Earned Income Tax Credit (EITC). I identify the effects of the CTC on extensive margin labor supply using a difference-in-discontinuities design, exploiting the fact that parents lose eligibility for the credit when a child turns 17 before the end of a tax year. Focusing on the credit's effects among lower-income households in Census surveys and administrative tax data, I find that loss of the credit leads to a fall in parental employment, with elasticities comparable to those found in studies of the EITC. These findings suggest tax credits continue to have substantial labor supply effects for low-income households in the post-welfare reform era.

**Unemployment Insurance in Survey and Administrative Data**

Jeff Larrimore, *Federal Reserve Board of Governors*

Jacob Mortenson, *Joint Committee on Taxation*

David Splinter, *Joint Committee on Taxation*

Unemployment Insurance (UI) benefits were a central part of the social safety net during the Covid-19

recession. UI benefits, however, are severely understated in surveys. Using administrative tax data, we find that over half of UI benefits were missed in major survey data, with a greater understatement among low-income workers. As a result, 2020 official poverty rates were overstated by about 2 percentage points, and corrected poverty reached a six-decade low. We provide data to correct underreporting in surveys and show that, compared to UI benefits, the UI exclusion tax expenditure was less targeted at low incomes.

## **Putting the Paycheck Protection Program into Perspective: An Analysis Using Administrative and Survey Data**

Michael Dalton, *Bureau of Labor Statistics*

After matching over 3 million loans from the \$669 billion Paycheck Protection Program to administrative wage records, I estimate a doubly robust dynamic difference-in-difference event study showing robust, causal impacts of the loans on employment, wages, and opening status of establishments 7 months after PPP approval. Doing back-of-the-envelope calculations, I find a range of \$20,000 to \$34,000 of PPP spent per employee-month retained, with about 24% of the PPP money going towards wage retention in the baseline model. The longer-term impact shows approximately a 10% reduction in permanent closures because of the first round of PPP money. Small and low-wage establishments show the largest impact from PPP.

**Tuesday, October 25**

**1:15PM**

---

### ***Session B-5: Information Quality Standards and Processes***

---

Organizer: Xuan Pham, *Risk Management Agency*

Chair: Carrie Jones, *Economic Research Service*

Discussant: Julie Parker, *Economic Research Service*

Room 140B

## **Evolution of Methodology and Quality Measures Reports at USDA-NASS**

Lindsay Drunasky, *National Agricultural Statistics Service*

Over the last several years, USDA's National Agricultural Statistics Service (NASS) has been striving to publicly document their survey methodology and publish select quality measures for their reports. Each Methodology and Quality Measures (M&QM) report is customized to the publication and follows the OMB Standards and Guidelines for Statistical Surveys. They are published on a recurring basis typically within 30 days of the corresponding report's release and contain such metrics as sample size, survey response rate, coefficient of variation (CV), and weighted item response rate. Development, creation, and challenges of the M&QM reports will be discussed.

## **USDA ERS: Data Product Quality Standards and Processes**

Katherine Ralston, *Economic Research Service*

USDA-Economic Research Service is a Principal Federal Statistical Agency whose function is the compilation, analysis, and dissemination of information encompassing a broad spectrum of agriculture, food, the environment, and rural development for statistical purposes. ERS embeds six quality standards, as defined by OMB and USDA directives, in publication of 80+ data products: purpose, utility, objectivity, transparency, integrity, and accessibility. The ERS Data Product Review Council (DPRC) oversees and implements these standards through biannual data product quality review rounds. In this session, a DPRC members will discuss ERS's data quality standards; agency's procedures and practices for conducting data quality reviews; and creating a culture of continuous improvements. Session will highlight specific examples of how ERS develop and maintain data quality excellence.

## Transparency and Data Quality

Dan Gillman, *Bureau of Labor Statistics*

The recent panel report on Transparency in Statistical Information for NCSES and All Federal Statistical Agencies defines transparency as the provision of sufficient documentation. Metadata are formalized documentation, so we can assess transparency through adherence to a metadata specification corresponding to the documentation needs for transparency in each situation. Data quality arises as a transparency concern itself, and it arises as a result of transparency in general – the provision of sufficient documentation. Quality data are interpretable, meaning they are sufficiently documented. But how good is the documentation? This is the question of the quality of metadata. Here, the use of metadata standards helps – with data and process interpretability, metadata quality, data quality, and documenting data quality. The Data Documentation Initiative (DDI) Lifecycle standard is designed to document the design, collection, analysis, and dissemination of statistical data, including data quality. In this talk, all these ideas are discussed and woven together.

**Tuesday, October 25**

**1:15PM**

---

### ***Session B-6: Leveraging Probability Panels for Federally Sponsored Statistical Data Collections***

---

Organizer: David Dutwin, *NORC at the University of Chicago*

Chair: Edward Mulrow, *NORC at the University of Chicago*

Room 144ABC

## **Transparency in Documentation for Statistical Surveys using a Probability Panel**

J. Michael Dennis, *NORC at the University of Chicago*

Jennifer Benz, *NORC at the University of Chicago*

Ed Mulrow, *NORC at the University of Chicago*

Commercially available survey panels have been criticized for having proprietary “black box” sampling, data collection, and statistical weighting processes and solutions and for incomplete disclosure of key metrics such as the AAPOR response rate obtained for each stage of the panel recruitment and retention process.

Proprietary panels are often silent or unclear about any sample lists, advertising campaigns, or social media used to identify potential panelists. The lack of appropriate disclosure and transparency, it can be argued, discourages the federal statistical community from more fully embracing such survey panels.

This paper will address descriptively the steps taken by NORC’s AmeriSpeak Panel since its founding in 2015 to provide transparency in methods and response rate reporting. The steps include, but are not limited to, detailed annual updates to publicly available documentation specifying the processes used in the sampling and recruitment of research subjects to the panel; complete AAPOR response rate reporting; documentation on within-panel sampling methodology; data editing and cleaning procedures; and statistical weighting protocols, as well as measures taken to protect the rights of human research subjects.

## **Case Study in the Use of a Probability Panel for Federally Sponsored Research**

Paul Scanlon, *National Center for Health Statistics*

The Research and Development Survey (RANDS), which began in 2015, is a series of cross-sectional surveys from probability-sampled commercial survey panels. RANDS has been used for methodological research at the National Center for Health Statistics (NCHS), including the use of close-ended probe questions and split-panel experiments for evaluating question-response patterns and for developing statistical methodology to calibrate survey estimates that leverage the strength of national survey data. Survey results have been used to evaluate estimation approaches for health outcomes from recruited survey panels, including propensity score adjustment and calibration. Additionally, the targeted embedded probe questions in RANDS have expanded findings from NCHS’ Collaborating Center for Questionnaire Design and Evaluation Research cognitive interviews to a wider sample.

This paper addresses the use of NORC's AmeriSpeak Panel as the sample source for recently conducted RANDS projects. The paper will cover the purposes of the research, the OMB review process, the use of phone mode to supplement web mode responses, transparency in documentation, and the use of supplemental non-probability panel samples in a special iteration of RANDS named "RANDS during COVID-19." Using AmeriSpeak sample, two rounds with a longitudinal design were conducted during summer 2020 (RANDS during COVID-19 Round 1 and RANDS during COVID-19 Round 2), and a third round from an independent sample was conducted during spring 2021 (RANDS during COVID-19 Round 3). The RANDS during COVID-19 surveys were used to publicly release a set of experimental estimates on selected topics, including loss of work due to illness with COVID-19, telemedicine access and use before and during the pandemic, and reduced access to specific types of health care. In addition to the RANDS during COVID-19 surveys, the specific topics covered in rounds of RANDS have included access to health care and utilization, chronic conditions, food security, general health, health insurance, opioid use, physical activity, psychological distress, smoking, and disability.

## **Strategies to Recruit and Represent Low Incidence Population Groups and Engender Sample Diversity in a Probability Panel**

David Dutwin, *NORC at the University of Chicago*  
Ipek Bilgen, *NORC at the University of Chicago*

Perhaps as never before, research has placed ever greater importance on highly representative samples of survey respondents. A diversity, equity and inclusion perspective in survey research has underscored the importance of not just interviewing traditionally under-represented subgroups, but representation of such groups adequate enough for the development of stable point estimates of, and within, a wide range of subpopulations. For it is not just that a study can adequately represent, for example, the Asian American (AAPI) population in the aggregate, but as well, to ensure that the disaggregated survey estimates are stable and valid for the many subgroups of the AAPI population (such as by heritage, language, income, education, and other metrics).

The AmeriSpeak probability panel has made wide-ranging efforts to recruit, maintain, and represent a range of populations of interest. This paper will detail: 1) the use of Big Data predictive models to recruit hard-to-reach populations, 2) the results of a major effort in 2021 to recruit Spanish-speaking respondents, 3) details on efforts to recruit a robust sample of AAPI, 4) efforts to retain key subpopulations with "1st 6th month engagement" protocols, 5) the use of the TrueNorth hybrid samples methodology and calibration to oversample hard-to-reach populations. Each of these are a key contributor to a major effort to ensure sample diversity in the AmeriSpeak Panel.

## **The Importance of Nonresponse Follow-up in Probability Panels**

Ipek Bilgen, *NORC at the University of Chicago*  
J. Michael Dennis, *NORC at the University of Chicago*  
David Dutwin, *NORC at the University of Chicago*  
Josh Lerner, *NORC at the University of Chicago*

Federal surveys for decades have rightly held high standards for low nonresponse. In today's age of low response rates, an effective strategy for maximizing a weighted response rate is a nonresponse follow-up (NRFU) protocol. The AmeriSpeak probability panel has utilized in-person NRFU to recruit approximately half of its ~50,000 panelists. This paper will explore the benefits of NRFU to the AmeriSpeak panel. First, we document and report how NRFU recruits hard-to-reach populations, particularly, those of low income, Spanish speakers, non-Whites, and persons with low educational attainment. Second, we explore the impact of NRFU on the distribution of attitudes, opinions, and behaviors reported in AmeriSpeak surveys, again documenting important changes that strongly suggest the value in data quality of conducting NRFU and recruiting respondents via NRFU. Third, we conceptualize "sample diversity" and explore the degree to which NRFU generates more diverse samples than more basic single-stage recruiting efforts. Overall we find NRFU to be not just a method to achieve response rates more acceptable to Federal survey research, but as well, a key method by which to increase sample representativeness, demographically, attitudinally, and behaviorally. Further, research subjects recruited as a result of the investment in NRFU tend to have more politically conservative policy positions on issues, as evidenced by NRFU respondents being more likely than initially

recruited respondents to report that they believe the country “spends too much” on improving and protecting the nation’s health, on Social Security, on protecting the environment, and other priorities.

**Tuesday, October 25**

**3:15PM**

---

***Session C-1: Methods for Refining and Supplementing Food Security Statistics and Importance to Policy Research***

---

Organizer: Alisha Coleman-Jensen, *Economic Research Service*

Chair: Jeffrey M. Gonzalez, *Bureau of Labor Statistics*

Discussant: Anna Vaudin, *Food and Nutrition Service*

Room 146A

**Analysis of the September 2020 Current Population Survey Food Security Supplement Split Panel Test**

Alisha Coleman-Jensen, *Economic Research Service*

Matthew P. Rabbitt, *Economic Research Service*

The Current Population Survey Food Security Supplement (CPS-FSS) survey questions had been largely unchanged since collection began in 1995. To ensure the continued relevance of the survey questions, a review the CPS-FSS instrument as it was currently implemented was conducted, changes were developed as needed, and cognitive testing of revisions to the survey instrument was completed. Based on the review and cognitive testing, some survey items were updated or modified, and some new questions were added. A one-time split panel test to assess the revised CPS-FSS survey instrument was conducted in 2020. The data from this one-time split panel test were used to assess the functioning of the updated questionnaire and the effect of modifying some survey items on key outcomes of interest, like the measure of food security. Based on the analysis comparing the findings from the standard and test instruments, the test instrument performed well. Observed differences in the test and standard instrument for food spending and community nutrition assistance were expected given changes in the survey items. The Rasch analysis confirms that the minor changes to the food security section are unlikely to impact the measurement of food insecurity, or impact comparability of estimates from year-to-year.

**Predicting “Real-Time” Food Hardship with Retail Food Spending: A New Approach Linking Scanner Data and the Census Household Pulse Survey**

Katie Fitzpatrick, *University of Delaware*

Kejda Llana, *University of Delaware*

Food insecurity is a vital national statistic that informs federal policy. The reliance on measuring food insecurity annually with the December Current Population Survey (CPS) data, however, creates challenges for understanding food hardship at the state or sub-state level, as well as responding to changes in food hardship quickly. This research uses “big data” from grocery store scanner data linked to measures of food insufficiency in the new Census Household Pulse data to determine if changes in total food expenditures, categories of expenditures, and/or specific goods are a leading indicator of food insufficiency. By creating a model of how changes in retail food purchases relate to changes in food insufficiency, it provides a new tool to understand food hardship that would complement the annual CPS estimates and be used by policymakers at all levels of government.

**Estimating Long-Run Trends in Food Security: A Structural Economic Approach**

Matthew P. Rabbitt, *Economic Research Service*

M. Taylor Rhodes, *Oregon State University*

We develop a new measure of household food hardship, structural food security (SFS), to study the economics of food security in the United States. SFS is an economically based measure of food hardship that maps a household’s income, subjectively assessed subsistence food needs, and preferences to their food

security and can, in turn, be used to determine the probability of a household experiencing food hardship equivalent to any of the U.S. Department of Agriculture's food security status classification categories (e.g., food insecurity). Initially, we use data from the 1995-2019 Current Population Survey Food Security Supplement to calibrate the economic model, determine the validity of the SFS measure, and assign each household a probability of experiencing food insecurity. After, we use the calibrated model along with data from the 1968-2019 Panel Study of Income Dynamics (PSID) to examine long-run trends in food insecurity based on the SFS measure.

## **The Effect of Pandemic EBT on Measures of Food Hardship**

Lauren Bauer, *Brookings Institution*

Krista Ruffini, *Georgetown University*

Diane Whitmore Schanzenbach, *Northwestern University*

The COVID-19 pandemic led to widespread school closures in the 2019-20 and 2020-21 school years, resulting in the loss of prepared school meals for millions of children. The Pandemic EBT program accounted for this forgone nutritional assistance by providing eligible families with a grocery voucher, similar to the Supplemental Nutrition Assistance Program (SNAP), but program details and disbursement dates varied across states. This paper leverages cross-state variation in the rollout of Pandemic-EBT over the 2019-20 and 2020-21 school years matched to high-frequency data on food and economic hardship. We find that Pandemic EBT reduced multiple definitions of food hardship among low-income families by about 20 percent in the 2020-21 school year. These reductions are largest in states that had relatively high rates of school closures.

**Tuesday, October 25**

**3:15PM**

---

### ***Session C-2: Furthering Research through Data Linkages***

---

Organizer: Ann Miller, *Department of the Interior*

Chair: Ann Miller, *Department of the Interior*

Discussants: Brian Knop, *U.S. Census Bureau* and Michaela Dillon, *U.S. Census Bureau*  
Room 146B

## **Internal Migration in the U.S. During the COVID-19 Pandemic**

Thomas B. Foster, *U.S. Census Bureau*

Mark Ellis, *University of Washington*

Lee Fiorio, *University of Washington*

Did internal migration in the U.S. increase or decrease during the COVID-19 pandemic? Census survey estimates suggest a decline in migration, but narratives in the popular press suggest the opposite. We answer this question using linked administrative data from the U.S. Postal Service (USPS), Internal Revenue Service (IRS), and Social Security Administration (SSA) to describe changes in migration during the pandemic. Voluntary submissions to the USPS National Change of Address registry serve as the backbone of our data and analysis, providing monthly data on temporary and permanent domestic moves prior to and during the pandemic. To this backbone we add income information from IRS 1040s and demographic information from the SSA Numident file using anonymous individual identifiers assigned by the Census Bureau. We use these data to describe changes in interstate, intrastate, and intracounty migration in 2020 and 2021 relative to pre-pandemic period and disaggregate these trends by age, race/ethnicity, and income. The broader implications of our findings for the pandemic and continued recovery from it are discussed.

## **Race, Class, and Mobility in U. S. Marriage Markets**

Caroline Walker, *U.S. Census Bureau*

Ariel Binder, *U.S. Census Bureau*

Marta Murray-Close, *U.S. Census Bureau*

Jonathan Eggleston, *U.S. Census Bureau*

We link the universe of Internal Revenue Service (IRS) 1040 income tax forms to household survey data from the 2011-2019 American Community Survey to study the relationship between childhood family income and adult family structure. We consider differences by childhood family income percentile in the probability that adults observed between their late 20s and mid-30s are married or living with an unmarried partner. We improve upon the partnership linkages identified in the ACS by supplementing them with partnerships linkages available in the tax data. For those who partner, we also assess the degree of assortative matching on childhood family income. Preliminary findings reveal substantial heterogeneity in these relationships across racial-ethnic groups and at different parts of the childhood family income distribution. We test several common hypotheses proposed in the literature to explain these differences in matching patterns.

## **Who gets married and who breaks up? Tracking Relationship Transitions After the Introduction of Legal Same-Sex Marriage in the U.S. Using Linked Survey and Administrative Data**

Marta Murray-Close, *U.S. Census Bureau*

Brian Knop, *U.S. Census Bureau*

This research uses linked survey and administrative data to construct longitudinal relationship trajectories for a large national sample of same-sex and different-sex couples, to compare the relationship histories of these groups on the eve of marriage equality, and to study the response of same-sex couples to marriage access. We match married and cohabiting couples from the 2013 American Community Survey to their 1974-2019 income tax records to determine when couples first shared a household, if and when they married, and if and when they broke up. We focus on couples who lived in states where same-sex couples could not legally marry in 2013 and where they therefore gained access to both state and federal marriage recognition between 2013 and 2015. We find that, in 2013, the average same-sex couple in these states had been together for fewer years than the average different-sex married couple but for more years than the average different-sex cohabiting couple. By 2018, marriage and long-term cohabitation had proven equally popular among same-sex couples: 40 percent of female same-sex couples and 38 percent of male same-sex couples had married, while 32 percent and 42 percent remained cohabiting.

## **The Impact of Manufacturing Credentials on Earnings and the Probability of Employment**

Maggie R. Jones, *U.S. Census Bureau*

Vanessa Brown, *National Student Clearinghouse*

Gardner Carrick, *National Association of Manufacturers*

Nikolas Pharris-Ciurej, *U.S. Census Bureau*

John Voorheis, *U.S. Census Bureau*

Caroline Walker, *U.S. Census Bureau*

This paper examines the labor market returns to earning industry-certified credentials in the manufacturing sector. Specifically, we are interested in estimating the impact of a manufacturing credential on wages, probability of employment, and probability of employment specifically in the manufacturing sector post credential attainment. We link students who earned manufacturing credentials to their enrollment and completion records, and then further link them to their IRS tax records for earnings and employment (Form W2 and 1040) and to the American Community Survey and decennial census for demographic information. We present earnings trajectories for workers with credentials by type of credential, industry of employment, age, race and ethnicity, gender, and state. To obtain a more causal estimate of the impact of a credential on earnings, we implement a coarsened exact matching strategy to compare outcomes between otherwise similar people with and without a manufacturing credential. We find that the attainment of a manufacturing industry credential is associated with higher earnings and a higher likelihood of labor market participation

when we compare attainers to a group of non-attainers who are otherwise similar.

**Tuesday, October 25**

**3:15PM**

---

### ***Session C-3: Advances in Small Area Estimation***

---

Organizer: Yang Cheng, *National Agricultural Statistics Service*

Chair: Yang Cheng, *National Agricultural Statistics Service*

Room 146C

#### **A Model-based Allocation Approach for Estimation of Adult Competency in Small Domains**

Andreea Erciulescu, *Westat*

Weijia Ren, *Westat*

Jianzhu Li, *Westat*

Tom Krenzke, *Westat*

Leyla Mohadjer, *Westat*

Robert Fay, *Westat*

The Program for the International Assessment of Adult Competencies (PIAAC) is a multicycle international survey of adult skills and competencies sponsored by the Organization for Economic Cooperation and Development (OECD). For the U.S. PIAAC, survey data alone is not sufficient to produce reliable adult competency estimates in small domains defined as counties, states, or counties or states by age or education groups, with some domain-level sample sizes being as small as zero. Small area estimation model-based approaches that combine sparse survey data with rich data from auxiliary sources have been proven useful to improve the reliability of adult literacy and numeracy estimates for such domains. Hierarchical Bayes models were developed to produce small area estimation model-based period estimates of adult literacy and numeracy for counties, states, and states by age or education groups, using survey data collected in the first three rounds of U.S. PIAAC, during 2012, 2014, and 2017. Then, an allocation approach was developed to produce estimates of adult literacy and numeracy for counties by age or education groups. This allocation approach is presented here, along with selected results. Plans are to include the estimates in the PIAAC Skills Map, available on the U.S. Department of Education National Center for Education Statistics at the Institute of Education Sciences webpage.

#### **An Interactive Data Tool to Explore State and County Model-based Estimates of Adult Literacy and Numeracy Skills**

Tom Krenzke, *Westat*

Weijia Ren, *Westat*

Andreea Erciulescu, *Westat*

Leyla Mohadjer, *Westat*

The Program for the International Assessment of Adult Competencies (PIAAC) is a survey of adult skills sponsored by the Organization for Economic Cooperation and Development (OECD). The survey examines a range of basic skills such as literacy, numeracy, and digital problem solving, and assesses these adult skills in a consistent manner across participating countries. In the United States, PIAAC is administered by the National Center for Education Statistics (NCES) as a nationally representative survey of the U.S. adult population 16–74 years old. The U.S. PIAAC Skills Map is an interactive data tool that provides state- and county-level data on the literacy and numeracy proficiency of adults ages 16–74 in all 50 states, all 3,141 counties, and the District of Columbia. The tool reports estimates of adult skills as an average score on the PIAAC scale (0–500) and as the proportion of adults at the lowest levels of proficiency (at or below Level 1), medium proficiency (at Level 2), and the highest levels of proficiency (at or above Level 3). Users can explore state and county summary cards that investigate demographic data to provide context for the model-based estimates. These estimates were produced using small area estimation models and are based on combined

PIAAC data from 2012, 2014, and 2017 and data from the American Community Survey (ACS). Users can also compare a county to a state, a state to the nation, and counties (or states) to each other using statistical testing. The Skills Map will soon also include state and county-level model-based estimates for six age groups and four education groups. The presentation will cover some methodological aspects, a demonstration of the PIAAC Skills Map, and present parts of the associated guidance on using these estimates.

### **Model-based Estimates for Farm Labor Quantities**

Lu Chen, *National Agricultural Statistics Service*

Nathan B. Cruze, *NASA Langley Research Center*

Linda J. Young, *National Agricultural Statistics Service*

The United States Department of Agriculture's (USDA's) National Agricultural Statistics Service (NASS) conducts the Farm Labor Survey to produce estimates of numbers of workers, duration of workweek, and wage rates for all workers). In the traditional process, the survey estimates for each current surveyed reference week at different levels are reviewed by the Regional Field Offices (RFOs), headquarter (HQ) statisticians, and finally by the Agricultural Statistics Board (ASB). The ASB also considers the relationship between current year survey estimates and previous year official values. Alternatively, implementing small area models for integrating survey estimates with additional sources of information provides more reliable official estimates and quantifies the uncertainty associated with each type of estimate. In this paper, several hierarchical Bayesian sub-area level models are developed in support of different estimates of interest in the Farm Labor Survey. A 2020 case study illustrates the improvement of the direct survey estimates for areas with small sample sizes by using auxiliary information and by borrowing information across areas and sub-areas. The resulting framework provides a complete set of coherent estimates for all required geographic levels; these methods were incorporated into the official Farm Labor publication for the first time in 2020.

### **Small Area Beverage Price Indexes**

Abigail Okrent, *USDA*

Chen Zhen, *University of Georgia*

Lan Mu, *University of Georgia*

Gauri Datta, *University of Georgia*

Although a handful of studies have demonstrated that spatial price variation is an important factor in food policy analysis, limited public-use spatial price information is available to researchers. The objective of this research is development of a public-use panel food price database for Census places in the contiguous United States for 2008–2017. These price indexes are based on the rolling-window Gini-Elteto-Koves-Szulv (GEKS) formula, which allows the index to be transitive within the rolling window, and hence, largely drift free. Index calculations are based on the IRI retail scanner data on sales and prices of 450,000 food barcodes at over 40,000 food retailers. Because some of the information are only available at large marketing areas rather than Census places, we use store characteristics from the TDLinx database to disaggregate store sales at the marketing level into Census places. The analysis is first applied to over 30 beverages within the Food and Nutrient Database for Dietary Studies nonalcoholic beverage category (i.e., 92) as a test case and gradually be expanded to other product groups as information on food nutritional characteristics become available.

### **Developing a Data-Driven System for Identifying Vulnerable and Resilient Neighborhoods across the United States: Opportunities for Improvement through Small Area Estimation**

Katherine Ann Willyard, *U.S. Census Bureau*

Gabriel Amaro, *U.S. Census Bureau*

Due to policymaker efforts to bolster equity and community resilience to natural hazards, there is considerable federal interest in developing precise methods of identifying communities most vulnerable to a disaster. However, measures using publicly available data are limited. The objective of this paper is to explain issues with how common social vulnerability indices use Census data and how Census' new Community Resilience Estimates (CRE) overcomes these concerns. Using restricted 1-year American Community Survey (ACS) data combined with population estimates, we demonstrate how small area modeling can produce more

precise measures of the communities most at-risk to disasters. Using the 2019 CRE as a case study, we demonstrate how small area estimates of social vulnerability and community resilience can be used to compare the vulnerable population rate of different geographic areas, such as Census regions and divisions, urban/large rural/small rural/isolated areas, historically disenfranchised/not historically disenfranchised areas, communities facing persistent poverty/those that aren't, and toxic communities, and those that aren't. When comparing the relative error of 1-year ACS estimates to the relative error of CRE for 71,666 Census tracts across the United States in 2019 using a t-test with a 90% confidence interval, we find that small area modeling significantly decreases the error of estimates. Also, using a t-test with a 90% confident interval, we find the vulnerable population rate is higher in small rural and isolated areas, historically disenfranchised areas, communities facing persistent poverty, and toxic communities. Since CRE produces population estimates along with margins of error, it is the only index that can be used to determine if the vulnerable population rate between places is statistically different. CRE provides a more precise illustration of how at-risk every neighborhood in the United States is to the impacts of a disaster. In comparison to CRE, if other national indices were to be used to distribute resources, communities would have more difficulty planning because estimates are less stable and less reliable. In developing a response to bolster community resilience, decisionmakers should rely upon CRE.

**Tuesday, October 25**

**3:15PM**

---

### ***Session C-4: Strengthening the Autonomy of the Federal Statistical Agencies***

---

Organizer: Steve Pierson, *American Statistical Association*

Chair: Edward Wu, *American Statistical Association*

Discussant: Jonathan Auerbach, *George Mason University*

Room 140A

Federal statistical agencies have been subjected to increasing political intrusion over the past decade. Decisions over what statistics should be collected and through which operations have at times been reduced to partisan talking points for the sole purpose of political gain. Yet protecting the federal statistical system against such partisan intrusions is complicated by the legitimate role of the legislative and executive branches to set priorities for data collection and exercise oversight over statistical agency performance. In this session, panelists discuss the autonomy protections necessary to keep official statistics both objective and accountable. This session follows recent research by former leaders of the federal statistical system including Constance F. Citro, Katherine Smith Evans, J. Steven Landefeld, Jeri Mulrow, Thomas Petska, Charles J. Rothwell, John Thompson, James L. Woodworth, and others.

#### **Panelists:**

Constance Citro, *Committee on National Statistics*

Ron Jarmin, *U.S. Census Bureau*

Nancy Potok, *NAPx Consulting*

---

***Session C-5: Advancements in Sexual Orientation and Gender Identity Measurement***

---

Organizer: Jennifer Ortman, *U.S. Census Bureau*

Chair: Jennifer Ortman, *U.S. Census Bureau*

Room 140B

**Results of a Mixed-Method Study to Evaluate 2-Step Gender Identity Measures**

Valerie Ryan, *National Center for Health Statistics*

Kristen Miller, *National Center for Health Statistics*

Stephanie Willson, *National Center for Health Statistics*

This presentation covers a mixed-method study to investigate the performance of a 2-step gender identity measure whereby respondents are first asked sex assigned at birth followed by current gender identity. The study design consisted of a cognitive interviewing component in which patterns of error and interpretation were identified, followed by a 2x2 split-sample survey experiment. The experiment examined question order and two versions of the gender identity question. Additionally, the survey contained embedded follow-up questions (developed from cognitive interview findings) to identify cases of false positive and false negative error as well as reasoning for that error.

Results of the cognitive interview component suggest that, while there is a broad range of understanding as to what constitutes the concepts of ‘sex’ and ‘gender,’ some respondents see little, if any, difference. While no outright error was discovered, several interpretative patterns were identified that would explain causes of both false positive and false negative error among survey respondents. For the split sample survey experiment, results show both types of error occurring (regardless of version), with 22.9% of potential gender minorities being false positive cases. In covering the study’s key findings, the presentation will illustrate the impact of the identified response error on key health characteristics also included on the survey and point toward directions for improved gender identity measurement.

**Asking Questions About Sexual Orientation and Gender Identity on the Experimental Household Pulse Survey: Evidence from Initial Evaluations**

Thom File, *U.S. Census Bureau*

Zachary Scherer, *U.S. Census Bureau*

To better understand the pandemic experiences of sexual and gender minority (SGM) populations, the Household Pulse Survey (HPS) began asking a series of three sexual orientation and gender identity (SOGI) questions in July 2021, including survey items about sex assigned at birth, current gender identity, and sexual orientation. This presentation will summarize a series of evaluations designed to better understand SOGI question performance on the HPS. First, to inform future data dissemination strategies, we will evaluate the impact of removing allocated sex at birth responses from analyses of the transgender population. Second, we will explore whether certain groups of respondents understand the SOGI questions better than others by investigating the characteristics of respondents who “correct” their sex at birth and/or current gender identity answers when prompted by an automated “check” question. Third, we will assess whether asking about sexual orientation before gender identity (as opposed to after) has any measurable effect on reported SOGI outcomes or survey breakoffs. Finally, we will use crosstabulations and logistic regression models to create a demographic profile of non-respondents to SOGI questions and those with responses of “something else” or “I don’t know” to the sexual orientation question and “none of these” to the gender identity question.

## **How do you Describe Yourself in the Workplace? Asking Teachers about their Sexual Orientation and Gender Identity in a School Survey**

Maura Spiegelman, *National Center for Education Statistics*

Allison Zotti, *U.S. Census Bureau*

Shawna Cox, *U.S. Census Bureau*

Many federal surveys, including the Census Household Pulse Survey, National Crime Victimization Survey, National Health Interview Survey, and High School Longitudinal Study, ask respondents about their sexual orientation and/or their gender identity (SOGI). These questions allow respondents to accurately describe themselves, statistical agencies to fully describe their populations, and researchers to explore differential outcomes for these groups. Past research has found that respondents are comfortable with these questions when contacted at their home, but there is limited research to address whether respondents will answer these questions when contacted at their workplace. In 2022, the U.S. Department of Education's National Teacher and Principal Survey (NTPS) conducted a split-ballot test to determine whether public school teachers would be willing to answer SOGI questions when contacted at their schools. This presentation discusses the results of cognitive interviews conducted in 2018 and 2022 that asked school staff about question wording, their comfort answering these questions, and their colleagues' perceived comfort. For the split ballot test, we compare response rates and breakoff rates when SOGI questions are included or omitted. When SOGI questions are included, we examine item response rates against other demographic items, the performance of a web survey verification question when teachers report different answers for their gender and sex assigned at birth, and write-in responses and other respondent feedback.

## **Assessing Open-Ended Self-Reports of Sexual Orientation and Gender Identity: Is There Room For Improvement?**

David Richards, *National Center for Education Statistics*

Elise Christopher, *National Center for Education Statistics*

Maura Spiegelman, *National Center for Education Statistics*

Over many decades, the National Center for Education Statistics has collected data from multiple cohorts of postsecondary students, including demographic information such as gender. In recent sample surveys of postsecondary students, NCES employed items designed to better measure sexual orientation and gender identity. These "SOGI" items include categorical response options, but also allow sample members to provide open-ended responses. Though these items are validated by cognitive tests prior to full-scale data collection, they can be further validated by analysis of full-scale data, to better confirm how well existing categorical response options are capturing SOGI data. We analyze open-ended responses collected from three cohorts of postsecondary students. Responses are coded and then the frequency and quality of open-ended responses are analyzed to determine how often respondents are providing written responses and the usefulness of those responses. Results of these analyses have implications for future collections of SOGI data, particularly for the improvement of existing SOGI measures and whether different categorical response options should be provided to survey respondents.

---

**Session C-6: Lifting the Fog: Editing and Imputation**

---

Organizer: Tiandong Li, *Housing and Urban Development*

Chair: Tiandong Li, *Housing and Urban Development*

Room 144ABC

**Banff and Simputation: A Comparison Using BEA's Multinational Enterprise Surveys**

Larkin Terrie, *Bureau of Economic Analysis*

The Bureau of Economic Analysis (BEA) employs automated data editing and imputation systems to process a subset of its multinational enterprise (MNE) surveys. Until recently, all BEA auto-editing systems were based on Statistics Canada's Banff system for data editing and imputation, which runs in SAS. To enhance its flexibility, BEA has explored the feasibility of creating auto-editing systems using R packages created by researchers at Statistics Netherlands. Previous research has shown that Banff produces accurate imputations on BEA's MNE surveys, so a key question is whether these R packages' imputations are as accurate as Banff's imputations. Among these R packages, Simputation is responsible for almost all imputation functionality. This project uses simulations to assess the relative accuracy of the imputations produced by Banff and Simputation. The results indicate that Simputation is sufficiently accurate to make the Statistics Netherlands R packages a viable alternative to Banff. The results also suggest that Simputation might produce more accurate imputations for survey instruments that collect few data items, while Banff might be more accurate for more complex forms.

**Imputing Missing Ferry Boarding Counts for Estimating National Totals**

Young-Jun Kweon, *Bureau of Transportation Statistics*

Aubrey Nguyen, *Bureau of Transportation Statistics*

Clara Reschovsky, *Bureau of Transportation Statistics*

The National Census of Ferry Operators (NCFO) is conducted every two years by the Bureau of Transportation Statistics (BTS). Data collected on passengers, vehicles, and route miles are used by the Federal Highway Administration (FHWA) for ferry funding allocation. Data edit and rule-based imputation are performed to increase data quality and usability for other analyses. Annual boarding counts is one of the most important pieces of information and also the hardest to impute because of its variability and wide range of values. Although NCFO data with imputed boarding counts may not be appropriate for analysis at an individual ferry operation level, it would be adequate for estimating the national totals. To this end, BTS applies several imputation models on multiple simulated datasets, generates complete datasets with imputed values, calculates national boarding totals for each dataset, and selects the best model based on accuracy and reliability in estimating the national totals. The presentation will discuss about data edit and preparation, imputation models, model comparison, and imputation results.

**Automating the Data Edit Process for the National Census of Ferry Operators**

Aubrey Nguyen, *Bureau of Transportation Statistics*

Young-Jun Kweon, *Bureau of Transportation Statistics*

Clara Reschovsky, *Bureau of Transportation Statistics*

The National Census of Ferry Operators (NCFO) data is collected by the Bureau of Transportation Statistics (BTS) every 2 years. Previously, extensive amount of time was spent to process the data once the census was concluded. There was little automation involved due to the data's complex nature: erroneous and missing information and free-text responses. We facilitate the data extraction, processing, and cleaning by developing repeatable automated programs using R Markdown that allow users to reproduce documents and generate reports almost instantly.

10 R Markdown programs are created to process the NCFO data. A single click executes all the codes and generates a report with interactive features: table of contents, graphs, and tables. If reported information in the census differs from BTS's records, the report captures the discrepancies and exports to a separate file

notifying that additional research is required.

This automation streamlines the data edit process and significantly reduces staff time required to manually review the information. The repeatable procedures ensure consistency and reliability regardless of when or by whom the edits are performed.

## **The Effects of Editing and Imputation on Measured Plant-level and Aggregate Productivity Growth in a Panel of U.S. Manufacturing Plants.**

Kirk White, *U.S. Census Bureau*

Hang Kim, *University of Cincinnati*

Martin Rotemberg, *New York University*

We investigate the effects of editing and imputation in the U.S. Census Bureau's Annual Survey of Manufactures (ASM) on measures of aggregate and plant-level productivity growth. Previous research found that the Census Bureau's imputations in the ASM, particularly regression imputes, tend to decrease measured within-industry productivity dispersion (White, Reiter and Petrin 2018; Rotemberg and White 2021). We add longitudinal ratio-of-ratios edits to the simultaneous Bayesian edit-imputation method of Kim et al. (2015), applied to the 2009-2013 and 2014-2018 panels of the ASM. We estimate plant-level and aggregate productivity growth in the manufacturing sector in the raw reported data (i.e., before editing and imputation), the Census Bureau's edited-imputed data, and the Bayesian edited-imputed data.

## **Scrubbed Clean: Does Data Cleaning Improve the Quality of Analytic Models?**

Megan A. Hendrich, *Ipsos Public Affairs*

Randall K. Thomas, *Ipsos Public Affairs*

Frances M. Barlas, *Ipsos Public Affairs*

Many researchers believe that data cleaning leads to better data quality and often clean out participants exhibiting sub-optimal behaviors (e.g., speeding, response non-differentiation, or nonresponse), sometimes using aggressive cleaning criteria that removes up to 20% of their sample. Surprisingly, most research has failed to find that data cleaning reduces bias for point estimates (e.g., proportions, means). In this study, we were interested in assessing if data cleaning affects covariance, specifically in multiple regression models. In an online study with over 9,000 completes from three different sample sources (a probability-based sample and two opt-in samples), we examined regression coefficients for two different multiple regression models (political attitudes predicting party identification and demographics and life experiences predicting life satisfaction). We deleted cases in gradations from 2.5% up to 50% of the sample based on speed of completion, weighted each dataset, and then ran the regression analyses. We found that small to moderate amounts of data cleaning did not substantially affect the direction or degree of coefficients; however, some coefficients became more unstable at 30% deletion and higher. We urge caution in any data cleaning protocols that might eliminate a higher proportion of participants since this may actually increase bias in covariance.

---

## ***Abstract Listings for Wednesday, October 26***

---

- |                         |                  |
|-------------------------|------------------|
| • Plenary Session II    | 8:30 – 9:30 am   |
| • Concurrent Sessions D | 10:00 – 11:45 am |
| • Concurrent Sessions E | 1:15 – 3:00 pm   |
| • Concurrent Sessions F | 3:15 – 5:00 pm   |

**Wednesday, October 26**

**8:30AM**

---

***Plenary Session II***

---

Organizer: Linda J. Young, National Agricultural Statistics Service  
Room 146ABC

**Building Representative Miniatures out of Non-representative Big Data: An Interplay of Data Quantity, Quality, and Privacy**

Xiao-Li Meng, *Harvard University*

The magical power of learning reliably about the opinions of 300 million people by a poll of 3,000 comes from the assumption that the 3,000 opinions expressed would form a representative sample of the 300 million opinions. But any magic involves illusion. In the era of “big data”, the illusion is to either believe or hope that the data can be magically representative by themselves without careful design, or that since the data sizes are so large (e.g., 3 million), the issue of representativeness is only for those statisticians who cry wolf. Recent theoretical and empirical investigations and findings reveal, however, that the wolves not only are real but also are much larger than anyone could have imagined. For example, the loss of sample size due to lack of representativeness exceeds 99% in some large national COVID vaccination surveys, suggesting that one should favor much smaller representative samples over much larger unrepresentative data. The common practice has been to defer the challenge of dealing with data defect to the analysis stage, via estimation methods such as weighting, super-population modeling, double-robust estimations, etc. The unbalanced nature of the data quality-quantity trade-off suggests an alternative strategy, at least for large databases that are released mostly for research purposes. That is, it may be more cost effective that the same information, assumptions, and technical ingenuity utilized for these estimation methods is used to build a miniature as representative as possible of the target population, out of the large unrepresentative data and any other helpful data sources. In addition to the ease of analysis for the users (no more “weight, weight, don’t tell me ...”), releasing a miniature instead of the parental data sources can greatly reduce statistical disclosure risk, at least for the vast majority of individuals whose data are not released verbatim. To avoid any new illusion, the proposed miniature cannot lead to less bias than those estimation methods with the same set of assumptions. It also induces another trade-off: lack of efficiency for particular studies in exchange for broader applicability and validity. These issues are similar in type to those that result from using imputations to fix the “holes” in data before releasing them for public use or forcing the users to apply a same set of weights regardless of the goal of their studies. Clearly, larger studies than this miniature one will be needed to investigate their impact on the proposed strategy, and of course the feasibility of the proposal itself.

**Wednesday, October 26**

**10:00AM**

---

***Session D-1: Advancing Efficiency and Accuracy with Data Science Techniques***

---

Organizer: Ellen Galantucci, *Federal Maritime Commission*  
Chair: Ellen Galantucci, *Federal Maritime Commission*  
Room 146A

**Improving Efficiency in Non-Data-Collection Activities with Data Science Tools: The Case of Census's Count Question Resolution Operation**

Nelson Chung, *U.S. Census Bureau*

Big data is often associated with making sense of large troves of information. But the tools of data science can be employed to acquire efficiency gains from automating tasks in Census projects where no new data collection is involved. This concept is illustrated through the use of Python in the U.S. Census Bureau's Count Question Resolution program, an operation that allows tribal, state, and local governments to challenge Census housing counts and boundary designations. Python is used to substantially reduce the number of

steps in leveraging pre-existent data, making sense of and publishing internal analytics, and tracking workflows. Enabling Python to talk to Oracle and Excel, to scrape web data, to expeditiously scanning entire directories, along with analytics best practices, are discussed.

## **Toward a Semi-automated Item Nonresponse Detector Model for Open-response Data**

Kristen Cibelli Hibben, *National Center for Health Statistics*

Zachary Smith, *National Center for Health Statistics*

Travis Hoppe, *National Center for Health Statistics*

Valerie Ryan, *National Center for Health Statistics*

Ben Rogers, *National Center for Health Statistics*

Paul Scanlon, *National Center for Health Statistics*

Kristen Miller, *National Center for Health Statistics*

Open-ended survey questions or web probes can be valuable because they allow respondents to provide additional information without the constraints of closed-ended options. However, open-text responses are more prone to item nonresponse, as well as inadequate and irrelevant responses. Further, time and cost factors associated with processing large sets of qualitative data often hinder the use of open-text data. We developed an item nonresponse detection model drawing on recent technological advancements in combination with targeted human-coding. The model is based on a Bidirectional Transformer for Language Understanding model, fine-tuned using Simple Contrastive Sentence Embedding. This approach is powerful because it uses natural language processing as opposed to previous nonresponse detection approaches that have relied exclusively on rules or regular expressions. Using open-text responses from several web probes as case studies, we present our process of refining and evaluating the model, including sensitivity and specificity calculations to quantify model performance when compared with human review. Prospects for a generalized version for public release are also discussed. Data are from the multi-round Research and Development Survey During COVID-19 developed by the National Center for Health Statistics, which used probability- and nonprobability-based panels.

## **“Why do you say that?” A Case Study in Applying Topic Modeling to Analyze Open-ended Survey Probes**

Benjamin Rogers, *National Center for Health Statistics*

Valerie Ryan, *National Center for Health Statistics*

Kristen Cibelli Hibben, *National Center for Health Statistics*

Open-ended survey questions and probes can be used to examine validity and measurement error; however, including open ended questions and probes may increase respondent burden and resources needed for analyzing responses. Traditionally open-ended response data is coded by hand, which can be time-consuming and labor-intensive. One way to potentially reduce the burden of hand-coding is to use natural language processing techniques, such as topic modeling, to assign topics to responses after filtering out responses that may not be meaningful. While traditional topic modeling methods are effective with large sections of carefully edited text to discover themes, they perform less well on open-text survey response data.

In 2020, the NCHS Research and Development Survey (RANDS) embedded a series of open-ended probes to assess respondents' understanding of terms being used to measure the impact of COVID-19. The resulting data file included 18,662 respondents with responses for up to 9 open-ended probes per respondent. This presentation compares the utility of using Latent Dirichlet Allocation, (LDA), to a Top2Vec topic model run using Google's Universal Sentence Encoder to classify open-text responses from two web probes. Results from both topic models are compared to results using human coding and subject matter expert review.

## Using Language Models to Extract Data from Unstructured Text

Thomas Neil Kattampallil, *University of Virginia*

Aritra Halder, *University of Virginia*

Stephanie Shipp, *University of Virginia*

Gary Anderson, *National Center for Science and Engineering Statistics*

John Jankowski, *National Center for Science and Engineering Statistics*

Billions of gigabytes of data are created globally everyday, and a significant portion of it is unstructured text data, which is difficult to process or use within traditional Relational Database systems. Several organizations collect large amounts of this 'dark data', data that is collected by an organization which is not used for analytics or insight generation, and is usually stored only for compliance purposes. Natural Language Processing methods and Language models can be applied to these unstructured data sets, in order to extract useful information. In our example, we use these methods to read through news articles to extract the company names and product names mentioned in the articles, and the process of how this can be applied in other areas is also discussed.

## Leveraging Natural Language Processing for Efficiencies without Increasing Nonresponse Bias

Sarah Grady, *Energy Information Administration*

Francisco Cifuentes, *Energy Information Administration*

S. Grace Deng, *Energy Information Administration*

Katie Lewis, *Energy Information Administration*

Surveys are increasingly using machine learning techniques like natural language processing (NLP) to increase efficiency of operations. The frame for the Residential Energy Consumption Survey (RECS) Energy Supplier Survey (ESS) is constructed using the names of energy suppliers (e.g., electric utilities) reported by respondents in the preceding RECS Household Survey. Historically, human coders manually compared name variations against a reference list of known suppliers. In the 2020 RECS, an algorithm was developed using NLP to increase efficiency in the final supplier assignment process. Staff reviewed matching output, yielding a 50 percent match rate. This presentation will explore how match rates differ across RECS-respondent geographic and demographic characteristics. Certain characteristics may be underrepresented, and as such, some cases may contribute more to the representativeness of the final ESS sample. Given time and other resource constraints, we explore the tradeoff between a case's expected ease of matching and contribution to sample representativeness. We provide recommendations for prioritizing cases and planning efficient operations that ameliorate rather than exacerbate nonresponse bias.

**Wednesday, October 26**

**10:00AM**

---

### ***Session D-2: Dying for Better Data: Using Administrative Records to Improve the Quality and Utility of Health Survey Data***

---

Organizer: Scott Boggess, *U.S. Census Bureau*

Chair: Scott Boggess, *U.S. Census Bureau*

Discussant: Paul Beatty, *U.S. Census Bureau*

Room 146B

## Assessing Data Quality in Population-Based Surveys Linked to End-Stage Renal Disease Administrative Data

Jonathan Aram, *National Center for Health Statistics*

Crescent B. Martin, *National Center for Health Statistics*

Lisa B. Mirel, *National Center for Health Statistics*

Integrating health data creates new resources that can be used to answer key health and policy related questions. Before such analyses are performed, it is important to understand the quality of the linked data.

The National Center for Health Statistics recently updated the linkage of two population-based surveys, the National Health Interview Survey (NHIS) and the National Health and Nutrition Examination Survey (NHANES), to End-Stage Renal Disease (ESRD) data from the United States Renal Data System (USRDS). Even though ESRD patients account for less than 1% of the total Medicare population, the condition has accounted for 7% of Medicare fee-for-service spending in recent years, making it an important topic in health research. The survey data provide information about topics such as access to health care, health insurance, obesity, nutrition, and hypertension, among others. Linkage to USRDS ESRD data enables investigations of the associations of these and other exposures with developing ESRD. In this analysis we will present quality metrics, including comparisons of linked data estimates and national estimates from the USRDS Annual Report. These results will be presented by survey and age groups and then compared to the totals in the ESRD registry for similar time periods and ages. Last, there will be a discussion of potential analyses using these and other linked survey data.

## **Linking the Medical Expenditure Panel Survey-Insurance Component to Administrative Records: A New Approach to Inferring Establishment Links**

Thomas Hegland, *Agency for Healthcare Research and Quality*

Ed Miller, *Agency for Healthcare Research and Quality*

Alice Zawacki, *U.S. Census Bureau*

This research introduces a new data linking approach which we used to link Medical Expenditure Panel Survey-Insurance Component (MEPS-IC) data on business establishments (i.e., particular work locations or worksites) to Decennial Census data and administrative tax records on those establishments' employees. The key challenge to this linkage is that while Internal Revenue Service W2 data does identify which business employs each worker, it does not necessarily identify the particular business establishment at which each employee works. Our approach to linking these data sources begins with matching establishments' parent firms to their workforces using identifiers appearing in tax records. Our linkage approach then seeks to match establishments to their own workforces by identifying the subset of their parent firm's workforce that best matches the expected size, total payroll, and residential geographic distribution of the establishment's workforce. This optimal subset search is a complex, multistep process, and requires different algorithms depending on the number of establishments from the same parent firm that must be simultaneously linked to their workforces. In this research, we also present statistics characterizing the match rate obtained using this approach and compare data from the newly linked dataset to benchmarks. Key results include that match rates are consistently high (exceeding 90%) across nearly all data subgroups, including subgroups where different matching algorithms were employed, and that the matched data exhibit a reasonable distribution of employment, payroll, and worker commute distances relative to expectations and external benchmarks. Notably, employment measures derived from tax records, but not used in the match itself, correspond with high fidelity to the employment levels that establishments report in the MEPS-IC.

## **Race and Ethnicity Modeling Applied to Linked Health Data**

Lisa B. Mirel, *National Center for Health Statistics*

Jessie L. Parker, *National Center for Health Statistics*

Dean Resnick, *NORC at the University of Chicago*

Christine Cox, *NORC at the University of Chicago*

Linking data creates new resources that cannot be obtained from a single source alone, including augmenting data with key health equity variables. The National Center for Health Statistics' (NCHS) Data Linkage Program links survey and administrative data to expand the analytic utility of its surveys by combining data from multiple sources. A recent linkage of NCHS' National Hospital Care Survey (NHCS) to the National Death Index (NDI) highlights an example where a linked administrative source augmented race and ethnicity information that was frequently missing from NHCS patient encounter records. However, since race and ethnicity data were only available for patients who died, critically important data gaps remained. To expand the research capabilities of the NHCS, along with its linked data resources, we modeled missing NHCS patient race and ethnicity data using name and geographic race and ethnicity frequencies from 2010 Census data. This presentation will describe the imputation methodology and the validation methods used to assess imputed race and ethnicity and will demonstrate the power of this imputation to further health equity research goals.

## **Developing a Physician-Employer Linked Dataset**

Dennis Linders, *U.S. Census Bureau*

Victoria Udalova, *U.S. Census Bureau*

Alice Zawacki, *U.S. Census Bureau*

Kate Vavra-Musser, *University of Southern California*

The U.S. Census Bureau's Enhancing Health Data (EHealth) research team is developing a new dataset covering the universe of U.S. physicians and their places of employment. These new data will deepen our understanding of the supply side of U.S. healthcare beyond existing sources by identifying the increasingly complex organizational relationships among physicians and their employers and by adding physician and business characteristics unavailable elsewhere, including information on self-employed physicians. This project relies on unique data available at the Census Bureau, such as comprehensive coverage of workers, tax filers, and employers that enables linkages across data sources at the individual and employer tax unit levels. We are able to enhance existing data on physicians by attaching demographic (such as race/ethnicity), medical specialty, geography, income, and other characteristics and we are able to enhance data on their employers by attaching business characteristics such as payroll, employment, ownership, and industry/organizational classifications. These data will support research in examining hierarchies of business relationships (including common ownerships of health systems); tracking physician characteristics over time (physician retirements, deaths, migration, employment and self-employment trends); developing new survey frames of physicians and group practices; and independently verifying and complementing existing datasets on the structure of the U.S. healthcare delivery system.

**Wednesday, October 26**

**10:00AM**

---

### ***Session D-3: Evaluation of EEOC Collection of Pay Data: Quality for Intended Use and Recommendations for the Future***

---

Organizer: Jennifer Park, *Committee on National Statistics*

Chair: William Rodgers III, *Federal Reserve Bank of St. Louis*

Discussant: Meghan Maury, *White House Office of Science and Technology Policy*

Room 146C

#### **Evaluation of Pay Data in EEO-1 Component 2**

Claudia Goldin, *Harvard University*

Siwei Chang, *New York University*

Assessment of pay concepts, measures, findings and recommendations will be provided.

#### **Measures of Sex, Race/Ethnicity, and other Protected Groups in EEO-1 Component 2**

Lee Badgett, *University of Massachusetts, Amherst*

Valerie Wilson, *EPI*

Authorities, measures, findings and recommendations will be presented.

#### **Use of EEO-1 Component 2 Data for Initial EEOC Investigations**

Elizabeth Hirsh, *University of British Columbia*

Kristen Olson, *University of Nebraska*

Don Tomaskovic-Devey, *University of Massachusetts, Amherst*

Procedure, assessment, findings and recommendations will be presented.

## **Use of EEO-1 Component 2 Data for Employer Self-Assessment and National Comparisons**

Judy Hellerstein, *University of Maryland*  
Lisette Garcia, *Pennsylvania State*

Assessment, approach, findings, and recommendations will be provided.

**Wednesday, October 26**

**10:00AM**

---

### ***Session D-4: Statistical Issues in Research to Inform Evidence-based Nutrition Security Policymaking***

---

Organizer: Andrea C. Carlson, *Economic Research Service*

Chair: Elina T. Page, *Economic Research Service*

Discussant: Sara Bleich, *Food and Nutrition Service*

Room 140A

#### **Data Required to Inform Nutrition Security Policy**

Kevin Meyers Mathieu, *Food and Nutrition Service*

Policymaking which supports nutrition security, such as the reevaluation of the Thrifty Food Plan, relies on the maintenance and continued improvement of a robust and interconnected food data system. To develop and implement evidence-based policy that supports nutrition security, USDA relies on nationally representative dietary intake data, and corresponding data on food composition and prices. This presentation will include a discussion of the data which supported the re-evaluation of the Thrifty Food Plan, the characteristics of these data which made the re-evaluation possible, and opportunities for the food data system to further support nutrition security policymaking.

#### **Capturing Nutrition Information for Food Items Reported in FoodAPS-2: Survey Methodological Enhancements for Evidence-based Policymaking and Research**

Linda Kantor, *Economic Research Service*

The National Food Acquisition and Purchase Survey (FoodAPS) collects comprehensive data about household food acquisitions. These data are then linked to USDA and proprietary databases containing information on food item level nutrition attributes. These linkages yield a richer set of outputs for evidence-based policymaking and research on nutrition security than FoodAPS could alone. As survey and data development efforts for the second round of FoodAPS (FoodAPS-2) are currently underway, this presentation will give an overview of the methodological enhancements being evaluated for FoodAPS-2 and discusses the implications of each for evidence-based policymaking and research on nutrition security.

#### **Statistical properties of the Purchase-to-Plate Crosswalk: A comparison of Healthy Eating Index scores across Federal surveys**

Andrea C Carlson, *Economic Research Service*

The Purchase to Plate Crosswalk (PPC), a part of the Purchase to Plate Suite, allows users to measure how well food purchases recorded in food scanner data align with the recommendations in the Dietary Guidelines for Americans. This alignment is measured using the Healthy Eating Index (HEI), originally developed by USDA and the National Cancer Institute to use with dietary recall data and has been applied to other types of data. The PPC is now available for multiple years of scanner data, but validation has only been done for individual years. In this presentation we will compare HEI scores estimated from the food scanner data to other estimates, such as dietary recall data, in order to validate using the PPC over time to conduct policy-oriented research.

## **Comparison of Purchase to Plate National Average Prices for NHANES to Other Federal Food Expenditure and Price Index Measures**

Christopher Lowe, *Economic Research Service*

The Purchase to Plate National Average Prices for NHANES (PP-NAP) builds on the Purchase to Plate Crosswalk (PPC) to develop national average prices for foods reported ingested by participants in What We Eat in America, the dietary component of the National Health and Nutrition Examination Survey (WWEIA/NHANES). These prices can be used to conduct policy-oriented research using WWEIA/NHANES. In this presentation we explore the statistical properties of the PP-NAP. This analysis includes currently includes two parts: 1) Comparing expenditures of NHANES households (calculated using PP-NAP) to other expenditure data series; 2) Comparing price changes over time for specific categories of foods in PP-NAP to the Bureau of Labor Statistics Consumer Price Index data. The objective is to validate the PP-NAP to support future policy-oriented research using these data.

**Wednesday, October 26**

**10:00AM**

---

### ***Session D-5: Reimagining NCES and the Education Data Ecosystem for the 21st Century***

---

Organizers: Melissa Chiu, *Committee on National Statistics*

Marie Marcum, *National Center for Education Statistics*

Chair: Melissa Chiu, *Committee on National Statistics*

Discussant: Anthony Nerino, *Office of Management and Budget*

Room 140B

## **A Vision and Roadmap for Education Statistics: Overview of the CNSTAT Study**

Amy O'Hara, *Georgetown University*

Melissa Chiu, *Committee on National Statistics*

The education landscape in the United States has been changing rapidly in recent decades: student populations have become more diverse; there has been an explosion of data sources; there is an intensified focus on diversity, equity, inclusion, and accessibility; educators and policy makers at all levels want more and better data for evidence-based decision making; and the role of technology in education has increased dramatically. With awareness of this changed landscape the Department of Education asked the National Academies of Sciences, Engineering and Medicine to provide a vision for the National Center for Education Statistics (NCES)—the nation's premier statistical agency for collecting, analyzing, and disseminating statistics at all levels of education.

The National Academies study panel reimagines NCES as a leader in the 21st century education data ecosystem, where it can meet the growing demands for policy-relevant statistical analyses and data to more effectively and efficiently achieve its mission, especially in light of the Foundations for Evidence-Based Policymaking Act of 2018. Key to achieving a new vision is for NCES to act strategically in all aspects of the agency's work, including incentivizing innovation and creative partnerships, to assist education decision makers at every level of government. The presentation discusses these and other key priorities.

## **Embracing a New Vision for Education Statistics**

Peggy Carr, *National Center for Education Statistics*

More than three decades since the National Academies published its landmark report, "Creating a Center for Education Statistics: A Time for Action" (1986), NASEM has once again produced an insightful, comprehensive report outlining a contemporary roadmap for how a federal statistical agency should tackle education statistics in the 21st Century. As our nation's leader in Education Statistics, the National Center for Education Statistics (NCES) has welcomed and is now embracing the ambitious, bold recommendations provided in NASEM's "A Vision and Roadmap for Education Statistics." NCES has carefully considered the fifteen recommendations of the NASEM panel and has begun planning and implementation of a

transformation as inspired by the report's recommendations. For this segment, the NCES Commissioner will respond to each of the panel's recommendations and provide an update on plans to implement those recommendations. These plans include better coordination of survey operations, tracking, and sampling methods with the goal of reducing response burden for schools and local education agencies who respond to NCES surveys; developing the agency's capacity to support the Commissioner's role as the Statistical Official in response to the Evidence Act; and rethinking our quality management and review practices to create more timely and relevant statistical products.

## **Private Sector Roles in the Future Education Data Ecosystem**

Rachel Dinkes, *Knowledge Alliance*

The recent National Academies study, *A Vision and Roadmap for Education Statistics*, reimagines the National Center for Education Statistics (NCES) as a leader in the 21st-century data ecosystem. As an integral part of how NCES currently accomplishes its mission, contractors and external experts also play an important role in the Center's future. The study panel recommends that NCES "utilize contractors and creative staffing arrangements to work collaboratively with staff to build internal capacity" such as for innovation. The report also discusses aligning both NCES's workforce and contractors to other recommendations such as strategic goals and embedding diversity, equity, inclusion, and accessibility considerations throughout the work. The study panel further recommends leveraging the expertise of external experts by establishing a joint statistical research program and a nimble ongoing consulting body. This presentation discusses private sector perspectives on the study's recommendations, its partnerships with NCES, and its role in the future education data ecosystem.

**Wednesday, October 26**

**10:00AM**

---

### ***Session D-6: Evaluating Methods for Evaluating Questionnaires***

---

Organizer: Doug Williams, *Bureau of Labor Statistics*

Chair: Doug Williams, *Bureau of Labor Statistics*

Room 144ABC

## **Understanding the Interviewer in Cognitive Interviewing: Reflexivity and Improving Data Quality**

Zachary Smith, *National Center for Health Statistics*

Cognitive interviewing studies are used to improve question design and contextualize survey data. In contrast to the vibrant literature on survey interviewer effects, however, little systematic analysis of the role of the cognitive interviewer exists. To address this gap, this study draws on the qualitative research evaluation criterion of reflexivity, the critical examination of a researcher's own role and influence in the generation, analysis, and reporting of data. The use of reflexivity is intended to increase the data quality of cognitive interviewing reports by clearly denoting how subjective decisions and unconscious processes guide the co-construction of data between interviewer and respondent.

This study analyzes the potential gains from use of reflexivity in the context of cognitive testing of questions about cannabis use. It draws on two sources: written reflections by interviewers on their perceived effects on individual interviews ( $n = 90$ ) and in-depth interviews with each member of the interview team ( $n = 6$ ). The results include a framework for assessing the role of the cognitive interviewer and recommendations for integration into cognitive interviewing data analysis.

## **A Comparison of Questionnaire Evaluation Methods to Assess the Airborne Hazards and Open Burn Pit Registry**

Darby Steiger, *Westat*

April Fales, *Westat*

Many questionnaire evaluation methods are used to improve survey data quality. Expert reviews can identify question problems and other potential measurement errors (Olsen, 2010). Cognitive interviews explore the respondent's cognitive processing as they determine how to answer a question (Willis, 2004). Usability testing uncovers design issues that respondents may encounter when interacting with the instrument (Couper, 1999). Debriefing interviews collect respondent feedback after completing a questionnaire (Martin, 2006). An ongoing question is whether these methods produce similar findings (Rothgeb, Willis and Forsyth, 2007).

This paper will compare results from six questionnaire evaluation techniques to guide improvements to the Airborne Hazards and Open Burn Pit Registry, which monitors potential health effects from military exposure to airborne environmental hazards during the wars in Iraq and Afghanistan. Techniques included methodologist review, expert review, cognitive interviews, exploratory interviews, usability interviews, and debriefing interviews. Findings will focus on convergence and divergence of recommendations based on the different types of feedback.

## **Effects of and Reactions to Response Formats for Race and Ethnicity Measurement in the U.S.**

Frances M. Barlas, *Ipsos Public Affairs*

Kip Davis, *Ipsos Public Affairs*

Randall Thomas, *Ipsos Public Affairs*

Megan Hendrich, *Ipsos Public Affairs*

Much has been made about the large change in the racial composition of the United States as indicated by the latest Decennial Census figures with many more Americans identifying as more than one race compared to 10 years ago. Typically, a question about Latino origin is asked prior to a question about racial identity. An often-discussed modification is to merge the two questions (race and Latino ethnicity) into a single question. In addition, another modification has proposed adding a Middle Eastern/North African category (MENA). We conducted a study testing 6 versions of race and ethnicity questions, randomly assigning one per respondent, to test combining race and Latino ethnicity as well as the provision of a MENA option. Over 7,000 respondents from Ipsos' KnowledgePanel, a probability-based online panel, completed this web-based experiment. We report on the similarities and differences we obtained between formats. We further asked respondents about their experiences answering questions about race and ethnicity generally and about the question format that they were asked in this study. Latino respondents across all question formats had the most difficulty in answering the question, particularly the more traditional two question format. High rates of respondents of color indicated that they had a difficult time answering the question, that they could not locate themselves in response lists, that they were concerned about data privacy, that they were concerned about answering the questions, and that they are offended when being asked such questions. Given the increasing diversity of the United States in our surveys and our interest in ensuring the most representative samples, it is critical that we consider revised, more inclusive, and more relevant race and ethnicity questions for our respondents.

## **Correlates of Item Nonresponse to Open-Ended Web Probes**

Zachary Smith, *National Center for Health Statistics*

Kristen Cibelli Hibben, *National Center for Health Statistics*

Travis Hoppe, *National Center for Health Statistics*

Valerie Ryan, *National Center for Health Statistics*

Ben Rogers, *National Center for Health Statistics*

Paul Scanlon, *National Center for Health Statistics*

Kristen Miller, *National Center for Health Statistics*

Open-ended web probes are often used in designing survey questions. While these allow respondents to provide unconstrained information, open-ended probes are also more prone to item nonresponse, and coding open-text data is labor-intensive. To address these challenges, we developed an item nonresponse detection model that, in contrast to existing rule-based approaches, draws on natural language processing to categorize inputs as valid, low quality, or nonresponse.

For the model's use in question design, however, understanding the correlates of item nonresponse is crucial. If those who provide nonresponses systematically differ from those who provide valid responses, questions designed using analysis of valid responses may perpetuate the exclusion of population and methodological subgroups. This paper examines the variance of sociodemographic and survey characteristics in probe non-response across and within four web probes on a variety of topics. It discusses the relationship between demographic and methodological subgroups and nonresponse and, further, compares the characteristics of responses incorrectly categorized as nonresponse with those categorized as valid.

## **Facilitating Innovation in Survey Systems with Human-Centered Design**

Jean E. Fox, *Bureau of Labor Statistics*

Scott Arden, *Bureau of Labor Statistics*

Tod Sirois, *Bureau of Labor Statistics*

Many survey organizations are applying human-centered design methods to maximize the usability of their collection instruments and analysis systems. In addition to these important benefits, human-centered design methods can also help organizations facilitate innovation. By applying methods such as interviews, focus groups, and usability tests before any development starts, organizations can identify users' actual needs, uncover their real pain points, and gather their ideas and recommendations for innovative improvements to systems and procedures. This information aligns team members with a common understanding of the users' needs, which then helps them to (1) prioritize the focus of their development work, (2) streamline the design process, and (3) incorporate creative new solutions they would not have identified on their own.

In this presentation, we will discuss how applying a human-centered design approach can foster innovation. We will also share methods that are most useful and show examples from two programs at BLS that have benefitted from this approach.

**Wednesday, October 26**

**1:15PM**

---

### ***Session E-1: Statistical Modeling***

---

Organizer: Joe Conklin, *Bureau of Justice Statistics*

Chair: Joe Conklin, *Bureau of Justice Statistics*

Discussant: Andreea Erciulescu, *Westat*

Room 146A

## **Applying a Model of Income Distribution, Adopted as Econophysics, to Estimating Non-Compliance with the Individual Federal Income Tax**

John Angle ORCID, *Tresoria, LLC*

The Inequality Process (IP) is a mathematical model of income and wealth distributions and other statistics, tested and empirically confirmed many times since 1983. The IP is a stochastic particle system. The IP has been adopted by econophysicists, physicists and other applied mathematicians dedicated to renovating

economics, to make economics more like a science. See Byrro Ribeiro's (2020) textbook "Income Distribution Dynamics of Economic Systems: an Econophysics Approach", Cambridge University Press, pp. 154-158, under the heading "The Angle Process", i.e. the IP. Prof. Byrro Ribeiro cites 29 IP papers and traces a whole econophysical literature the IP has inspired. The IP's parameters can be estimated from parts of an income distribution (e.g., near the median) where observations are more numerous and/or more credible, or just a few statistics. These parameter estimates imply expected relative frequencies in less accessible/trustworthy parts of the distribution, for example, those of large incomes. Estimating IP parameters and their implied IP distributions also creates baselines to highlight anomalous subsets of incomes. This presentation shows that an IP model can fit income and wealth distributions using the entire interval on which IP particle parameters are defined. The IP fits distributions of labor income (gammoidal distributions), the heavier-tailed income distributions from tangible assets (approximately variance-gamma distributed), and distributions of proprietor's income (in between). See Angle (2022) "Generalizing the Inequality Process' gamma model of particle wealth statistics" in The Journal of Mathematical Sociology. Baselining reported income amounts with the IP is less intrusive than auditing samples of taxpayers' returns in the Taxpayer Compliance Measurement Program (TCMP). The IP was discovered by searching for the simplest mathematical model that accounts for the widest scope of patterns in statistics of income and wealth, i.e., the method of science. Comparative baselining of individual reported incomes on tax returns might be a way of raising greater tax revenue without increasing nominal tax rates. The cost of trying the model out that way on existing data sets of reported incomes is next to nothing for the U.S. Treasury.

## **A Bayesian Model of the Probabilistic Role of Weather Variations on Crop Yield Potential**

Noé J Nava, *Economics Research Service*

I develop a Bayesian hierarchical model of crop yields that integrates agronomic evidence of how weather variations affect crops' growth that is parametrized to study unit-level production trends. My formulation is employed to estimate posterior distributions relevant to corn and soybean production in the U.S. I simulate production counterfactuals that employ weather climate scenarios from the most current Intergovernmental Panel of Climate Change report. I present crop yield projections from U.S. Eastern counties for 2022, 2027 and 2032. Results indicate that farmers in Midwestern counties will likely shift soybean production for corn since soybean yield potential will decline among while corn yield potential will increase among these counties. Overall productivity trends indicate that the U.S. will continue its upward trend for the next decade, corroborating existing evidence regarding U.S. agricultural outlooks.

## **Methods and Assumptions of the CPS ASEC Tax Model and Imputing 2020 Stimulus Payments**

Daniel Lin, *U.S. Census Bureau*

This presentation will describe the methods of the Census Bureau's CPS ASEC Tax Model and discuss the imputation of the 2020 Economic Impact Payments using the Tax Model.

The Census Bureau's CPS ASEC Tax Model imputes federal and state taxes and credits for the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) and for the Supplemental Poverty Measure (SPM) on an annual basis. The CPS ASEC Tax Model began in the early 1980s and had a major methodological update in 2004. This paper provides useful documentation for users of the CPS ASEC Tax model. It documents the methodology and assumptions of the model and describes in detail the formation of tax units, construction of income, statistical match to IRS microdata, iterative calculations of federal and state taxes and credits, requirements for filing, and a description of the output variables. The presentation discusses the limitations of the CPS ASEC Tax Model and presents aggregate statistics for comparisons with tabulations from the Internal Revenue Service.

This presentation will conclude by giving an application of the tax model by showing the estimation of the first two rounds of the stimulus payments received by households in 2020. It will present the model used in the 2021 CPS ASEC and two additional models with different underlying assumptions that affect take-up rates and discuss their impacts on SPM poverty rates.

## **Place-Based Tax Incentives and Minority Employment: Evidence from New Market Tax Credit (NMTC) Program**

Anil Rupasingha, *Economic Research Service*

Alexander Marre, *Federal Reserve Bank of Cleveland*

This paper studies impacts of the New Market Tax Credit (NMTC) program between 2010-2019 time-period on minority employment. We also investigate whether the minority jobs that can be attributed to the program stayed in program recipient neighborhoods or whether those jobs were occupied by non-residents. The NMTC program was implemented by the Community Development Financial Institution (CDFI) Fund of the U.S. Department of the Treasury and was designed to stimulate economic development in distressed urban and rural communities by encouraging private investment. We combine NMTC program data from the CDFI Fund with demographic data from the American Community Survey (ACS) and employment data from the Census Longitudinal Employer-Household Dynamics Local Origin Destination Employment Statistics (LODES). A quasi-experimental research design that combines matching with regression analysis is employed to identify the causal effect of the NMTC investments on employment outcomes based on race and ethnicity and place-based and residence-based employment. The statistically significant results show that there is clear evidence that NMTC investments increased both place-based and residence-based employment in treated tracts except residence-based employment for Hispanics. The results further show that while NMTC program increased the number of jobs being available in a program recipient tract (workplace employment), those individuals who live out-side the tract are holding many of these jobs available due to the program.

## **Skill Profiles and Portability of Credentials for the Technical Workforce**

Cesar Montalvo, *University of Virginia*

Haleigh Tomlin, *Washington and Lee University*

Leonel Siwe, *University of Virginia*

Vicki Lancaster, *University of Virginia*

The National Science Board estimated that the US economy will have 3.4 million unfilled skilled technical workforce jobs by 2022. Technical jobs are essential to maintain a competitive economy and, also provide an opportunity for individuals to enter the middle-class. This research seeks to study how technical skill prevalence shapes the salary returns and the connection between credentials and career pathways for skill-intensive occupations. The analysis builds a model where salaries are explained by the proportion of technical skills, and experience across occupations. The model uses Burning Glass Technology information for the State of Virginia and focuses on five major occupation groups including Construction, Maintenance, Production, Transportation and Healthcare. We determined the skill “genome” of occupational groups and construct a metric for skill importance to classify skills according to technical component. We provide returns for clusters of technical skills using variable selection through Monte Carlo method for model validation. In addition, we used network analysis to establish a metric for credential portability. The identification of returns for technical skills within occupations will help to inform individuals about career pathways that may lead to technical occupations, through reskilling and credential seeking.

---

**Session E-2: Mosaic Data Linkage**

---

Organizer: Xuan Pham, *Economic Research Service*Chair: Jeffrey Gonzalez, *Bureau of Labor Statistics*

Room 146B

**Estimating Foreign-Born Emigration Using Longitudinal Tax and Social Security Data**Thomas B. Foster, *U.S. Census Bureau*Anthony Knapp, *U.S. Census Bureau*Erik Vickstrom, *U.S. Census Bureau*

Emigration from the United States is an important component of population change, but measuring it directly is difficult and current methods of estimation rely on dated and often unreliable data sources. This research takes the initial steps towards an estimate of foreign-born emigration derived from more recent linked individual-level administrative records from the Internal Revenue Service and the Social Security Administration. We use these linked data to estimate foreign-born emigration rates for the 2010 to 2020 period at the national level, as well as by age, state, tenure in the U.S., and select countries of birth, using a traditional cohort-component accounting method. Preliminary results are promising and show that emigration estimates derived from linked IRS-SSA records are comparable to those currently produced by the U.S. Census Bureau's Population Estimates Program, both overall and with respect to the distribution of emigrants across age, state, and tenure in the U.S. categories. Directions for improvement and future research are discussed.

**Mobility of U.S.-trained Early Career Research Scientists and Engineers**Christina Freyman, *National Science Foundation*Wan-Ying Chang, *National Science Foundation*

A novel approach combining machine learning and author disambiguation techniques was used to link respondents of the Survey of Doctorate Recipients (SDR) to the publication database Web of Science (WoS). The survey data provides information on graduate field, doctoral institution, and employment characteristics to enable comparative analyses among key subgroups. The author affiliation and citation topic information contained in the metadata of publications provide unique longitudinal data to explore the mobility of Ph.D. scientists and engineers throughout their publication careers. Using this data set, we analyzed geographic mobility of a cohort of early career STEM doctorate graduates from 2009 – 2011. The cohort is classified into analysis groups by an indicator of doctoral institution located in jurisdictions targeted for capacity-building federal investment and whether they stayed or moved for post-graduation employment as indicated in publication affiliation data or survey responses. The level and impact of scientific productivity in terms of publications by analysis group is examined using descriptive summaries and regressions with potential factors such as doctorate field of study, the 2005 Carnegie classification of the doctoral institution, post-graduation employment; sex, race and ethnicity, and other background information accounted for.

**Exploration of NAEP-SSOCS linkage: Investigating the Association between Behavioral Incidents, School Behavioral Climate, and NAEP Mathematics Achievement**Saki Ikoma, *American Institutes for Research*Paul Bailey, *American Institutes for Research*Jizhi Zhang, *American Institutes for Research*Kelly Guinn, *ICF*Markus Broer, *American Institutes for Research*

Research on educational inequality has been rigorously studied using National Assessment of Educational Progress (NAEP); data linkage between NAEP and another survey by the National Center for Education Statistics (NCES) provides a further opportunity to study inequality. Data linkage of any given NCES survey to

another/non-NCES data requires caution since NCES surveys are based on complex survey designs with plausible values (PVs)—multiple imputations of a latent trait—to estimate summaries of student scores. The purpose of this study is to report benefits and limitations of linking NAEP and School Survey on Crime and Safety (SSOCS). We link NAEP and SSOCS and use the linked data to examine associations between school-level crime, violence, and behavioral indicators and NAEP student achievement. To avoid underestimating the associations of NAEP student achievement and external variables, we used the R EdSurvey and Dire packages to draw PVs conditioned with the SSOCS variables for further analyses. Results focus on comparisons among the direct estimation method, analysis with original PVs, and analysis with the conditioned PVs. Addressing the concerns of NCES data linkage and its usage, this study not only provides a guideline for researchers interested in NCES data linkage but also give insights on the relationship between NAEP achievement and these important external variables.

**Wednesday, October 26**

**1:15PM**

---

### ***Session E-3: Modernizing Bias Reduction Techniques***

---

Organizer: Richard Chard, *Office of Research, Evaluation, and Statistics*

Chair: Doug Williams, *Bureau of Labor Statistics*

Room 146C

#### **Seroprevalence of SARS-CoV-2 Antibodies: Representativeness of Nonprobability Serology Samples from Multiple Commercial Laboratories in the United States**

Yun Kim, *ICF*

Ronaldo Iachan, *ICF*

Adam Lee, *ICF*

Lee Harding, *ICF*

Kristie Clarke, *CDC*

Our study focuses on measuring the seroprevalence of SARS-CoV-2 antibodies and therefore to estimate the proportion of the population that has been infected with the virus that causes COVID-19. Serology specimens were collected from commercial lab visits for non-COVID-19 related reasons in all 50 states, the District of Columbia, and Puerto Rico. The seroprevalence of SARS-CoV-2 antibodies was estimated by state and age, and overall, for every data collection round. To assess potential bias, we examined the geographic and socio-demographics representation of the nonprobability sample. We developed mapping tools that can be used to adjust the sample distribution in every round. Due to the absence of individual socio-demographics data, we aggregated the data to county-level American Community Survey data by age, sex, race, income, and education. We developed logistic regression models for the likelihood of being represented in the sample, controlling for socio-demographics. We discuss the implications for weighting adjustments as well as for the new phase of data collection focused on the pediatric population.

#### **Finding and Handling Bias in Clustered Randomized Trials**

Bradford Chaney, *National Academies of Sciences, Engineering, and Medicine*

A common view is that data from randomized control trials and clustered randomized trials are not subject to bias and do not require statistical adjustments for bias. The lack of bias is one of statistical expectation: any individual sample may be biased, even if simply from bad luck in the randomization, but across large samples and repeated replications one expects no bias. Clustered randomized trials face a greater risk of bias because people are not distributed equally in terms of demographic characteristics, and the number of clusters in an evaluation is more important in affecting the degree of bias than the number of individuals within the clusters. Using simulations, this paper finds that bias is common in random cluster samples of schools. For example, when comparing ten treatment group schools with ten control group schools, half of 1,000 samples showed differences greater than 11 percent in the percentage of students who were Black. We next examine multiple statistical models to evaluate their success in managing bias. Well-specified models successfully handled the bias, but not a standard black box approach, which produced incorrect and misleading estimates.

## **Analyzing Nonresponses for National Census of Ferry Operators**

Young-Jun Kweon, *Bureau of Transportation Statistics*

Aubrey Nguyen, *Bureau of Transportation Statistics*

Clara Reschovsky, *Bureau of Transportation Statistics*

The National Census of Ferry Operators (NCFO) is a biennial census of all ferry operators operating in the United States and the Bureau of Transportation Statistics (BTS) has conducted the census since 2006. To improve the quality and usability of the census, BTS has been engaging in several efforts such as modernizing the survey instrument to an online collection, reducing the need for follow-up, and automating the data edit process. BTS performs its first nonresponse bias study on the 2020 NCFO for further improvement and explores several ways of calculating response rates and analyzing units and item nonresponses to accommodate unique characteristics of the census. Since the NCFO is a census, not a survey, the nonresponse bias study performed is different from that on a survey such as weight adjustments not being considered as part of the study. The presentation introduces the NCFO, discusses its unique characteristics and response/nonresponse decision process, and present analysis findings and their implications.

**Wednesday, October 26**

**1:15PM**

---

### ***Session E-4: Topics in Income Volatility Measurement***

---

Organizer: Mark A Klee, *U.S. Census Bureau*

Chair: Rebecca Heller, *Congressional Budget Office*

Discussant: Charles Hokayem, *U.S. Census Bureau*

Room 140A

## **Reconciling Trends in Male Earnings Volatility: Evidence from the SIPP Survey and Administrative Data**

Michael Carr, *University of Massachusetts, Boston*

Robert Moffitt, *Johns Hopkins University*

Emily Wiemers, *Syracuse University*

As part of a set of papers using the same methods and sample selection criteria to estimate trends in male earnings volatility across survey and administrative datasets, we conduct a new investigation of male earnings volatility using data from the Survey of Income and Program Participation (SIPP) survey and SIPP-linked administrative earnings data (SIPP GSF). We find that the level of volatility is higher in the administrative earnings histories in the SIPP GSF than in the SIPP survey but that the trends are similar. Between 1984 and 2012, volatility in the SIPP survey declines slightly while volatility in the SIPP GSF increases slightly. Including imputations due to unit nonresponse in the SIPP survey data increases both the level and upward trend in volatility and poses a challenge for estimating a consistent series in the SIPP survey data. Because the density of low earnings differs considerably across datasets, and volatility may vary across the earnings distribution, we also estimate trends in volatility where we hold the earnings distribution fixed across the two data sources. Differences in the underlying earnings distribution explain much of the difference in the level of and trends in volatility between the SIPP survey and SIPP GSF.

## **Projecting Monthly Poverty Rates in the United States: New Estimates from the Survey of Income and Program Participation**

Lewis Warren, *U.S. Census Bureau*

Ani Silwal, *U.S. Census Bureau*

This paper proposes a methodology for projecting monthly poverty rates from the Survey of Income and Program Participation (SIPP). SIPP is the only major household survey in the United States that tracks monthly poverty over time. Information needed to determine poverty status in the survey is collected for the

calendar year preceding each SIPP interview, meaning that the lag in releasing SIPP poverty rates can exceed a year. This can significantly limit the value of the SIPP poverty data when trying to make real-time policy decisions, such as during the COVID-19 pandemic. To address this limitation, this paper utilizes data from Basic Monthly Current Population Survey files, which are released monthly for the proceeding calendar month, to project monthly poverty rates.

### **A Stitch in Time: Evaluating Seam Bias in the Redesigned SIPP**

Neil Bennett, *U.S. Census Bureau*

Mark A Klee, *U.S. Census Bureau*

Robert Munk, *U.S. Census Bureau*

Seam bias, the tendency for transitions to occur disproportionately at the precise meeting point between two reference periods of a longitudinal survey, is a common data quality concern for survey methodologists. We study seam bias in employment transitions using the 2014 panel of the Survey of Income and Program Participation (SIPP). We document the level of seam bias and quantify how much of seam bias comes from four sources: cross-sectional imputation, misreporting by a proxy respondent, misreporting in the presence of dependent interviewing, and misreporting in the absence of dependent interviewing. The 2014 SIPP panel is a fertile ground for studying seam bias, because it features fewer seams and longer recall periods than prior SIPP panels.

**Wednesday, October 26**

**1:15PM**

---

### ***Session E-5: Implementing Title III of the Evidence Act: A Framework to Expanding Data Access for Evidence Building***

---

Organizer: John Finamore, *National Center for Science and Engineering Statistics*  
Room 140B

### **ICSP's Vision for Implementing Title III of the Evidence Act**

Emilda Rivers, *National Center for Science and Engineering Statistics*

### **The Trust Regulation: Documenting the Fundamental Responsibilities of Recognized Statistical Agencies or Units**

Robert Sivinski, *Office of Management and Budget*

### **The Access Regulation: Expanding Secure Access to CIPSEA Data Assets**

Spiro Stefano, *Economic Research Service*

### **The Standard Application Process: Creating Uniformity in the Process for Requesting Access to Restricted Data**

Alex Marten, *Environmental Protection Agency*

**Wednesday, October 26**

**1:15PM**

---

***Session E-6: Measuring Sex, Gender Identity, and Sexual Orientation:  
Findings of the National Academies***

---

Organizer: Tara Becker, *National Academies of Sciences, Engineering, and Medicine*

Chair: Katrina Stone, *National Academies of Sciences, Engineering, and Medicine*

Discussant: Christina Dragon, *National Institutes of Health Sexual and Gender Minority Research Office*

Room 144ABC

**Guidelines for the Measurement of Sex and Gender**

Nancy Bates, *U.S. Census Bureau (Retired)*

Sex and gender are often conflated in surveys and other data collection instruments. These two variables, however, are distinct and multidimensional constructs that are both contextually and temporally fluid and may not align with each other. When measures fail to provide clarity on the dimensions of sex and/or gender that they seek to assess, respondents are left to decide how to interpret the question, resulting in mismeasurement and poor data quality. To address this data quality problem and to advance efforts to field measures that can identify transgender populations, the National Academies of Sciences, Engineering, and Medicine have convened a consensus study committee to develop standards for the measurement of sex and gender. The report also assesses the measurement of sexual orientation, which is linked to sex and gender by the degree to which sexual orientation is commonly defined by the gender of a person's romantic partners in relation to the person's own gender. This presentation will discuss the committee's approach to defining the constructs of sex, gender, and sexual orientation, as well as the principles the committee developed to guide its deliberations on strategies to improve measurement of these constructs. It will also detail the committee's conclusions and recommendations regarding the measurement of sex and gender in surveys and other research, administrative contexts, and clinical settings.

**Measuring Sexual Orientation Identity and Gender Identity: Findings from the National Academies**

Kellan Baker, *Whitman-Walker Institute*

Sexual orientation identity is the dimension of sexual orientation most consistently tied to experiences with material forms of discrimination and noted explicitly in non-discrimination laws and policies aimed at protecting or harming sexual minorities. Moreover, gender minorities experience greater health and economic well-being disparities, patterns of stigma, victimization and violence, lifetime suicide attempts, and substance use disorders compared to non-sexual minorities; however, the routine collection of this data to monitor such trends is scarce. The National Academies' Committee on Measuring Sex, Gender Identity, and Sexual Orientation evaluated existing approaches to measuring sexual orientation identity and gender identity with regards to construct validity and item performance. Specifically, we document the current measures used to assess sexual orientation and gender identity, the pros and cons associated with each, and the empirical evidence of their validity as measures in order to develop recommendations on which measure(s) to use. This presentation will discuss data supporting the final recommendation for how to measure sexual orientation and gender identity, as well as areas in which more research is needed to further efforts to comprehensively and meaningfully represent the range of sexual and gender identities in the U.S. as the terminology and social and political landscape around sexual and gender identity continues to evolve.

## **Developing Measures of Intersex Status/Differences in Sex Development at the Population Level**

Tara Becker, *National Academies of Sciences, Engineering, and Medicine*  
Katharine Dalke, *Penn State University*

The history of intersex persons—those with differences in sex development (Intersex/DSD)—in the United States has largely been one of invisibility and erasure. Though this population faces unique challenges, there have been almost no efforts to document and measure at the population level the social, economic, and health-related disparities and inequities they experience. This absence of information is driven by the underdevelopment of survey measures that can be used to identify this population. Though several measures have been proposed to identify Intersex/DSD populations, to date, none have been tested and validated within both Intersex/DSD populations and the general population. The National Academies convened the Committee on Measuring Sex, Gender Identity, and Sexual Orientation to review the research literature in the U.S. and other English-speaking countries to identify and assess the full range of measures that have been proposed and tested for use in identifying Intersex/DSD status. For each measure, we evaluate the complexity of the question, the populations in which it has been tested, and the results of standard methods of question validation, such as cognitive interviewing and item nonresponse. We then identify the most promising measure(s) that can be used to identify Intersex/DSD populations. We further note that Intersex/DSD status can be considered sensitive health information and provide guidance for when such data collections are appropriate and the care that data administrators should use in protecting this population from disclosure of this information.

**Wednesday, October 26**

**3:15PM**

---

### ***Session F-1: Applications of Statistical Methods in Economic Areas***

---

Organizer: Chris Moriarity, *National Center for Health Statistics*

Chair: Chris Moriarity, *National Center for Health Statistics*

Room 146A

## **What Most Affects the Probability of Receiving Public Assistance? Examining the Effect of Family Background and Educational Attainment on Receiving Public Assistance with Multivariate Regression Analysis**

Patricia I. Vargas, *U.S. Department of State*

Understanding poverty as a state of being poor with regard to concretely being unable to meet basic needs (e.g., water, food, clothes, and shelter), its concerted study in the United States—especially amidst global challenges like COVID-19—is crucial for soundly informing and assessing law, policy, and programs that ensure a robust society. I examine how family background and educational attainment interact to jointly affect poverty in the United States using data from a nationally representative panel study sponsored by the National Center for Education Statistics (NCES). After proxying poverty with receiving public assistance, I tested 18 independent variables consistent with demographic and family background by conducting a multivariate regression analysis. Ten variables were found to have a statistically significant effect on the probability of receiving public assistance with dependents, recent unemployment, and being female being among the strongest predictors of receiving public assistance ( $p < 0.001$ ). Roughly 36% of the variation in receiving public assistance is explained by the 18 independent variables tested. More importantly, and given the unique unemployment rate of 2020, results signaled a need to buttress public programs through at least 2023.

## **Alternative Data and Hedonic Price Indexes in the U.S. Consumer Price Index: A Review of Recent Research**

Craig Brown, *Bureau of Labor Statistics*

Jeremy Smucker, *Bureau of Labor Statistics*

We explore the use of alternative data sources and methods for adjusting price indexes for quality change to modernize elementary price indexes in the U.S. Consumer Price Index. Our motivation is to obtain measures of quality adjusted price change at scale using automated techniques where possible. We combine web-scraped and household expenditure survey data and use these data to estimate hedonic regression models. We examine price estimation methods that consider the valuation of both observable and unobservable characteristics at the time of product turnover. We employ a machine learning algorithm to assist in the selection of model regressors to yield more precise predictions of price. We apply these methods with data for wireless telephone services and televisions. The resulting hedonic regression models have high predictive accuracy and are used to construct hedonic price indexes.

## **Skilled Technical Workforce Classification**

Guy Leonel Siwe, *University of Virginia*

Cesar Montalvo, *University of Virginia*

Vicki Lancaster, *University of Virginia*

Haleigh Tomlin, *University of Virginia*

Skilled technical workforce are occupations that require technical knowledge but no bachelor's degree for entry. In 2015, Rothwell proposed the metrics for classifying STW occupations using education and knowledge data from the Occupational Information Network Content Model. These data present a number of fitness-for-use issues: small sample sizes, data over a decade old, 196 occupations that have no education and/or knowledge data; plus these metrics are calculated ignoring the standard errors of the estimates. The STW is a function of the nature of work which is rapidly changing due to emerging technologies. We propose a new approach for classifying STW occupations. We use labor market information which provides detailed information at the occupation level, is updated daily, minimum education requirement, and detailed information on technical skills. As a proxy for knowledge, we evaluate the use of skills listed in these job-ads and in lieu of a discrete criterion for knowledge we use the machine learning method, Non-negative Matrix Factorization, to separate occupations into STW and non-STW.

## **United States Inflation Experience across the Income Distribution**

Joshua Klick, *Bureau of Labor Statistics*

Any Stockburger, *Bureau of Labor Statistics*

The Bureau of Labor Statistics (BLS) produces the Consumer Price Index (CPI) as a measure of price change (or inflation) faced by consumers. The headline CPI targets the inflation experience of urban consumers (CPI-U) which may not reflect the inflation experience of an individual household or demographic group. Increasingly there is user demand for CPIs across the income distribution. In June 2021, an Interagency Technical Working Group convened by the Office of Management and Budget issued a report recommending the BLS produce a CPI to be used in the calculation of the U.S. Official Poverty Measure. In April 2022, The National Academy of Sciences issued a report recommending development of price indexes by income quintile to allow for comparisons across groups.

In the spirit of these reports, this paper presents consumer price indexes for the lowest quintile, highest quintile, and collapsed 2nd to 4th quintile income groups. For the lowest quintile, this paper compares results with unadjusted income versus equivalized income to account for household size. Historically, inflation estimates for lower income households are generally higher than higher income households. This paper explores data through 2021 to demonstrate whether historical patterns hold during this updated period.

---

**Session F-2: Innovations from the Criminal Justice Administrative Records System (CJARS) Project**

---

Organizer: Keith Finlay, *U.S. Census Bureau*

Chair: Keith Finlay, *U.S. Census Bureau*

Discussant: Anne Morrison Piehl, *Rutgers University*

Room 146B

**Hierarchical Approaches to Text-based Offense Classification**

Michael Mueller-Smith, *University of Michigan*

Jay Choi, *University of Michigan*

David Kilmer, *Measures for Justice*

Sema A. Taheri, *Measures for Justice*

Researchers and statisticians working with multi-jurisdictional crime data often must classify offense information into a common scheme. No comprehensive standard currently exists, nor a mapping tool to transform raw description information into offense type variables. This paper introduces a new offense classification schema, the Uniform Crime Classification Standard (UCCS), and the Text-based Offense Classification (TOC) tool to address these shortcomings. The UCCS schema draws from existing Department of Justice efforts, with the goal of better reflecting offense severity and improving offense type disambiguation. The TOC tool is a machine learning algorithm that uses a hierarchical, multilayer perceptron classification framework, built on 313,209 unique hand-coded offense descriptions from 24 states, to translate raw description information into UCCS codes. In a series of experiments, we test how variations in data processing and modeling approaches impact recall, precision, and F1 scores to assess their relative influence on the TOC tool's performance. The code scheme and classification tool are collaborations between Measures for Justice and the Criminal Justice Administrative Records System.

**CJARS Data Quality Automation**

Brian Miller, *University of Michigan*

The CJARS team collects data from over 300 distinct criminal justice agencies. Those raw administrative data are then harmonized into a common national data schema. In order to assess the quality of the harmonized data, the team has developed a suite of automated data quality validation tools to identify and triage data quality issues. These tools allow a small team to produce a high-quality microdata product even as the project scales up to include more and more agency data providers.

**Measuring Intergenerational Exposure to the U.S. Justice System: Evidence from Longitudinal Links between Survey and Administrative Data**

Brittany Street, *University of Missouri*

Keith Finlay, *U.S. Census Bureau*

Michael Mueller-Smith, *University of Michigan*

Intergenerational exposure to the justice system is both a marker of vulnerability among children and a measurement of the potential unintended externalities of crime policy in the U.S. Estimating the size of this population has been hampered by inadequate data resources to date including the inability to (1) observe non-incarceration sources of exposure, (2) follow children throughout their childhood, and (3) measure multiple adult influences in increasingly dynamic households. To overcome these challenges, we leverage over a terabyte of confidential administrative and survey data linked with the Criminal Justice Administrative Records System (CJARS). In doing so, we find substantially larger prevalences than previously reported: 9% of children born between 1999-2005 were intergenerationally exposed to prison, 18% to a felony conviction, and 39% to any criminal charge; charge exposure rates reach as high as 62% for Black children. We regress these newly quantified types of exposure on measures of child well-being to help gauge their importance. We

find that all types of exposure (parent vs. non-parent, prison vs. charges, current vs. previous) are negatively correlated with child outcomes and are often statistically indistinguishable from each other, suggesting more children are harmed by crime and criminal justice than previously thought.

**Wednesday, October 26**

**3:15PM**

---

***Session F-3: Drawing Optimal Statistical Inferences from Nonprobability or Low Response-rate Probability Surveys: Some Recent Advances and Applications***

---

Organizer: Yulei He, *National Center for Health Statistics*

Chair: Yulei He, *National Center for Health Statistics*

Discussant: Phillip S. Kott, *RTI International*

Room 146C

**Variable Inclusion Strategies for Adjusting the Weights of Surveys Subject to Selection Bias**

Katherine Irimata, *National Center for Health Statistics*

Yan Li, *University of Maryland*

Yulei He, *National Center for Health Statistics*

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

As web-based surveys have become increasingly available to evaluate time-sensitive priority topics, various propensity score (PS)-based adjustment methods have been developed to improve population representativeness for nonprobability- or probability-sampled web surveys subject to selection bias. Conventional PS-based methods construct pseudo-weights for web samples using a higher-quality reference probability sample. The bias reduction, however, depends on the outcome and variables collected in both the web and reference samples. A central issue is identifying the variables for inclusion in the PS-adjustment. Rather than including all available variables as suggested by past literature, this research demonstrates that only certain types of variables are needed in the PS-adjustment. This is illustrated by evaluating the prevalence of asthma in the US from the National Center for Health Statistics' Research and Development Survey.

**Generating Survey Weights Using a Machine-learning Method and the Entropy Balancing Calibration Technique**

Joseph Kang, *U.S. Census Bureau*

We demonstrate how to build weights that adjust for non-response bias in a non-probability survey data set using a machine-learning method and the entropy balancing calibration technique in two steps: 1) fit a survey response probability model using a machine-learning method called the generalized boosted model to create weights with the inverse of the probability; and 2) calibrate the weights using the entropy balance calibration. The first step is to handle problems of the association between the response status and covariates, which harbor non-linearity, collinearity, and complex interactions among them. The second step is to ensure the evidence of balanced distributions among the covariates between the respondents and non-respondents. The contribution of this two-step procedure is to illustrate the application of a machine-learning method and novel entropy-balancing calibration technique to attenuate or remove survey non-response bias in the American Community Survey (ACS) data set.

## **Methods for Combining Probability and Nonprobability Samples Under Unknown Overlaps**

Terrance D. Savitsky, *Bureau of Labor Statistics*

Matthew R. Williams, *RTI International*

Julie Gershunskaya, *Bureau of Labor Statistics*

Vladislav Beresovsky, *National Center for Health Statistics*

Nels G. Johnson, *U.S. Department of Agriculture*

Nonprobability (convenience) samples are increasingly sought to stabilize estimations for one or more population variables of interest that are performed using a randomized survey (reference) sample by increasing the effective sample size. Estimation of a population quantity derived from a convenience sample will typically result in bias since the distribution of variables of interest in the convenience sample is different from the population. A recent set of approaches estimates conditional (on sampling design predictors) inclusion probabilities for convenience sample units by specifying reference sample-weighted pseudo likelihoods. This paper introduces a novel approach that derives the propensity score for the observed sample as a function of conditional inclusion probabilities for the reference and convenience samples as our main result. Our approach allows specification of an exact likelihood for the observed sample. We construct a Bayesian hierarchical formulation that simultaneously estimates sample propensity scores and both conditional and reference sample inclusion probabilities for the convenience sample units. We compare our exact likelihood with the pseudo likelihoods in a Monte Carlo simulation study.

## **New Measures for Assessing Non-ignorable Selection Bias in Non-probability Samples and Low Response Rate Probability Samples**

Rebecca Andridge, *Ohio State University*

Brady T. West, *University of Michigan*

Recent developments in survey statistics have yielded simple, novel measures of the non-ignorable selection bias in estimates of means, proportions, and regression coefficients that may arise due to deviations from ignorable sample selection, where these deviations might be introduced by the sampling mechanism (e.g., non-probability sampling) or survey nonresponse. Responsive survey designs rely on active monitoring of sound indicators of survey errors to inform real-time design decisions, and these new measures, which are easy to compute at any point in time during a data collection, have the potential to serve as useful indicators of the possible selection bias in estimates of interest. This presentation will review the computation of these indicators, the data required to compute them, software tools for computing them, and how they might be actively monitored in real time to inform design decisions in responsive survey designs.

**Wednesday, October 26**

**3:15PM**

---

***Session F-4: The Need for a National Study of Post-Secondary Faculty***

---

Organizer: Hans-Joerg Tiede, *American Association of University Professors*

Chair: Felice Levine, *American Education Research Association*

Discussant: Valerie Martin Conley, *Idaho State University*

Room 140A

**Defining, Counting and Understanding Faculty in the United States: Strengths and Limitations of Current Higher Education Data**

Caren Arbeit, *RTI*

KC Culver, *University of Southern California*

Who are “faculty”? Changes to the definitions and approaches used to identify and classify faculty in federal data, as well as changing institutional practices for employing faculty, have created challenges to our accurate understanding of who faculty are. These issues have particular relevance in terms of the representativeness of non-tenure track and contingent faculty, especially those who are part time. We will discuss how the faculty role was defined in NSOPF 2004, and look at examples of current definitions. In this paper, we present a comparative analysis of definitions and statistics from public-use and published data from IPEDS, ECDS, and SDR and other surveys to demonstrate the need for more comprehensive data on faculty.

**Data: The Foundation of Efforts to Advance Equity and Diversity among our Nation’s Faculty**

Laura Perna, *University of Pennsylvania*

“Relative to their representation among bachelor’s degree recipients, women continue to be underrepresented among the highest-ranking faculty and people of color continue to be underrepresented among faculty at all ranks. There are also important variations in representation by field and institutional type. In order to increase equity and diversity among our nation’s faculty, institutional leaders and researchers need nationally representative data that describe the characteristics of work and life of faculty of different demographic groups who hold different positions, in different disciplines, and different institutional types. This paper discusses the questions college and university leaders need to ask to improve faculty equity and diversity, and the types of individual-level data that are needed to address those questions.

**The Importance and Value of Discipline-Level Data**

Robert Townsend, *American Academy of Arts and Sciences*

To understand the roles and challenges of the professoriate, it is vital to understand the differences between the fields. Robert Townsend will discuss why and how national-level data are not sufficient to answer vital questions about the work being performed by faculty members, as well as the current and future needs of a healthy and productive higher education sector.

**NSOPF Institution-Level Data on Personnel Policies and Practices**

Hans-Joerg Tiede, *American Association of University Professors*

In addition to individual faculty members, NSOPF included the administration of questionnaires to institutions selected in the first stage of sampling in order to inquire about tenure practices, fringe benefits, and personnel actions taken by those institutions during the preceding five years. We will review survey findings related to personnel policies and practices from the four iterations of the survey and compare selected findings to currently available data. We will then consider questions that cannot be easily administered by organizations that do not have the standing of NCES, because they are “institutionally sensitive” and could result in high item nonresponse when administered by organizations such as the AAUP.

**Wednesday, October 26**

**3:15PM**

---

***Session F-5: Classifying Settlement Form and Function: Urban, Rural, Metropolitan and Micropolitan Area Definitions and Classifications***

---

Organizer: Michael Ratcliffe, *U.S. Census Bureau*

Chair: Michael Ratcliffe, *U.S. Census Bureau*

Room 140B

**The Census Bureau's Urban and Rural Classification and Overview of 2020 Census Urban Area Criteria**

Michael Ratcliffe, *U.S. Census Bureau*

The Census Bureau's urban and rural classification provides an important baseline for other federal agency definitions of urban and rural for both analytical and programmatic purposes. In addition, the Census Bureau's urban areas provide the cores for the Office of Management and Budget's Core Based Statistical Areas. In this presentation, I provide an overview of the Census Bureau's urban and rural classification and discussion key changes to the concept and criteria implemented with the delineation of areas based on 2020 Census data. In addition, I provide a few thoughts regarding the need to consider moving from the current urban-rural dichotomy to a continuum of settlement categories.

**The 2020 Standards for Delineating Core Based Statistical Areas**

Paul Mackun, *U.S. Census Bureau*

The Office of Management and Budget (OMB) is responsible for the delineation of core based statistical areas (CBSAs)--the collective term for metropolitan and micropolitan statistical areas--as well as for related statistical areas. These areas are used by federal statistical agencies, such as the U.S. Census Bureau, for the tabulation and dissemination of statistical data. Every decade OMB reviews and, if warranted, revises the standards used for delineating these areas. The presentation will describe and explain the 2020 standards for delineating CBSAs. It will begin by describing the different types of statistical areas identified under the standards, including CBSAs and related statistical areas. It then will outline the review process associated with the development of the 2020 standards. The presentation will conclude by summarizing the standards used to qualify these areas

**Understanding Rural Definitions**

John Cromartie, *Economic Research Service*

The effectiveness of any rural research project, policymaking endeavor, or economic development program depends, in large part, on selecting an appropriate definition of rural. The U.S. Census Bureau provides a widely used, statistically consistent definition, but many other versions are currently employed to define rural, just among federal agencies. Here we present analysis to aid in choosing appropriate rural definitions, beginning with understanding key differentiating features. Rural definitions vary in terms of underlying concepts, geographical building blocks, criteria, and thresholds. By analyzing the characteristics of these components and comparing a sample of rural definitions, we can clear up much of the confusion surrounding this issue and enable better decision-making in choosing definitions that best fit the needs of specific applications.

**The Effects of Geography on Data Quality**

Rolf R. Schmitt, *Bureau of Transportation Statistics*

Ed Strocko, *Bureau of Transportation Statistics*

The way in which data are aggregated by classification systems affects utility and other aspects of data quality. Geographic classification systems are no exception. The effects of geography are noted in the Framework for Data Quality, published by the Federal Committee on Statistical Methodology to provide an inventory of possible threats to data quality and a common language for describing and understanding those threats. After a brief overview of the Framework, the effects of geographic classification on data quality are explored.

Wednesday, October 26

3:15PM

---

***Session F-6: Reporting on the Pulse of America – Rapid Data and Collection and Information Sharing Initiatives in the Federal Government***

---

Organizer: Chris Chapman, *National Center for Education Statistics*

Chair: Stephen Blumberg, *National Center for Health Statistics*

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

Room 144ABC

**Household Pulse Survey. Foundation For a Rapid Response, High-Frequency Household Data System**

Jason Fields, *U.S. Census Bureau*

Jennifer Hunter Childs, *U.S. Census Bureau*

Cassandra Logan, *U.S. Census Bureau*

Household Pulse Survey (HPS) was designed by the Census Bureau and partner agencies to measure the impact of the coronavirus (COVID-19) pandemic. From the experimental survey's inception on March 23, 2020, the Census Bureau staff worked with experts from the five partner agencies to launch the HPS. The ability to understand how individuals experienced the pandemic was deemed critical to the governmental and non-governmental response in light of business curtailment and closures, stay-at-home orders, school closures, changes in the availability of consumer goods and consumer patterns, and other abrupt and significant changes to American life. The survey was designed to quickly collect and provide rapid estimates to support agency priorities. This concept and the more than 2 years of subsequent implementation has generated the foundation for a new data collection framework, a high-frequency rapid-response program to support data needs. The initial program leveraged email and SMS text message contacts as the only contact strategy, this presentation highlights where the HPS program has been and where it is going.

**School Pulse Panel and Development of a Rapid Response Data Collection System for Establishment Studies - Design, Challenges, and Accomplishments**

Rachel Hansen, *National Center for Education Statistics*

Ryan Iaconelli, *National Center for Education Statistics*

Rebecca Bielamowicz, *National Center for Education Statistics*

The School Pulse Panel (SPP) was designed by the National Center for Education Statistics at the Institute of Education Sciences and involves ongoing coordination with multiple agency and administration partners. SPP is designed to measure the impact of COVID-19, broadly defined, on U.S. public schools and public education. Data are collected from schools and reported on each month. The presentation will provide information on operational challenges and successes associated with designing, fielding, and releasing information from a high frequency establishment survey (of schools) while working within the traditional federal study clearance, data collection, and statistics release system. How lessons learned and processes developed for SPP can be used more broadly beyond measuring the impact of the pandemic in public schools will also be presented.

## **Leveraging Commercial Online Survey Panels for Timely and Actionable Health Statistics**

Stephen Blumberg, *National Center for Health Statistics*

Jennifer Parker, *National Center for Health Statistics*

Paul Scanlon, *National Center for Health Statistics*

Jonaki Bose, *National Center for Health Statistics*

Amy Brown, *National Center for Health Statistics*

The COVID-19 pandemic highlighted the need for timely critical health and health care data from population-based surveys. Legacy surveys such as the National Health Interview Survey could not meet this need due to both the disruption to data collection from the pandemic and the longer processing times prior to data release. Even as legacy surveys return to the proven survey operations that have made them gold standards for accurate and reliable data collection, the demand is likely to continue for faster actionable data. To meet this demand, the National Center for Health Statistics (NCHS) is developing a program of quarterly surveys, using commercial probability-sampled online survey panels to collect data on emerging topics of interest to NCHS, other Centers for Disease Control and Prevention (CDC) programs, and the Department of Health and Human Services (DHHS). This new “rapid surveys” program builds on the success of NCHS’s Research and Development Survey (RANDS), a survey platform started in 2015 that was deployed during the pandemic to collect data on loss of work because of illness with COVID-19, telemedicine access and use before and during the pandemic, and inability to get specific types of health care because of the pandemic. This presentation will provide a brief overview of RANDS and then highlight goals, key features, and challenges of the Rapid Surveys program.

---

## ***Abstract Listings for Thursday, October 27***

---

- |                         |                     |
|-------------------------|---------------------|
| • Concurrent Sessions G | 8:30 – 10:15 am     |
| • Concurrent Sessions H | 10:30 am – 12:15 pm |
| • Concurrent Sessions I | 1:30 – 3:15 pm      |
| • Concurrent Sessions J | 3:30 – 5:15 pm      |

**Thursday, October 27**

**8:30AM**

---

***Session G-1: Insights from the Congressional Watchdog on Equity in Evaluation Design***

---

Organizer: Sirin Yaemsiri, *Government Accountability Office*

Chair: Jared Smith, *Government Accountability Office*

Discussant: Frances Tirado, *Government Accountability Office*

Room 140A

**Geospatial Analysis and Regression Modeling to Assess Social and Demographic Disparities in Locations of Lead Water Pipes and Pavement Conditions of the National Highway System**

Sonya Vartivarian, *Government Accountability Office*

Mark Braza, *Government Accountability Office*

Caitlin Cusati, *Government Accountability Office*

Diane Raynes, *Government Accountability Office*

Infrastructure ensures the economic vitality and quality of life in the country, from the water we drink to the roads that connect communities. This talk will describe two studies examining infrastructure in communities across the country, with an application to cities served by lead service lines in drinking water systems and pavement conditions of the National Highway System. These studies use geospatial analysis to overlay the infrastructure data with community indicators of social vulnerability, such as family poverty and race/ethnicity from the American Community Survey. Using generalized linear regression models, we examined whether disparities in infrastructure exist for socially vulnerable communities after accounting for other drivers. We will cover approaches to identifying potential drivers, conducting geospatial analysis, developing regression models to simultaneously adjust for various drivers, and evaluating the results in light of the focus on advancing equity among executive agencies.

**Assessing Disparities in FEMA Flood Map Investments using Multivariate Analysis**

Lijia Guo, *Government Accountability Office*

Joel Aldape, *Government Accountability Office*

Mark Braza, *Government Accountability Office*

John Vocino, *Government Accountability Office*

Flooding is the most frequent severe weather threat and the costliest natural disaster in the United States. The Federal Emergency Management Agency (FEMA) has invested millions every year in mapping flood hazard areas to help guide the National Flood Insurance Program as well as the flood mitigation efforts of property owners, community planners, insurers, and others. As part of our evaluation of FEMA's flood map program outcomes from 2012 to 2020, we identified and examined equity issues by linking administrative data to publicly available databases. We used geospatial analyses to overlay measures of flood risk and social vulnerability to FEMA communities, a geographic region delineated for the purposes of flood mapping. We used regression models to estimate the relationships between social vulnerability and program outcomes related to flood map investments. Our analysis found that FEMA's mapping investments for fiscal years 2012 to 2020 were greater where flood risks were higher. However, investments were lower for areas with higher percentages of socially-vulnerable populations, while accounting for areas with similar flood risk.

**Survival Analysis Modeling of Promotions in the State Department's Workforce**

Moon Parks, *Government Accountability Office*

Nisha Rai, *Government Accountability Office*

Melinda Cordero, *Government Accountability Office*

The U.S. Department of State has expressed a commitment to maintaining a workforce that reflects the diverse composition of the United States and has undertaken efforts to increase representation of diverse

groups at all levels of their workforce in its Civil and Foreign Services. We assessed these efforts by developing a discrete survival analysis model to assess differences in promotion outcomes for various demographic groups in the State Department's workforce while accounting for unique characteristics of State's promotion process. Using administrative workforce and payroll data at the employee-level, we estimated differences in promotion likelihoods across demographic groups by race/ethnicity and gender. We also accounted for individual and occupational factors other than race/ethnicity and gender that could influence promotion, such as serving in hardship assignments and having long-term leave in the previous year. Our analysis found generally lower promotion likelihoods for non-white race/ethnicity groups than for whites in both the Civil and Foreign Services. We also found that promotion outcomes for women compared with men were lower in the Civil Service and generally higher in the Foreign Service.

## **Imputing Race and Ethnicity in Federal Administrative Data: Two Applications of Bayesian Improved Surname Geocoding (BISG)**

Jeff Tessin, *Government Accountability Office*

Congress and other stakeholders often want to know how federal programs operate and perform across diverse populations. Yet data systems for some programs do not measure the race and ethnicity of participants, in part because these variables are not explicitly necessary to administer programs. To supplement existing administrative data, recent GAO evaluations have imputed race and ethnicity using the surnames and residential locations of program participants—an established method known as Bayesian Improved Surname Geocoding (BISG). This talk will describe the statistical assumptions and structure of BISG; review the literature on its accuracy; and summarize two recent GAO reviews of how customer service at the Social Security Administration and the use of COVID small business tax credits varied across imputed racial and ethnic groups. We illustrate the strengths and weaknesses of imputation and contrast it with other data collection methods, such as surveys and linkage to other federal datasets that might provide similar evidence for administration and oversight.

**Thursday, October 27**

**8:30AM**

---

## ***Session G-2: Highlights of CNSTAT Report on Transparency in Statistical Information***

---

Organizer: Michael L. Cohen, *Committee on National Statistics*

Chair: Dan Kasprzyk, *NORC at the University of Chicago*

Discussant: Emilda Rivers, *National Center for Science and Engineering Statistics*

Room 140B

## **Transparency of Official Statistics: Part of the Recent Emphasis on Reproducibility of Scientific Activities**

Sarah Nusser, *Iowa State University*

Much attention has recently been given to improving \ reproducibility and other forms of rigor in scientific research, including a recent report of the National Academy of Sciences. Since official statistics depend on the application of scientific methods, there is a strongly related need for data and methods to be accessible for examination and reuse. These forms of transparency have a number of benefits to statistical agencies, their stakeholders and the research community. We discuss transparency in this context and provide an overview of practices statistical agencies can adopt.

## **The Use of Modern Software Development Tools for Documentation of Methods Used**

Lars Vilhuber, *Cornell University*

There are several tools that are now commonly used in academia and industry for the development of software by a number of co-developers, as well as tools for validating software code, etc. These tools could also be used to help in documenting the current methods used by statistical agencies for data collection, data treatment, estimation, and for assessment of the quality of such estimates. We provide an overview of the benefits that could be gained through their greater use in the federal statistical system.

## **The Use of Metadata Standards and Tools for Greater Transparency of Official Statistics**

Dan Gillman, *Bureau of Labor Statistics*

Metadata standards such as several versions of DDI, SDMX, and GSBPM are currently used in a large number of national statistical offices to help document and reuse data and methods. What these standards achieve and how they can be brought into a statistical agency are discussed.

**Thursday, October 27**

**8:30AM**

---

### ***Session G-3: Using Spatial Data to Inform Official Statistics***

---

Organizer: Denise A. Abreu, *National Agricultural Statistics Service*

Chair: Matthew Deaton, *National Agricultural Statistics Service*

Room 143A

## **Bayesian Nonparametric Multivariate Spatial Mixture Mixed Effects Models with Application to American Community Survey Special Tabulations**

Scott Holan, *University of Missouri; U.S. Census Bureau*

Ryan Janicki, *U.S. Census Bureau*

Andrew M. Raim, *U.S. Census Bureau*

Jerry Maples, *U.S. Census Bureau*

Leveraging multivariate spatial dependence to improve the precision of estimates using American Community Survey data and other sample survey data has been a topic of recent interest among data users and federal statistical agencies. One strategy is to use a multivariate spatial mixed effects model with a Gaussian observation model and latent Gaussian process model. In practice, this works well for a wide range of tabulations. Nevertheless, in situations in which the data exhibit heterogeneity within or across geographies, and/or there is sparsity in the data, the Gaussian assumptions may be problematic and lead to underperformance. To remedy these situations, we propose a multivariate hierarchical Bayesian nonparametric mixed effects spatial mixture model to increase model flexibility. The number of clusters is chosen automatically in a data-driven manner. The effectiveness of our approach is demonstrated through a simulation study and motivating application of special tabulations for American Community Survey data.

## **Model-based Estimation of the Number of U.S. Farms and Land in Farms from Survey, Spatial and Administrative Data**

Habtamu Benecha, *National Agricultural Statistics Service*

Tara Murphy, *National Agricultural Statistics Service*

Grace Yoon, *National Agricultural Statistics Service*

Arthur Rosales, *National Agricultural Statistics Service*

Luca Sartore, *National Agricultural Statistics Service*

Denise A. Abreu, *National Agricultural Statistics Service*

USDA's National Agricultural Statistics Service (NASS) produces annual estimates of the number of U.S. farms and land in farms by using the June Area Survey (JAS). The JAS is a survey conducted every year in June based on a sample selected from an area frame that covers all land in the continental U.S. To account for misclassification of farm tracts within sampled segments during pre-screening and data collection processes, estimation from the JAS uses a capture-recapture method. In this approach, records from an independent frame, the NASS list frame, are matched to the area-frame tracts. Weighting adjustments are then estimated by using logistic regression models fitted to the resulting matched dataset. In recent years, decreasing JAS sample sizes and other factors have been affecting the quality of the model-based estimates and the estimation process. This paper explores ways in which available spatial and administrative data can supplement the JAS to mitigate the impact of smaller sample sizes and improve estimates of the number of farms and land in farms. Results from simulation studies are presented with application of the methods to NASS data.

## **Bayesian Spatial Hierarchical Models for Differentially Private Measurements of Census Tabulations**

Ryan Janicki, *U.S. Census Bureau*

Andrew Raim, *U.S. Census Bureau*

Scott Holan, *University of Missouri; U.S. Census Bureau*

Kyle Irimata, *U.S. Census Bureau*

James Livsey, *U.S. Census Bureau*

Official statistical agencies have an obligation to protect the confidentiality of their respondents. Traditionally, the task of disclosure avoidance (DA) has consisted of a variety of different approaches, including cell suppression and swapping, among others. More recently, official statistical agencies (e.g., Census Bureau) have employed differential privacy (DP) methods to facilitate disclosure avoidance. A major component of this method involves the addition of noise to the observed measurements based on an appropriate DP noise distribution and a pre-specified privacy-loss budget. The consequences of implementing DP are reduced precision of published estimates, as well as the inability to publish more granular Census data, such as counts for some detailed race, ethnicity, and tribal groups at low levels of geography, that have been produced in the past. Consequently, it may be desirable to develop statistical models for the DP noisy measurements. One natural way to achieve this is using a Bayesian hierarchical model which utilizes auxiliary publicly available data sources, such as American Community Survey data and past Census data, and leverages spatial correlation to improve precision of estimates. We show that accurate, model-based estimates of the number of persons in a detailed race group in a target geography, which may be misaligned from the source data, such as an American Indian or Alaska Native (AIAN) area, can be made. As an example, the problem of estimating the number of Choctaw persons in counties and AIAN areas in Oklahoma using only observed DP measurements of the number of Choctaw persons residing in counties in Oklahoma is discussed.

**Thursday, October 27**

**8:30AM**

---

***Session G-4: Connecting the Dots: Leveraging Extant Data for Research***

---

Organizer: Maura Spiegelman, *National Center for Education Statistics*

Chair: Elise Christopher, *National Center for Education Statistics*

Room 143BC

**Creating the 21st Century Census Curated Data Enterprise (CDE) – Establishing Criteria for Choosing Case Studies to Develop the CDE**

Stephanie S. Shipp, *University of Virginia*

Joseph Salvo, *University of Virginia*

The Census Bureau has a long history of innovation. We propose accelerating these innovations by developing a Curated Data Enterprise (CDE). The CDE concept directly addresses changing economic and social conditions through exploiting multiple data sources across sample surveys, censuses, and administrative and private-sector data to produce more robust, granular, timelier, and comprehensive measures. This will allow addressing time-sensitive questions and require curating associated data and processes. For example, managing the impact of pandemics, wildfires, hurricanes, and changes in work on vulnerable populations require real-time and geographically detailed data to address stakeholder questions. The CDE evolution will benefit from selected case studies to exercise its capabilities through development, highlighting the challenges of discovering, integrating, and using data in unexpected new ways. We have sought and will share input from technical experts for criteria and priorities to define case studies that could accelerate the development of the CDE.

**Enhancing Race and Ethnicity Information in Medicaid Data: The Role of Census Bureau Data**

Aubrey Limburg, *U.S. Census Bureau*

Victoria Udalova, *U.S. Census Bureau*

Adam Kurczewski, *U.S. Census Bureau*

Medicaid data are used to investigate racial and ethnic disparities in health. However, there is considerable variation in completeness of race/ethnicity information in Medicaid data across US states. To address this problem, the Census Bureau's Enhancing Health Data (EHealth) research team assessed the feasibility and effectiveness of linking Medicaid race/ethnicity data with federal population-level data. This research linked Medicaid data to Census records to populate race/ethnicity information for beneficiaries missing this information and to validate race/ethnicity information in both data sources. Additionally, we evaluated the degree to which bias was introduced into estimates of health disparities (e.g., all-cause mortality) when Medicaid beneficiaries missing race/ethnicity are omitted from analysis. Our findings suggest that researchers should be cautious about using Medicaid data for certain states with a high degree of bias. This project demonstrates significant potential for using Census data to complement existing health records that commonly lack important demographic characteristics, such as race/ethnicity. We find that enhancing race/ethnicity information in Medicaid data with population-level Census records is feasible and can advance understanding of population health.

**Address Frame Enhancement with Administrative Data using an Interactive Mapping Dashboard**

Lee Fiorio, *NORC at the University of Chicago*

Ned English, *NORC at the University of Chicago*

Chang Zhao, *NORC at the University of Chicago*

Peter Herman, *NORC at the University of Chicago*

Frames for in-person household surveys are typically sourced from lists of addresses such as the USPS Delivery Sequence File (CDS or CDSF). However, due to CSD under-coverage, some additional housing unit listing is still necessary to enhance the CDS and minimize overall coverage bias. While address listing in-

person is costly, the growing availability of administrative data and advancements in interactive web mapping applications have opened the potential for CDS address list enhancement to be conducted in-office using GIS. This paper will offer an assessment of NORC's efforts to integrate CDS with spatial data including satellite imagery, building footprint data and county assessor tax parcel data improve the address frame for NORC's 2020 Decennial National Master Sample. In a sample of low CDS-coverage areas, we evaluate address lists generated from an interactive mapping dashboard in which in-office listers add addresses from administrative data that appear missing from the CDS. Results from our analysis demonstrate the contexts in which interactive administrative data enhancement of the CDS performs well and highlights some limits of remote housing unit enumeration.

## **Building the Social Impact Data Commons: Data Discovery and Acquisition Lessons Learned**

Joanna Schroeder, *University of Virginia*  
Kathryn Linehan, *University of Virginia*  
Micah Iserman, *University of Virginia*  
Joel Thurston, *University of Virginia*  
Cesar Monatlavo, *University of Virginia*  
Vicki Lancaster, *University of Virginia*  
Hanna Charankevich, *University of Virginia*  
Neil Kattampallil, *University of Virginia*  
Leonel Siwe, *University of Virginia*  
Stephanie Shipp, *University of Virginia*  
Aaron Schroeder, *University of Virginia*  
Sallie Keller, *University of Virginia*

Local decision-makers require data to inform action on access and equity issues. To address this need, we developed the Social Impact Data Commons, an open knowledge repository that co-locates data, curates data insights, and tracks issues over time. Guided by our Community Learning through Data Driven Discovery process, local stakeholders engaged to identify important policy areas, including food security, digital equity, and health access. We present opportunities and challenges associated with the discovery, collection, documentation, and dissemination of data from various sources to create new metrics at novel geographic units. The Data Commons infrastructure includes a data and metadata collection pipeline designed to track quality, transparency, and version using current standards. We use open-source tools, such as Dataverse and GitHub, to disseminate raw and statistically transformed data. We draw on best practices in information visualization and worked iteratively with stakeholders and experts to collect feedback on our data dashboard. Next steps include sustainably collecting opportunity data (e.g., local staple food availability) and improving user experience.

## **Finding Alternative Data: VIIRS Nightfire Oil and Gas Well Flare Data Evaluation**

Jeramiah Yeksavich, *Energy Information Administration*  
Nathan Agbemenyale, *Energy Information Administration*  
Caitlin Steiner, *Energy Information Administration*  
Michael Winkler, *Energy Information Administration*

Oil and natural gas producers flare gas during the production process for safety and economic reasons. Flared gas (FG) volumes reduce the amount of gas brought to market and contribute to energy industry related greenhouse gas emissions. The U.S. Energy Information Administration (EIA) does not survey FG volume data and relies on administrative and third-party data sources for volume data to support publications and analysis. The data collection and dissemination of each of these sources are based on different methodologies and vary in quality. Using the satellite-based Visible Infrared Imaging Radiometer Suite (VIIRS) sensor, National Oceanographic and Atmospheric Administration and the Colorado School of Mines Payne Institute for Public Policy researchers developed a methodology, dubbed Nightfire, for estimating annual FG volumes. A data source with a single methodology covering various geographic levels would be useful to EIA. EIA researchers completed an evaluation of the Nightfire methodology compared with other data sources used by EIA. EIA researchers will present their findings and highlight approaches for evaluating a new data source for an uncertain FG population.

Thursday, October 27

8:30AM

---

## ***Session G-5: Building a Sampling System for the Annual Integrated Economic Survey***

---

Organizer: Laura Bechtel, *U.S. Census Bureau*

Chair: Thomas Kirk White, *U.S. Census Bureau*

Discussants: David Marker, *Marker Consulting, LLC*; Ron Jarmin, *U.S. Census Bureau*  
Room 144ABC

### **Developing the Sample Design for the New Annual Integrated Economic Survey**

Katherine Jenny Thompson, *U.S. Census Bureau*

In the near future, the economic directorate of the U.S. Census Bureau will introduce the Annual Integrated Economic Survey (AIES), an economy wide survey that replaces a suite of six independently designed ongoing surveys. The AIES sample design requirements are informed by the user community's longstanding data needs (e.g., national and sub-national tabulations), as well as by extensive respondent research on collection. This paper provides an end-to-end high-level overview of the AIES probability sampling design, including determination of the sampling unit, computation of unit-level inclusion probabilities for the implement probability proportional to size sampling, stratification, allocation, and sample selection and validation. Throughout, we selectively highlight specific challenges of developing a multi-purpose business survey whose collection covers a wide range of economic sectors. The paper concludes with a brief discussion on post-sampling activities designed to improve the quality of the tabulated domain estimates, including calibration weighting and proposed imputation methodology.

### **Developing a Sampling Frame for the Annual Integrated Economic Survey (AIES) in the Economic Directorate of the U.S. Census Bureau**

Justin Z. Smith, *U.S. Census Bureau*

Adam Smeltz, *U.S. Census Bureau*

James Burton, *U.S. Census Bureau*

In the near future, the economic directorate of the U.S. Census Bureau will begin survey processing for the Annual Integrated Economic Survey (AIES). The AIES will replace six independently designed annual surveys. The primary objective of the AIES is to provide accurate economy-wide estimates common to all industries, as well as a set of industry-specific estimates. The AIES will primarily publish domain estimates with different detailed North American Industry Classification System industry levels for national and subnational estimates. Given that the AIES will be a recurring survey, survey redesign and selection of new samples will be an ongoing requirement to ensure that the AIES sample is representative of the U.S. economy. This paper describes the requirements and challenges of building a sampling frame for the AIES. The coverage criteria will be consistent across all six surveys included in the project and with the Economic Census wherever possible. Additionally, this paper discusses the uses of various frame variables in sampling processes, such as measure of size creation, initial allocation, probability of selection, joint inclusion probabilities, and variance estimation.

### **Winsorizing Low Inclusion Probabilities from an Unequal Probability Sample in a Multi-Purpose Annual Survey**

Nicole Czaplicki, *U.S. Census Bureau*

Yeng Xiong, *U.S. Census Bureau*

The Annual Integrated Economic Survey (AIES) will combine several of the U.S. Census Bureau's current annual surveys into one cohesive survey to better meet the needs of data users and respondents. The AIES sample will be selected using a sequential probability proportional to estimated size (PPES) method with inclusion probabilities accounting for two different levels of tabulation (national and subnational) using

annual payroll as the measure of size (MOS). PPES assumes a relationship between MOS and the variable of interest to assign larger companies higher inclusion probabilities. However, this relationship may not be strong enough at the lower end of the distribution to warrant preserving it for smaller units. Furthermore, sampled units with low probabilities of inclusion will have large sampling weights, increasing the variability of variance estimates. Therefore, we develop a data driven one-sided Winsorization procedure to identify and treat the inclusion probabilities for these units, systematically replacing the original values with averaged (larger) values within sampling strata. For sufficiently large strata, a machine learning clustering approach is applied to determine which inclusion probabilities to replace; otherwise, a standard stratification algorithm is used. Measures are also taken to ensure that the PPES nature of the overall sample design is not violated.

**Thursday, October 27**

**10:30AM**

---

## ***Session H-1: Data Science Improving Statistical Products at the National Center for Education Statistics***

---

Organizer: Chris Chapman, *National Center for Education Statistics*

Chair: Peggy Carr, *National Center for Education Statistics*

Room 140A

### **How to Improve Agency Data Visualization Strategies**

Josue DeLaRosa, *National Center for Education Statistics*

The National Center for Education Statistics has continued to take meaningful progress in increasing the public's understanding education information. These efforts include continued technical and procedural enhancements which support the dissemination of information through easier to digest data visualizations. This presentation will provide an overview of updates to data visualization techniques and procedures used by NCES to boost public understanding. Additionally, this presentation will provide key details about the statistical agency's new handbook for the development of data visualizations.

### **Using Geospatial Mapping Techniques and Data Modeling to Build Better Information**

Douglas Geverdt, *National Center for Education Statistics*

The National Center for Education Statistics (NCES) provides information to the public about each and every one of the country's elementary and secondary schools and school systems every year or every other year. Apart from providing summary information about the schools and school system characteristics such as enrollments and finances, NCES links vast amounts of information to geographies around the schools and school systems to improve understanding of the context in which the institutions operate for research purposes and to help families locate schools that are best for them. Such work requires extensive data matching and modeling operations to build complex data sets that can be used by data mapping tools. The presentation will provide information about recent improvements to the data matching and mapping work of the Center.

### **Mining Paradata to Improve Understanding of Test Taking**

Emmanuel Sikali, *National Center for Education Statistics*

The National Center for Education Statistics (NCES) conducts a large number of complex assessments of student academic skills and knowledge. The most well known of the assessments is the National Assessment of Education Progress (NAEP). Assessments like NAEP have traditionally only captured student responses to assessment questions. However, recent developments in computing power and analytic tools has made collection of vast amounts of paradata, or process data, both feasible to collect and meaningfully analyze. The process data show how quickly students move through assessments, what items they might take more or less time with or return too, and how often students access supporting information and tools available to answer

questions. Analyzing the terabytes of resulting information provides NCES with a great deal more substantive information about student problem solving processes. The nature of the data and how they are collected requires highly nested data structures that makes them inaccessible to traditional statistical software packages. NCES has developed techniques to effectively use such data to give a glimpse of students' interactions with the physical computer upon which tests are taken, the testing environment, and the individual items that comprise the assessment. This presentation will discuss the state of this work.

## **Web Scraping to Improve Establishment Employee Surveys - the Case of the National Teacher and Principal Survey**

Louis Avenilla, *U.S. Census Bureau*

As part of the Census Bureau's goal to use "big data" or "administrative records" to augment the reach of traditional data sources, researchers are utilizing data science methods to explore how alternative data sources can be used to supplement survey data collection. These data science methods include web scraping, or the process of systematically collecting information publicly available on the internet, as an alternative to collecting the same information directly from a survey respondent. Researchers at the Center for Optimization and Data Science (CODS) at the Census Bureau are using web scraping to collect teacher and associated school data from the websites of schools sampled for the National Teacher and Principal Survey (NTPS). Historically, selected schools are asked to provide a roster of teachers, which is used to select teachers to receive a teacher questionnaire. In the academic year following when NTPS is conducted, teachers are also contacted to complete the Teacher Follow-up Survey (TFS), and schools are then asked to provide an updated teacher roster. With web scraping, our goal is to use the school's existing internet presence to obtain teacher roster information for both the NTPS and the TFS, thereby minimizing the response burden on school staff and helping the survey team adjust sampling methodology and recruitment efforts in a more timely and cost effective manner.

**Thursday, October 27**

**10:30AM**

## ***Session H-2: Responsible Creation, Handling, and Use of Sexual Orientation and Gender Identity Data***

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

Chair: Hubert Hamer, *National Agricultural Statistics Service*

Room 140B

### **SOGI Data Collection Principles and Examples**

Robert Sivinski, *Office of Management and Budget*

Questions about sexual orientation and gender identity ask respondents to provide sensitive information about themselves and their private lives. Agencies can draw from a variety of tools and approaches developed through government-wide experience collecting sensitive demographic data to make sure that they are collecting and handling SOGI data responsibly and in a way that maximizes the usefulness of the data.

### **Measuring Sex and Gender Identity in Department of Education Information Collections**

JB Simpson, *Department of Education*

The U.S. Department of Education has dozens of information collections that ask questions about sex or gender. Until recently, most had the binary response options of male or female. As the Department has received input from the public, changes have been made to the response categories for some of these information collection instruments. Identifying a need to coordinate among different offices, the Department's Data Governance Board formed a working group to discuss the potential of a consistent approach while carefully considering inclusivity, privacy, utility, clarity, and interoperability. The working group's intended outcomes were a shared understanding of terms, goals, and the current state of the

Department's information collections; the identification of the pros and cons, as well as expected costs and benefits, of potential changes; and the development of guidelines that could be broadly applicable.

## **The Impact of Sexual Orientation and Gender Identity and Disability Questions on Response Rates**

Linda J Young, *National Agricultural Statistics Service*

Barbara Rater, *National Agricultural Statistics Service*

The USDA National Agricultural Statistics Service (NASS) is studying the effect including sexual orientation and gender identity (SOGI) and disability questions might have on respondents' willingness to respond to the 2027 Census of Agriculture. As a first step in this process, NASS initiated an experimental study to assess the potential effect that the presence of these questions might have on unit and item response rates. A randomized sample of about 75,000 was drawn from the NASS's list frame of all active U.S. farms. A survey was administered by mail, on-line and over-the-phone. The four treatment groups were (1) control (none of the test questions were included); (2) only disability questions; (3) only SOGI questions; and (4) both disability and SOGI questions. A four-page questionnaire was used for all treatment groups, with some questions being deleted from the control questionnaire to provide space for the treatment (SOGI or disability) questions. The results of this study and their potential implications for the 2027 Census of Agriculture in particular and the use of SOGI and disability questions more broadly will be discussed.

## **Development of an X Gender Marker Definition for the U.S. Passport: Methodology and Implications for Asking about Gender**

Kristin Miller, *National Center for Health Statistics*

To include an X gender marker on U.S. Passports, The U.S. Department of State partnered with the Collaborating Center for Questionnaire Design and Evaluation Research at the National Center for Health Statistics to assess how potential passport applicants would interpret different versions of the definition as well as how various options would inform response choices. A total of 100 in-depth interviews were conducted with respondents filling out mock-up versions of the passport application, each with different parenthetical definitions of 'X' appearing next to the customary options of 'M' (male) and 'F' (female). Interviews focused on processes and understandings used by respondents to complete the form, including conceptualizations of 'sex' and 'gender' and relationships to their own identities. Analysis suggests that the primary driver of respondents' interpretations, concerns and, ultimately, their preference for a particular definition was in how they viewed the purpose and context of the application. Specifically, the passport would serve as a legal document to identify and affirm themselves as well as to be publicly presented in such places as airport security and customs, not only in the United States, but in countries with different laws and norms. In this regard, respondents focused on three major factors—inclusivity, privacy/safety, and clarity, themes that would eventually be used to frame the optimal definition. This presentation describes the study and discusses more broadly how the concepts of purpose, context, and privacy, along with the need for inclusion, should inform the ways in which the Federal government asks about gender.

---

**Session H-3: Practical Applications of Web-Scraping**

---

Organizer: Tyler Wilson, *National Agricultural Statistics Service*

Chair: Tyler Wilson, *National Agricultural Statistics Service*

Discussant: Mike Gerling, *National Agricultural Statistics Service*

Room 143A

**Comparing Web Scraped Establishment Survey Frames of Industrial Hemp Growers in Seven States: Costs, Contact Data, and Accuracy of Frame**

Samuel Chad Garber, *National Agricultural Statistics Service*

Mike Gerling, *National Agricultural Statistics Service*

Katherine Vande Pol, *National Agricultural Statistics Service*

Tyler Wilson, *National Agricultural Statistics Service*

The National Agricultural Statistics Service (NASS) recently conducted USDA's first national survey of Industrial Hemp Growers. To create the list frame of Industrial Hemp Growers, NASS began developing an in-house web-scraped list frame. A few months later, NASS hired a contractor with extensive web scraping experience to also create a web-scraped list frame in the hopes of obtaining additional records. NASS's in-house approach utilized a hybrid approach of manual and automated web scraping while the contractor relied more heavily on an automated approach. In this paper, the number of records scraped and the manner in which the data were cleaned, including identifying duplicate records and the handling of partial information records, will be compared. The resulting differences in quality of the web-scraped data from the two approaches will be evaluated. The implications that the methodological insights and practical observations have for other applications within the survey framework will be highlighted.

**Social Media as an Early Data Source for Emerging Substance Use Trends**

Alexander Preiss, *RTI International*

Mark Edlund, *RTI International*

Peter Baumgartner, *Explosion AI, GmbH*

Georgiy Bobashev, *RTI International*

Efforts to prevent and treat substance use disorders require timely information on substance use trends. Social media data are real-time, unstructured, and often anonymous, filling some of the gaps of traditional, survey-based data collection. In this study, we test the hypothesis that trends in discussion of substance use topics on the social media platform Reddit correlate with ground truth trends in their prevalence. We scraped millions of Reddit posts and comments from 2016 to 2020 using open-source tools and developed a word embedding-based keyword expansion tool to identify posts and comments discussing a given topic. We show that trends in discussion of several topics correlate strongly with ground truth trends drawn from various national surveys and databases. This suggests that social media data, along with advances in web scraping and natural language processing, can provide valuable leading indicators for outcomes which are unknown due to their novelty or delays in data collection and processing.

**Web Scraping to Collect U.S. School District Level COVID-19 Mask Mandate Information for Contextualizing Results from the National Assessment of Educational Progress**

Brian Cramer, *National Center for Education Statistics*

Sadaf Asrar, *Optimal Solutions Group*

NCES has been collecting online information about COVID-19 mask mandates in the U.S. at the school district-level using artificial intelligence to help understand the upcoming release of the achievement results from the 2022 National Assessment of Educational Progress administered by the National Center for Education Statistics. To collect the data, a customized web scraping tool using the Python package BeautifulSoup was

developed to automate Google searches of mask mandates in school districts in the U.S. These automated searches retrieved a pre-determined number of top search results in a tabular format. Next, Natural Language Processing (NLP) code was developed that read the search results and classified which school districts had implemented a mask mandate. This classification was achieved by developing and training a supervised machine learning algorithm using the search results data that were manually labelled by the authors. School district level mask mandate information for the state of Ohio was used to successfully pilot the tool. The algorithm trained using this data classified which school districts had implemented a mask mandate with an accuracy of 87 percent. The data predicted by the algorithm was also used to help validate the same data collected by ED through monthly surveys of a sample of public schools from December 2021 - January 2022. This presentation will discuss these methods as well as some of the challenges, refinements, and successes of the tool. Ultimately, the results of the tool demonstrate that large scale data collection and validation activities can be conducted with high accuracy at a low cost and can be repeated more frequently and quickly than surveys can without incurring any additional burden on potential respondents.

## **Web Scraping in Support of the U.S. Census Bureau's Public Sector Programs**

Hector Ferronato, *U.S. Census Bureau & Reveal GC*

Brian Dumbacher, *U.S. Census Bureau*

The U.S. Census Bureau conducts many surveys of state and local governments to collect data on government employment and finances. For these surveys, respondent data or equivalent-quality data can sometimes be found on government websites. Analysts currently obtain data from these sources manually, so an automated approach would be very useful. A long-term solution must handle various online formats, structures, and content. To this end, the Census Bureau is developing web scraping methods that find documents from multiple URLs, convert the documents to text and Excel files, and perform large-scale scraping using predetermined key terms. This work is being performed using the SABLE (Scraping Assisted by Learning) web scraping environment, which is based on the open-source software Python. This report details current web scraping efforts in support of State Government Finances and the Quarterly Summary of State and Local Tax Revenue.

**Thursday, October 27**

**10:30AM**

---

### ***Session H-4: If You Can't Beat Them, Join Them: Leveraging Administrative Data to Evaluate Survey Responses and Enhance the Value of Survey Data***

---

Organizer: Scott Boggess, *U.S. Census Bureau*

Chair: Scott Boggess, *U.S. Census Bureau*

Discussant: Katie Genadek, *U.S. Census Bureau*

Room 143BC

## **SNAP Receipt in the SIPP: Using Administrative Records to Evaluate Data Quality**

Michael D. King, *U.S. Census Bureau*

Katherine G. Giefer, *U.S. Census Bureau*

Veronica L. Roth, *U.S. Census Bureau*

The Survey of Income and Program Participation (SIPP) asks about Supplemental Nutrition Assistance Program (SNAP) receipt at the monthly level in an event history calendar. Between 2013 and 2019, the SIPP captured higher rates of annual SNAP receipt among households than the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) or the American Community Survey (ACS). However, SIPP estimates still fell below administrative benchmarks by between 3 and 6 million households. Prior research matching the CPS ASEC and ACS to state-level SNAP records found roughly 40 percent of survey respondents did not report receipt found in administrative records. Little, if any, work has linked recent SIPP data to state SNAP records. We link administrative records from select states to the SIPP for years 2013 to 2019. The paper evaluates annual receipt comparing administrative records to survey reports. The paper also assesses accuracy of months of receipt between the two sources. The SIPP, unlike the CPS ASEC and ACS, collects

monthly information making this a unique contribution. This is preliminary research that we hope to extend to improve imputation of SNAP receipt in the SIPP.

## **Household Indebtedness According to the Spanish Survey of Household Finances and the Central Credit Register: A Comparative Analysis**

Sandra García-Urbe, *Banco de España*

Laura Crespo, *Banco de España*

Olympia Bover, *Banco de España*

The aim of this study is to analyze the quality of the information on indebtedness gathered by the Spanish Survey of Household Finances (EFF). To this end, we match EFF data with the administrative data from the Central Credit Register (CIR), which every month details all outstanding loans in excess of €6,000 arranged by individuals with financial institutions in Spain. Given the differences between the two sources in terms of the information they gather, we construct and compare various measurements of household indebtedness. In order to minimize the differences associated with the discrepancies in household composition according to the municipal population register and the EFF, we analyze both the total linked sample and a subset of comparable households. Our findings show that, after controlling for the limitations of the link, indebtedness calculated with the EFF and the CIR is similar. 25.8% of households have mortgage debt according to the EFF, versus 29.9% according to the CIR. Within indebted households, the median mortgage debt recorded in the EFF is only 0.5% lower than the figure according to the CIR. Non-mortgage debt differences are bigger, but not substantial. 18% of households have non-mortgage debt according to the EFF, versus 23% according to the CIR, and the median debt is 10% lower in the EFF. Moreover, the detailed information provided by the survey on the characteristics of households and their respective debts makes it possible to identify the age of the reference person and the existence of debts shared with individuals who are not members of the household as being the characteristics that have the most bearing on the discrepancies between the EFF and the CIR. The findings of this analysis will help improve the gathering of information and the protocols for interviewing households for the EFF.

## **Concordance Between Self-Report of Medicaid Enrollment in the National Health Interview Survey and Medicaid Administrative Records**

Jessie L Parker, *National Center for Health Statistics*

Crescent B Martin, *National Center for Health Statistics*

The National Center for Health Statistics has a robust data linkage program that links survey data with administrative records. Recently, data from the National Health Interview Survey (NHIS), a population-based, nationally representative survey, were linked to research-optimized analytic files from the Transformed Medicaid Statistical Information System (T-MSIS) released by the Centers for Medicare and Medicaid Services (CMS). This linkage utilized an enhanced algorithm that blends probabilistic and deterministic techniques. In this paper, the linked NHIS-T-MSIS data are examined to compare respondent-level concordance between Medicaid and Children's Health Insurance Program (CHIP) enrollment as reported in NHIS and as indicated by administrative data from T-MSIS for calendar years 2016-2018. Describing concordance in this new linked data source by key subgroups is a critical step in demonstrating the value of the linked NHIS-T-MSIS data for public health researchers and survey methodologists.

## **Building Scientific Publication Profiles for U.S. Trained Doctorate Recipients**

Haoyi Wei, *National Center for Science and Engineering Statistics*

Eric Livingston, *Elsevier*

Christina Freyman, *National Science Foundation*

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

Author profiles of scientific publications are data assets with increasing potential to advance understanding of scientific research and impact made by individual researchers. When complete publication history can be constructed and linked to author attributes such as educational background, employment characteristics, and demographics collected in a longitudinal survey of STEM research doctorate holders, the linked data become a powerful resource for assessing the nation's strategic investment in doctoral training and informing science

policies. Two independent large scale data linkage projects matching respondents of the Survey of Doctorate Recipients (SDR), to two leading publication databases, Web of Science (WoS) and Scopus, have generated a wealth of data on researchers' publications and their career trajectories with different level of successes. In this paper, we develop methods to borrow strength from the SDR-WoS matched set and use auxiliary data from the SDR-Scopus matching process to improve the SDR-Scopus linkage. We then assess resulting SDR-Scopus matches using gold standard data. The overlap between the two matched sets is used to investigate the property of the Scopus author ID and the effect of splitting and lumping error on precision and recall. We conclude with illustrative applications and proposing a cost-effective approach to accelerate and expand the linkage when new sample and publication data become available.

**Thursday, October 27**

**10:30AM**

---

### ***Session H-5: Web Survey Design Standards for the Enterprise Data Collection System at the U.S. Census Bureau***

---

Organizer: Elizabeth Nichols, *U.S. Census Bureau*

Chair: Lora Rosenberger, *U.S. Census Bureau*

Discussant: Jean Fox, *Bureau of Labor Statistics*

Room 144ABC

The U.S. Census Bureau is developing an enterprise-wide system to streamline survey development across data collection modes and survey program areas. Survey analysts will utilize a questionnaire design system to input all survey content and metadata, and then with the push of a button, a survey instrument can be rendered, including web, CATI, CAPI, and paper questionnaires. These survey instruments will adhere to a standard designed to accommodate all the surveys and censuses the Census Bureau administers. This presentation will provide a brief overview of the enterprise-wide solutions for the electronic survey design component. Then, attendees will hear from the Web Standards Team about their approach to developing guidelines for the design of the web surveys. The guideline for each topic was created through a variety of methods, including summarizing empirical literature, conducting split panel studies, compiling evidence from usability studies, and relying on the expertise of usability websites including the U.S. Design Web System. Challenges to crafting the guidelines and suggestions for avoiding impasses will also be discussed.

#### **Data Ingest and Collection for the Enterprise**

Leslie Miller, *U.S. Census Bureau*

This presentation will provide a brief overview of the enterprise-wide solutions for the electronic survey design component, and the proposed timeline for the solutions to be operationalized. Discussion topics will include creation of Web standards sub-team charged with creating the guidelines, the methods used to create the guidelines, and the process to socialize these guidelines with staff responsible for each survey. Challenges to crafting the guidelines and suggestions for avoiding impasses will be discussed.

#### **Linear Navigation Guideline**

Elizabeth Nichols, *U.S. Census Bureau*

Several empirical studies have examined linear navigation design for web surveys. Survey navigation is also a focus during usability testing. To create the linear navigation guideline, we relied on these prior studies, usability pretesting experience, and new A/B studies conducted iteratively.

#### **Edit Validations**

Shelley Feuer, *U.S. Census Bureau*

Most current Census Bureau surveys have validations built into them. Validations trigger a message that appears to respondents with instructions to correct an entry. The functionality of the messages can present

programming challenges and frustration for the respondent. They are also difficult to test experimentally. We rely heavily on the U.S. Design Web System and the expertise of usability websites to form the validation message guideline.

## **Web Survey Branding**

Jonathan Katz, *U.S. Census Bureau*

The Census Bureau has current branding standards. Working with our Public Information Office, we designed a number of options for the header and footer for our surveys. We then embedded these options in an A/B study to gather respondent feedback. Heatmaps were used to gain information on where respondents expected features such as the language toggle, contact us, or FAQs. Several pages that meet the branding standards are offered for the header and footer based on this work.

## **Use of Progress Indicators**

Rachel Horwitz, *U.S. Census Bureau*

Numerous studies have examined the use of progress indicators in a survey. Progress indicators did not improve response rates in any of the studies, and in longer surveys, they were shown to increase breakoffs. However, progress indicators often pique the interest of survey sponsors. Our progress indicator guideline is based on the literature. However, we offer a design compromise for survey areas whose sponsor is interested in showing progress to respondents.

## **Personalizing Questions**

Erica Olmsted-Hawala, *U.S. Census Bureau*

Web surveys offer the ability to personalize questions. On the paper questionnaire, annual wage information is requested, but it might be easier for people to report daily, weekly, bi-weekly, or monthly wage instead. In A/B studies we compared different designs where the respondent selects the appropriate unit to see if one design is more usable than another. Based on these experiments, we offer a web survey guideline for questions where the respondent is offered the ability to designate the unit of measurement.

## **Input Field Formatting**

Temika Holland, *U.S. Census Bureau*

Formatting input fields is an important element in forms design. The team successfully used A/B studies to determine an optimal design for fixed-length data but experienced challenges when trying to design an experimental study to test input fields to collect monetary data for our business surveys. We rely on usability testing experience and current designs to form this guideline.

**Thursday, October 27**

**1:30PM**

---

***Session I-1: Foreign-Born Scientists and Engineers and the U.S. Workforce Evidence Building***

---

Organizer: Michael Atkinson, *America's Datahub Consortium*

Chair: John Finamore, *National Center for Science and Engineering Statistics*

Room 140A

As more countries offer their students reasons to stay in their own country for their education or to return home after earning a degree, the U.S. could face a shortage in a critical segment of its workforce." - (National Science Board, 2020) Gaining a better understanding of foreign-born scientists and engineers (FBSEs) in the U.S. is critical for policymakers and others. They are a key part of the U.S. science and engineering enterprise--both the number and proportion of such workers in U.S. science and engineering jobs have risen over time. Additionally, the training, education, and employment of this population in the U.S. affects not only our nation but also the global science and engineering enterprise. Finally, understanding what FBSEs do after they are trained is important to understanding the return on investment (ROI) the U.S. reaps from supporting their education and training. Data Gaps: There is currently limited information on scientists and engineers who are born and trained in foreign countries. Although we have information on FBSEs who earn their doctorates in the U.S., we lack information on FBSEs who come to the U.S. after their training, as well as those who come to the U.S. and earn associate's, bachelor's, and master's degrees but do not go on to complete a doctorate (including those who are part of the skilled technical workforce). We also lack information for FBSEs at all levels in terms of their patterns for staying in the U.S. after their education/training and their international mobility if they leave the United States. Key Policy-relevant context: Filling these data gaps by answering some key questions will ultimately provide data-based evidence about the impact of FBSEs on the U.S. scientific enterprise. NCSES has funded five efforts through the America's Datahub Consortium. Each team will provide an overview of their project to date and lessons learned in evidence building efforts. The presentations and its authors are as follows:

**Foreign-Born Scientists and Engineers and the U.S. Workforce Evidence Building  
Overall Effort and Goals**

John Finamore, *National Center for Science and Engineering Statistics*

**A Network Scale-Up and Ad-Tech Approach Examining Workforce Participation and International Mobility in the FBSE Population**

Ian McCulloh, *Accenture Federal Services LLC*

**Foreign-Born Scientists and Engineers and the US Workforce**

Anand Desai, *Clarivate*

**Creating a New Data Infrastructure for Foreign-Born Scientist and Engineers: Data, Analysis and Use**

Ahu Yildirmaz, *Coleridge Initiative*

**Building an Evidence-Based Foundation to Understand Foreign-Born Scientists and Engineers' Participation in the US Workforce**

Brandon Sepulvado, *NORC at the University of Chicago*

**Foreign-Born Scientists and Engineers and the U.S. Workforce**

Caren Arbeit, *RTI International*

**Thursday, October 27**

**1:30PM**

---

***Session I-2: A Marshall Plan for Reconstructing National Statistical Offices After Conflict: Practical Guidance from International Principles***

---

Organizer: Jennifer Park, *Committee on National Statistics*

Chair: Dominik A. Rozkrut, *Statistics Poland*

Discussant: Albert Kroese, *International Monetary Fund*

Room 140B

Remarkably, the National Statistical Office of Ukraine has continued to operate since the beginning of the Russian-Ukrainian war. This resilience is a testimony to the essential nature of objective, accurate, reliable, and timely national statistics to inform policy-making, and the steadfastness of the national statisticians behind the numbers. The healthy functioning of a national statistical office has implications for its relationships with bi-lateral and multi-lateral agreements with donor countries and organizations, and therefore, for the security of its country.

This year marks the anniversary of the Fundamental Principles of Official Statistics (1992), championed by many esteemed thought leaders, including Józef Oleński (former President of the Polish Central Statistical Office), Jean-Louis Bodin (INSEE, France), and Katherine Wallman (former, U.S. Chief Statistician and chair, UNECE CES), within the context of the United Nations Economic Commission for Europe Conference of European Statisticians as a way to support the production of national statistics among countries transitioning from centrally planned economies to market economies. The Fundamental Principles of Official Statistics subsequently was endorsed by the highest body of the UN, the General Assembly (2014).

There have been additional efforts to develop aspirational and practical guidance for national statistical offices to strengthen capacity. The European Statistics Code of Practice (2005), U.S. Office of Management and Budget's Statistical Policy Directive 1 (2014) (now embedded in the Evidence Act), OECD's Recommendation of the Council on Good Statistical Practice (2015), American Statistical Association's Ethical Guidance for Statistical Practice (2022) are but a few. There have also been efforts to develop implementation guidance for these principles; notably, FPOS (2011, 2015, and 2020).

After the Second World War, the Marshall Plan was implemented to assist in the reconstruction and strengthening of nation states affected by conflict. What would a Marshall Plan for national statistics in Ukraine look like? How can the parameters of the FPOS and other aspirational guidance inform practical steps of such a plan? What roles could various entities take to implement such a plan? What elements are essential in the short term? Over the longer term?

**Panelists:**

Włodzimierz Okrasa, *Cardinal Stefan Wyszyński University in Warsaw*

Misha Belkindas, *Open Data Watch Consulting, IAOS President*

Ron Wasserstein, *American Statistical Association*

**Thursday, October 27**

**1:30PM**

---

***Session I-4: Developing the Infrastructure to Build Data Infrastructure:  
The Decennial Census Digitization and Linkage Project***

---

Organizer: Katie Genadek, *U.S. Census Bureau*

Chair: Katie Genadek, *U.S. Census Bureau*

Discussant: David Johnson, *National Academy of Science*

Room 143BC

**Merging Newly Digitized Census Form Images with Previously Created Microdata**

John Sullivan, *U.S. Census Bureau*

Katie Genadek, *U.S. Census Bureau*

J. Trent Alexander, *University of Michigan*

Sohall Kamdar, *University of California, Berkley*

We developed linkage algorithms for merging newly recovered data from decennial census images to original microdata captured at the time the census was taken. While in one decennial census year (1990) the linkage is straight forward with a given linkage key, in the remaining years (1960-1980) the linkage is probabilistic and is based on various geographic and household information available in the newly captured and older data. This presentation will discuss the development of these linkage algorithms and the creation of the programs to be used for linking 850 million respondents' newly captured information to the original data.

**Communication and Iterative Development of Training Data**

Kelsey Drotning, *U.S. Census Bureau*

Cecile Murray, *U.S. Census Bureau*

An important component of training the machine learning algorithm for the handwriting data capture for the DCDL project was hand keying a small sample of images from each year which was made possible by the keying staff at the National Processing Center (NPC). To start, a small group of keyers tested the keying application, provided feedback, and tested changes that were made in response. Later, a large group of keyers was brought on at the same time. In order for this project to be a success, we utilized a Microsoft Teams channel where NPC and Econ staff could work through issues and ask questions in real time. This presentation will provide an overview of the development of this working relationship.

**Creating Comparable Geography Across Data files for Record Linkage**

David Bleckley, *University of Michigan*

J. Trent Alexander, *University of Michigan*

Katie Genadek, *U.S. Census Bureau*

The DCDL project will adapt well-established techniques to conduct record linkage across the 1960-1990 Censuses, using a modified version of the Census Bureau's production record linkage processes. These processes rely on residential geographic location in census and survey data and proximate administrative records, primarily using address. The 1990 Census includes full street address. However, in 1970 and 1980, the census data does not have full street address. This presentation will describe the harmonization of address and geographic records in from 1969 and 1979 tax records to geographic areas identified on each year's census data. The various data methods used to perform this work will be discussed, and the resulting geographic crosswalks and planned uses for linkage will also be shown.

**Big Data Processing and Storage using On Premise and Cloud Solutions**

Jonathan Morgan, *U.S. Census Bureau*

Katie Genadek, *U.S. Census Bureau*

Junied Sheik, *U.S. Census Bureau*

The DCDL project is in the process of creating 4 Petabytes of digital images. These data are stored on

premises and the information from the images are being captured simultaneously on cloud-based servers (for handwriting capture) and on premises servers (for bubbled data capture). This presentation will describe the data flow and processes for supporting and managing this work. We will discuss lessons learned while developing the storage and processing solutions for this big data project, and the resulting combination of manual and automated processes.

**Thursday, October 27**

**1:30PM**

---

### ***Session I-5: Innovations in Health Care Survey Data Collection***

---

Organizer: Stephanie Coffey, *U.S. Census Bureau*

Chair: Stephanie Coffey, *U.S. Census Bureau*

Room 144ABC

#### **Using Data from the National Health and Nutrition Examination Survey (NHANES) to Support Health Equity Research**

Denys Lau, *National Center for Health Statistics*

Cheryl Fryar, *National Center for Health Statistics*

Namanjeet Ahluwalia, *National Center for Health Statistics*

Matthew Kaufman, *National Center for Health Statistics*

Irma Arispe, *National Center for Health Statistics*

Ryne Paulose-Ram, *National Center for Health Statistics*

Health equity research looks beyond the existence of health disparities towards upstream adverse social determinants of health (SDOH) and systemic factors, such as racism, discrimination, and inequitable distribution of resources including fair pay, healthy food, quality education and housing, transportation, and safe neighborhoods. These factors drive and manifest health disparities among racial and ethnic minorities, and other disproportionately affected populations. NHANES is a complex, multi-stage, nationally representative sample survey that monitors the health and nutritional status of civilian, non-institutionalized adults and children across the United States. NHANES uniquely combines interviews with physical examinations, body scans, and biospecimen collection to identify chronic and infectious diseases, and assess nutritional status, physical activity, and environmental exposures. This presentation examines how NHANES data can be used to support research on SDOH and their disproportionate impact on various populations. Specifically, the presentation will describe (1) how various components of 1999-March 2020 NHANES – survey items, biomarkers/physical measurements, sampling design, and data linkage – support health equity and SDOH research; and (2) discuss future opportunities that can be explored to increase NHANES capacity to advance health equity science. NHANES has supported health equity research by including survey items on perceived SDOH, oversampling subgroups disproportionately affected by SDOH, and linking NHANES to external data sources on structural inequities and built environments. This paper will highlight resources within NHANES that can inform researchers and policymakers about health equity and SDOH.

#### **“Second Chance” Respondents to the National Survey on Drug Use and Health: A Natural Reinterview Reliability Analysis**

Peter Frechtel, *RTI International*

Lauren Klein Warren, *RTI International*

Adam Lee, *RTI International*

Due to the COVID-19 pandemic, the National Survey on Drug Use and Health (NSDUH) offered web response mode for the first time beginning in 2020. Of the approximately 70,000 respondents in 2021, 567 began the web survey, broke off at some point, and later completed a face-to-face interview. Due to the sensitive nature of the data, the incomplete web data were no longer available to these respondents upon subsequent entry into the survey. More than half of these respondents provided enough information via the web to be used in a reinterview reliability analysis to assess both mode and data quality differences. Preliminary results for substance use suggest that a nontrivial subset of these “second chance” respondents answered inconsistently in unexpected ways. We present analyses of inconsistent responses and consider some ways we can take

advantage of the research opportunities these cases provide. Possibilities include (1) determining methods for identifying intentional and unintentional changes to survey question responses, and (2) determining what these observed changes to survey responses mean for the follow-up protocol for web breakoffs.

## **Need Telephone Show Cards For Your In-Person Survey Due To A Pandemic-Induced Multimode Data Collection Shift? Path Study Lessons Learned 2020 to Present...**

Sarah Dipko, *Westat*

Victoria Castleman, *Westat*

The Population Assessment of Tobacco and Health (PATH) Study is a nationally representative longitudinal study which has interviewed respondents in-person since its inception in 2012. The study selects youth and adults for annual or biennial audio computer-assisted self-interviews (ACASI). These interviews contain a large volume of questions, many of which include on-screen images to illustrate a variety of tobacco products. Parent CAPI interviews also use on-screen images to provide response option lists for sensitive questions. In-person operations were suspended in March 2020 due to the COVID-19 pandemic. Work began immediately to transition to telephone administration. Show cards were developed to support telephone respondents through the new interviewer-administered response process, including creation of a nimble, flexible web design.

The PATH Study was able to resume data collection via telephone administration in July 2020. This paper reviews the issues and decisions involved in the initial creation of the PATH Study show cards, highlights how the instrument flow informed the web design, and notes the improvements made for subsequent rounds of the study.

## **Zooming into the Future with CAVI: An Update**

Jesus Arrue, *Westat*

Darby Steiger, *Westat*

Lena Centeno, *Westat*

Cindy Good, *Westat*

During COVID, Computer-Assisted Video Interviewing (CAVI) emerged as a new mode of data collection (Schober, et al. 2020) and has been found to offer similar benefits to in-person interaction (West, et al. 2021). The Medical Expenditure Panel Survey (MEPS) has traditionally been conducted as an in-person survey, with heavy reliance on show cards and medical records. During COVID, MEPS quickly shifted to telephone, and then returned to in-person as soon as it was feasible. CAVI was introduced in the spring 2022 cycle as a primary option for households that preferred not to have face-to-face interaction. Over 300 field interviewers were trained on CAVI over a six-week period in early 2022 (Steiger, et al, 2022).

As interviewer training progressed, weekly monitoring reports tracked the implementation and adoption of CAVI in comparison to in-person and telephone modes. In addition, interviewers provided qualitative feedback through focus groups and debriefing surveys to report on implementation successes and challenges. This paper presents results of our first data collection cycle of CAVI implementation on MEPS, interviewer feedback on the mode, and next steps for continued research.

Thursday, October 27

3:30PM

---

## ***Session J-1: Evolution of the U.S. Economic Census for 2022 – An Overview of Changes, Testing, and Innovations Employed to Support Continued Improvement***

---

Organizer: Michael Sprung, *U.S. Census Bureau*

Chair: Kimberly Moore, *U.S. Census Bureau*

Discussant: Brent Hueth, *Economic Research Service*

Room 140A

### **Assessing Stakeholder Data Needs**

Michael Sprung, *U.S. Census Bureau*

Meghan Harrison, *U.S. Census Bureau*

A comprehensive assessment of the industries and content captured in the Economic Census was conducted leading up to the 2022 cycle. This involved research on new industries and products combined with a concerted effort to identify, communicate with, and obtain input from stakeholders. The content changes implemented as a result of this endeavor will ensure the Economic Census remains relevant in 2022 and beyond.

### **Using Iterative Quick Turnaround Cognitive Interviewing to Integrate Quality Into the 2022 U.S. Economic Census**

Kristin Stettler, *U.S. Census Bureau*

Melissa Cidade, *U.S. Census Bureau*

To prepare for the 2022 Economic Census, the Census Bureau consulted with experts both inside and outside of the Federal government about content. Their recommendations were evaluated, and new proposed content underwent cognitive testing to ensure the questions met Census Bureau quality standards. Given the number of topics (19), the number of questions (65), and the need to get feedback from businesses of varying size and industry classification, the process for setting up and conducting cognitive interviews was complex. A total of 230 phone interviews over 3 rounds of testing were conducted over 4 months. This intensive and iterative testing strategy helped to improve the quality of the questions included in the 2022 Economic Census

### **Implementing Interactive Classification Tools**

Emily Wiley, *U.S. Census Bureau*

Daniel Whitehead, *U.S. Census Bureau*

Improving classification and modernizing the data review process are high priorities for the 2022 Economic Census. The development of dynamic industry and product classification tools for the electronic collection instrument supports both of these goals. Testing has shown these interactive machine learning applications will provide high quality classification information while reducing labor intensive coding of respondent provided descriptions.

### **Leveraging Respondent Input to Improve Program Outreach**

Charles Brady, *U.S. Census Bureau*

Ensuring that outreach materials reach and resonate well with their intended audiences is a critical component to successful data collection and dissemination. A series of market specific focus groups were conducted to inform the development of a comprehensive outreach plan designed to reach the diverse industries and locations included in the Economic Census.

---

**Session J-5: Analyzing Respondent Behavior and Data Quality in Web Surveys**

---

Organizer: Jessica Holzberg, *U.S. Census Bureau*

Chair: Jessica Holzberg, *U.S. Census Bureau*

Room 144ABC

**Developing a Self-Administered Eldercare Module for the American Time Use Survey**

Zoe Grotophorst, *NORC at the University of Chicago*

Stephanie Denton, *Bureau of Labor Statistics*

Lisa Lee, *NORC at the University of Chicago*

Justine Bulgar-Medina, *NORC at the University of Chicago*

The Bureau of Labor Statistics (BLS) introduced an eldercare module to the American Time Use Survey (ATUS) in 2011 with the goal of collecting quality data on eldercare provision in the United States. To date, BLS has conducted this module solely via Computer Assisted Telephone Interviewing (CATI). Now that BLS is developing a self-administered, web-based mode for the ATUS, there is also a need to develop a web version of the eldercare module. As part of this process, BLS and NORC at the University of Chicago conducted cognitive testing to find out whether respondents could, without the help of a telephone interviewer, self-identify as eldercare providers or non-providers and if they could accurately report their time spent providing eldercare. BLS and NORC programmed a web version of the module and conducted 50 cognitive interviews with respondents over Zoom. Interviews were conducted in two rounds, with refinements made between rounds based on initial results. This presentation will share the results from the cognitive testing, highlighting key changes made to enable accurate data collection in the self-administered web module.

**Analyzing Survey Duration Time and Its Relationship to Data Quality for a Household Survey**

S. Grace Deng, *Energy Information Administration*

L. Kaili Diamond, *Energy Information Administration*

The Residential Energy Consumption Survey (RECS) is a nationwide study of household energy characteristics and consumption use. This paper analyzes how survey duration time varies by household characteristics and its relationship to data quality for the 2020 RECS, which was conducted entirely via self-administered Web and paper modes. Survey duration is the amount of time it takes respondents to complete a survey. It was captured as part of the paradata via the 2020 RECS web instrument. Our analysis compared duration times across key demographic and other housing unit characteristics, including respondent age, education level, employment status, and housing type. We also analyzed questionnaire section timing and explored how duration time was related to data quality using editing failures. There are about 11,000 Web cases included in this analysis. Preliminary results indicate that respondent age has the largest effect on the duration time, with younger respondents completing the web survey in a shorter amount of time than older respondents. Younger respondents were also more likely to respond to the survey using a mobile phone. However, longer duration time does not seem to be an indication of better data quality, but may be related to other factors such as respondent knowledge on certain questionnaire items and/or the need to look up information. We provide suggestions on conducting additional pretesting to understand how to speed up the duration time for older respondents, as well as to improve the design of certain questionnaire items.

## **Itemization of Reported Expenditures from the CE's Test of a Fully Self-Administered Diary**

Graham Jones, *Bureau of Labor Statistics*

Nikki Graf, *Bureau of Labor Statistics*

Tucker Miller, *Bureau of Labor Statistics*

Wendy Carlton, *Bureau of Labor Statistics*

From November 2021 to January 2022, the Bureau of Labor Statistics's Consumer Expenditure (CE) Survey conducted a test of a self-administered diary in a probability-based online panel. The purpose of the test was to gauge the impact of shifting from the current diary placement protocols by analyzing results from a two-week online diary survey administered without the aid of an interviewer. This presentation will examine the quality of data provided in the test diary by comparing rates of reported expenditure itemization, as well as median survey time, against CE production data in the first two months of fielding. For this research, itemizing is defined as providing items purchased as individual entries, as opposed to reporting the receipt total of a shopping trip with a vague description like "groceries" or "clothes." Itemizing is generally associated with higher data quality, and CE research has found that online cases are associated with higher rates of allocation edits than in paper cases, which are applied when respondents provide insufficient detail to meet tabulation requirements. Rates of itemization by expenditure type, device type, and demographics will also be examined.

## **It's a Trap!: Use of Trap Questions and Data Quality**

Mina Muller, *Ipsos Public Affairs*

Frances M. Barlas, *Ipsos Public Affairs*

Randall K. Thomas, *Ipsos Public Affairs*

Megan A. Hendrich, *Ipsos Public Affairs*

Researchers often have concerns over data quality due to inattentive or unmotivated respondents. Various measures have been developed to assess whether respondents are providing accurate responses. Besides speeding and response non-differentiation, researchers sometimes imbed trap questions that can be used to detect when someone is not paying attention. A compliance trap directs a respondent to select a particular response (e.g., "Select 'Somewhat agree' for this item") regardless of the question. We studied whether cleaning out respondents who fail such traps would improve data quality. In a study with over 3,500 completes from an online probability-based panel, respondents were randomly assigned to experience two compliance traps or not. We examined respondent reactions following their presentation and we also looked at any reduction in bias using 10 demographic items for which we had benchmark values. We also looked at how trap failure was related to speed to complete (another indicator of sub-optimal behavior). For both trap conditions, we found that higher trap failure had a modest association with faster completion times. We found that there was no difference in average bias between the full samples versus the samples that had eliminated participants due to trap failures. It appears that trap failures are not as closely related to data quality as many have believed.

## **Device Usage on the Internet Instrument of the American Community Survey**

Gregory Mills, *U.S. Census Bureau*

Lindsay Longsine, *U.S. Census Bureau*

R. Chase Sawyer, *U.S. Census Bureau*

The internet instrument was first used in production for the American Community Survey in 2013. At that time, it was estimated that around 95 percent of logins were from personal computers (PCs), while 5 percent were from tablets and mobile phones. These figures were based on a 2011 ACS Internet Test. The purpose of this paper is to provide an updated analysis of general device usage trends on the ACS, as much time has passed since such a systemic analysis was performed, and in that time, usage trends of PCs, tablets, and mobile phones in wider society have changed considerably. We use ACS internet instrument paradata, which gives some details about the device type and operating system used by a respondent, paired with case disposition and response data. We first explore the prevalence of various device types and operating systems using the ACS internet instrument over time. We then examine some behavioral patterns that are associated

with device types, such as number of logins, device switching patterns, and likelihood of completing the survey in a single login. Finally, we analyze some demographic characteristics associated with usage of each device type using a logistic regression model. Plans for additional research into device usage on the ACS will also be discussed.