



Federal Committee on Statistical Methodology

2023 Research and Policy Conference

*Building Toward the Future: Strengthening and Expanding
the Capacity of the Federal Statistical Ecosystem*

October 24-26, 2023

**College Park Marriott Hotel and Conference Center
Hyattsville, MD**

Final Program

Hosted By:



COPAFS

Council of Professional Associations on Federal Statistics

Federal Sponsors

The 2023 FCSM Research and Policy Conference received generous support from the following federal agencies.



Sponsors

The 2023 FCSM Research and Policy Conference received generous support from the following companies and organizations.



Exhibitors



FCSM 2023 Research and Policy Conference Program Committee

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

Jessica Holzberg, Co-Chair, *U.S. Census Bureau*

Jessica Graber, Co-Chair, *National Center for Health 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*

Maura Bardos, *Energy Information Administration*

Alisha Coleman-Jensen, *Economic Research Service*

Alexia Cooper, *Bureau of Justice Statistics*

Gavin Corral, *National Agricultural Statistics Service*

Brian Cramer, *National Center for Education Statistics*

Kyle Hood, *Bureau of Economic Analysis*

Diba Khan, *National Center for Health Statistics*

Tamara Lee, *Department of Veterans Affairs*

Richard Levy, *U.S. Census Bureau*

Justin D. Nguyen, *U.S. Census Bureau*

Jennifer Rhorer, *National Agricultural Statistics Service*

Sharon M. Stern, *U.S. Census Bureau*

Ed Strocko, *Bureau of Transportation Statistics*

Suzanne Strong, *Bureau of Justice Statistics*

Valerie Testa, *Statistics of Income*

Julie Weeks, *National Center for Health Statistics*

Doug Williams, *Bureau of Labor Statistics*

Federal Committee on Statistical Methodology

Ann E. Carson, *Bureau of Justice Statistics*

Chris Chapman, *National Center for Education Statistics*

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

Tala Fakhouri, *U.S. Food and Drug Administration*

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

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 Science and Engineering Statistics*

Jennifer Nielsen, *National Center for Education Statistics*

Jennifer M. 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*

Ed Strocko, *Bureau of Transportation Statistics*

G. David Williamson, *Centers for Disease Control and Prevention*

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.

2023 FCSM Research and Policy Conference

The Federal Committee on Statistical Methodology (FCSM) was founded in 1975 by the Statistical and Science Policy Branch (formerly, the Office of Statistical Policy) in the Office of Information and Regulatory Affairs in the Office of Management and Budget (OMB) to assist in carrying out SSP/OMB's role in setting and coordinating statistical policy.

The FCSM serves as a resource for OMB and the federal statistical system to inform decision making on matters of statistical policy and to provide technical assistance and guidance on statistical and methodological issues. The FCSM sponsors regular conferences (in partnership with the Council on Professional Associations on Federal Statistics), hosts seminars and workshops (often in collaboration with partner organizations such as the Washington Statistical Society), writes reports aimed at the federal statistical system and the community, and creates subcommittees and interest groups focused on topics relevant to the federal statistical community. Recent news and resources from FCSM are available at www.fcsm.gov.

The FCSM Research and Policy 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 2023 FCSM Research and Policy Conference is: *Building Toward the Future: Strengthening and Expanding the Capacity of the Federal Statistical Ecosystem*. The conference provides a forum for experts and practitioners from around the world to discuss and exchange current methodological knowledge and policy insights about topics of current and critical importance to federal agencies as well as the Federal Statistical System as a whole.

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

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

Keynote Speaker

Katharine G. Abraham, University of Maryland

Tuesday, October 24

11:00AM

New Challenges and New Opportunities for the Federal Statistical System

Katharine G. Abraham, *Distinguished University Professor, University of Maryland*

Approaching the five-year anniversary of the signing of the Evidence Act, the Federal statistical system faces many challenges but also an exciting new set of opportunities. Under the Act, statistical agency leaders are being asked to play an expanded role as stewards of the data held by the Federal government. Working together and with the government's program agencies—and, in some cases, with holders of private sector data—the agencies can help to better meet the growing demand for reliable information about our society and economy and about the effectiveness of policy interventions to address societal concerns. Embracing these opportunities will increase the agencies' relevance in a rapidly changing world and position the federal statistical system well for the coming decades.

Katharine G. Abraham is a Distinguished University Professor at the University of Maryland. Her published research includes papers on the work and retirement decisions of older Americans; how government policies affect employers' choices concerning employment and hours over the business cycle; the effects of financial aid on the decision to attend college; discrepancies in alternative measures of employment, wages and hours; and the measurement of economic activity. She served as Commissioner of the Bureau of Labor Statistics from 1993 through 2001 and as a Member of the President's Council of Economic Advisers from 2011 through 2013. Abraham currently serves on standing academic advisory committees convened by the Congressional Budget Office, Bureau of Economic Analysis, and Federal Reserve Bank of Chicago. She is a Research Associate of the National Bureau of Economic Research and a Research Fellow of the IZA, a member of the National Academy of Sciences and of the American Academy of Arts and Sciences, a Distinguished Fellow of the American Economic Association, and an elected Fellow of the American Statistical Association and of the Society of Labor Economists. Abraham received her Ph.D. in economics from Harvard University in 1982 and her B.S. in economics from Iowa State University in 1976.

Schedule Overview

Tuesday, October 24

8:00 am – 3:30 pm **Registration Open**

8:00 am – 9:00 am **Continental Breakfast**, *sponsored by Westat*

9:00 am - 10:30 am CONCURRENT SESSIONS A

A-1. Leveraging Probability Online Survey Panels for Federal Data Collection (*Vessey 1*)

A-2. Using Administrative Data for Program Evaluation and Research (*Vessey 2*)

A-3. Perspectives from the COVID-19 Pandemic (*Patuxent*)

A-4. Statistical Products First – A Paradigm Shift for Federal Statistics (*2110*)

A-5. Exploring Machine Learning Techniques (*2100*)

A-6. Application of Scanner Data to Calculate Monthly Area Price Measures (*0105*)

10:30 am - 11:00 am **Coffee Break**, *sponsored by the Population Association of America*

11:00 am - 12:15 pm **PLENARY SESSION** (*Chesapeake Ballroom*)

- Opening Remarks
 - Karin Orvis, *Chief Statistician of the United States*
- Keynote Address
 - Katharine G. Abraham, *University of Maryland*

12:15 pm - 2:00 pm **Lunch on Your Own**

2:00 pm – 3:30 pm CONCURRENT SESSIONS B

B-1. Naturally of Course – Strengthening Statistical Programs Through NLP (*Vessey 1*)

B-2. Making Less Work: Monitoring and Adjusting for Nonresponse (*Vessey 2*)

B-3. Advanced Statistical Techniques for Survey Inference (*Patuxent*)

B-4. Equity in STEM, Health, Poverty, and Race and Ethnicity Identification (*2110*)

B-5. What do We Mean by "For Statistical Purposes Only" (*2100*)

B-6. How the Commodity Flow Survey Made its First 100 Million (*0105*)

3:30 pm – 3:45 pm **Break**

3:45 pm – 5:15 pm CONCURRENT SESSIONS C

C-1. Building Diverse Leadership for the Statistical System (*Vessey 1*)

C-2. Initial Proposals for Updating OMB Statistical Policy Directive No. 15 (*Vessey 2*)

C-3. Modernizing Data Dissemination at the NCHS (*Patuxent*)

C-4. How the Census Bureau is Using Admin Records Across the ACS Life Cycle (*2110*)

C-5. Parsing Policy: Analyzing Large Language Models and Conversational AI (*2100*)

C-6. Differential Privacy Population Estimates and Implications for Public Health (*0105*)

Wednesday, October 25

7:30 am – 3:30 pm **Registration Open**

7:30 am – 8:30 am **Continental Breakfast**

8:30 am - 10:00 am CONCURRENT SESSIONS D

D-1. Building Evidence by Linkage to Data at the U.S. Census Bureau (*Chesapeake A*)

D-2. NLSY97 COVID-19 Supplement (*Chesapeake B*)

D-3. New Perspectives and Methods on Privacy and Disclosure Control (*Chesapeake C*)

D-4. How Can U.S. Professional Guidelines Inform Federal Statistical Policy? (*Vessey 1*)

D-5. Approaches for Improving Survey Response (*Vessey 2*)

D-6. Application of Respondent-Centered Establishment Survey Design (*0105*)

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

10:30 am - 12:00 pm CONCURRENT SESSIONS E

E-1. Advancing Federal Data Collections: Qx Design, Mode, and Contact Strategy (*Chesapeake A*)

E-2. Leveraging Data: Using Administrative and Synthetic Information (*Chesapeake B*)

E-3. Inclusive Insights: Adding SOGI Questions to Large-Scale Data Collections (*Chesapeake C*)

E-4. Synthetic Data Generation for Survey Data (*Vessey 1*)

E-5. Ensuring Scientific Integrity in the Federal Statistical System (*Vessey 2*)

E-6. The Missing Data Puzzle: Exploring Imputation Methods (*0105*)

12:00 pm – 1:45 pm **Lunch on Your Own**

1:45 pm – 3:15 pm CONCURRENT SESSIONS F

F-1. Who's Keeping Score on . . . Data Quality (*Chesapeake A*)

F-2. Dynamics of Data Linkage: Techniques and Challenges (*Chesapeake B*)

F-3. Advancements in Sexual Orientation and Gender Identity Measurement (*Chesapeake C*)

F-4. Democratizing Data: A Search and Discovery Platform for Public Data Assets (*Vessey 1*)

F-5. Testing a New Combined Race and Ethnicity Question in Household Surveys (*Vessey 2*)

F-6. Applications of Modeling and Post Randomization to Improve Estimation (*0105*)

3:15 pm - 3:30 pm **Break**

3:30 pm – 5:00 pm CONCURRENT SESSIONS G

G-1. If We Modernize the Survey, Will They Respond? (*Chesapeake A*)

G-2. Directions in Measuring Disability: Washington Group Short Set (*Chesapeake B*)

G-3. Recent Advances in Data Privacy and Public Policy (*Chesapeake C*)

G-4. Advances in Implementing Title III of the Evidence Act (*Vessey 1*)

G-5. Innovations from the CJARS Project (*Vessey 2*)

G-6. America's DataHub Consortium: Building Capacity for the Future (*0105*)

Thursday, October 26

7:30 am – 3:30 pm **Registration Open**

7:30 am– 8:30 am **Continental Breakfast**

8:30 am - 10:00 am CONCURRENT SESSIONS H

H-1. Business Revenue and Employment During the COVID-19 Pandemic (*Chesapeake A*)

H-2. Linking NCES Data to Other Data Sets (*Chesapeake B*)

H-3. Data Integration Techniques for Improving Economic Statistics (*Chesapeake C*)

H-4. Data Quality: Disclosure, Deidentification, and Publicly Available Data (*Vessey 1*)

H-5. Advancing the Federal Statistical Ecosystem TODAY! (*Vessey 2*)

H-6. What is Consumer Inflation? A Family of Indexes Tells the Story (*0105*)

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

10:30 am - 12:00 pm CONCURRENT SESSIONS I

I-1. Data Driven Insights: Integrated Data from Federal Statistical Agencies (*Chesapeake A*)

I-2. An Updated Measure of Poverty: (Re)Drawing the Line (*Chesapeake B*)

I-3. Disclosure Avoidance in the Survey of Income and Program Participation (*Chesapeake C*)

I-4. Building on GDP: The Future of Economic Statistics (*Vessey 1*)

I-5. Enhancing Survey Programs by Using Multiple Data Sources (*Vessey 2*)

I-6. Innovations in Gathering Health Data: Pilots, Methods, and Lessons (*0105*)

12:00 pm - 1:45 pm **Lunch on Your Own**

1:45 pm – 3:15 pm CONCURRENT SESSIONS J

J-1. Enhancing Food Policy Research Through Administrative Data Linkages (*Chesapeake A*)

J-2. Putting Statistics on the Map (*Chesapeake B*)

J-3. Improving Price Index Measurement: New or Alternative Data Sources (*Chesapeake C*)

J-4. The Importance of Private Sector Data to Federal Statistics (*Vessey 1*)

J-5. Leveraging ML to Improve the Accuracy and Reliability of Official Statistics (*Vessey 2*)

J-6. Federal Statistical Infrastructure for Puerto Rico (*0105*)

3:15 pm – 3:30 pm **Break**

3:30 pm – 5:00 pm CONCURRENT SESSIONS K

K-1. Advances on Federal SOGISC Data (*Chesapeake A*)

K-2. Collecting Race and Ethnicity in Establishment Surveys (*Chesapeake B*)

K-3. Using Data to Overcome Supply Chain Challenges (*Chesapeake C*)

K-4. Measuring the Health of the Federal Statistical System (*Vessey 1*)

K-5. Seize the Data: Using Linked Survey and Administrative Data (*Vessey 2*)

K-6. Advances in Estimation and Forecasting (*0105*)

Abstract Listings for Tuesday, October 24

- **Concurrent Sessions A** 9:00 am – 10:30 am
- **Plenary Session** 11:00 am – 12:15 pm
- **Concurrent Sessions B** 2:00 pm – 3:30 pm
- **Concurrent Sessions C** 3:45 pm – 5:15 pm

Session A-1: Leveraging Probability Online Survey Panels for Federally Sponsored Statistical Data Collections

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

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

Discussant: Ed Mulrow, *NORC at the University of Chicago*

Location: Vessey 1

A Review of Methods to Combine Probability Surveys and/or Calibrate One Probability Survey with Official Federal Statistical Survey Data

Stas Kolenikov, *NORC at the University of Chicago*

Paul Scanlon, *National Center for Health Statistics*

James Dahllhamer, *National Center for Health Statistics*

Katherine Irimata, *National Center for Health Statistics*

Michael Yang, *NORC at the University of Chicago*

NCHS is endeavoring to better understand the viability of collecting survey data from multi-client probability-based survey panels to produce timely national estimates of population characteristics such as health outcomes and policy-relevant information on emerging topics. The NCHS Rapid Surveys System project is designed to continue the statistical advances in calibration weighting made by the Research and Development Survey (RANDS) program through combining, calibrating, or modeling probability panel data to benchmarks from NCHS flagship surveys, as well as to explore other techniques to develop weighted estimates from multi-client consumer probability panels that are fit-for-purpose for reporting health and health-related population characteristics. To commence this effort, NCHS commissioned NORC and Ipsos to prepare a detailed literature review serving two panel-relevant topic areas. This panel presentation reports on NORC's research regarding methods to combine probability surveys, leveraging federal gold standard data to adjust other sources for estimation. We focus particularly on the situation where at least one data source is a probability-based panel.

Assessing and Improving Calibration Weighting of Web Surveys Using the R-indicator

Rong Wei, *National Center for Health Statistics*

Van L Parsons, *National Center for Health Statistics*

Yulei He, *National Center for Health Statistics*

Recently, commercial panel-based web surveys have been developed to complement the ability of the federal statistical system in providing health information about the U.S. population. For example, the Research and Development Survey (RANDS) is a series of primarily probability-sampled, commercial panel surveys conducted by the National Center for Health Statistics. During the COVID-19 pandemic, a special series of RANDS was used to publicly release experimental estimates on the impact of the pandemic.

Despite their great potential, statistical inferences based on these web surveys might be subject to potential bias compared with NCHS' traditional, high-quality household surveys. For example,

web-based panel surveys may have additional errors due to larger nonresponse and potential coverage bias.

To mitigate these biases, calibrating the weights from the web survey to a benchmark survey may be useful.

We propose to use the R-indicator, originally suggested as a measure of quantifying "representativeness" of surveys with nonresponse, to assess and improve the quality of calibration weighting. In the development of target calibration weights for web surveys, this metric can be effectively used to identify possible calibration variables and compare alternative weighting strategies. Several examples are provided using RANDS surveys.

A Test of a “Federal” Approach to Increase Survey Response and Fit-for-purpose of Probability Panels for Federal Data Collection

David Dutwin, *NORC at the University of Chicago*

Ipek Bilgen, *NORC at the University of Chicago*

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

The AmeriSpeak Probability Panel was originally designed to, in part, service Federal research, given its unique (among panels) design that includes intensive nonresponse follow-up (NRFU) to recruit households to the panel via Federal Express mailings and in-person recruiting. Nevertheless, NORC at the University of Chicago has developed additional protocols, dubbed “AmeriSpeak Federal,” that include sampling with double the rate of households receiving the NRFU during panel recruitment and survey-specific protocols after the recruitment over and above the standard AmeriSpeak approach, including addition of prenotification and nonresponse reminder letters, additional call attempts, and a higher survey incentive. AmeriSpeak tested the impact of these efforts in a recent omnibus survey to assess response rate improvement, reduction in weight variation, and impact of survey estimates. We report these findings in this panel and illustrate the relative value of AmeriSpeak Federal for Federal data collection.

Combining and Standardizing Panel Surveys as Part of a Government Survey System - an Investigation

Van L Parsons, *National Center for Health Statistics*

Yulei He, *National Center for Health Statistics*

Katherine Irimata, *National Center for Health Statistics*

Bill Cai, *National Center for Health Statistics*

Commercial web-based panel surveys can be used to complement existing large scale government surveys to provide health information about U.S. populations. In particular, the National Center for Health Statistics has started utilizing commercially provided, probability-sampled, panel surveys to create auxiliary, and timely, health information systems. This new platform, the NCHS Rapid Surveys System, will consist of periodic surveys fielded by two different survey vendors. While the panel surveys are of adequate quality for many purposes, their use as part of government official statistics may require additional post-data-delivery standardization. Firstly, commonly used health estimators produced from a panel survey should reflect estimates calculated from a high-quality government reference survey. To meet this requirement, some calibration fine-tuning of the panel survey to the reference survey may be required. Secondly, multiple panel surveys may be combined to increase required sample sizes and representativeness. Estimation methods based on the combination of more than one data source must be easily automated for timely release. Our study considers some robust methods used in meta-analysis for data combination. Furthermore, any combined panel survey must still be “representative” of the reference survey. Evaluations of the proposed methods using data from the Research and Development Survey (RANDS) are presented.

Using the Household Pulse Survey to Identify Potential External Benchmarks of Economic, Social, and Health Well-Being

Priyam Patel, *National Center for Health Statistics*

Lauren Rossen, *National Center for Health Statistics*

Katherine Irimata, *National Center for Health Statistics*

Morgan Earp, *National Center for Health Statistics*

With the increased need to produce early estimates, the National Center for Health Statistics (NCHS) is turning towards the use of web-based panels for faster data collection and model based early estimation models. NCHS currently uses probability-sampled panel surveys to create auxiliary and timely, health information systems, including the Rapid Surveys System. One challenge is identifying benchmarks to improve the accuracy of NCHS estimates. This analysis examines the potential for using the Household Pulse Survey (HPS) to identify social and economic well-being indicators that not only trend with health but are also predictive of the prevalence of future health outcomes. The HPS is a web-based survey conducted by the U.S. Census Bureau, and it contains questions that were developed collaboratively with 13 federal agencies to collect data on a broad range of topics, such as childcare, employment, and housing security. We explored

several different methods including regularization and conditional forests to identify indicators that are highly correlated or tend to trend with health outcomes. These findings can help guide the selection of potential predictors for use in small domain and temporal models.

Tuesday, October 24, 2023

9:00 AM

Session A-2: Using Administrative Data for Program Evaluation and Research

Organizer: Justin Nguyen, *U.S. Census Bureau*

Chair: Amy Anderson Riemer, *U.S. Census Bureau*

Location: Vessey 2

Is EHRs Data Extraction for Select Long-Term Care Settings Possible?

Christine Caffrey, *National Center for Health Statistics*

Manisha Sengupta, *National Center for Health Statistics*

Angela Greene, *RTI International*

Michelle Dougherty, *RTI International*

Aiming to automate data sharing and integrate systems in new and innovative ways, the Division of Health Care Statistics in CDC's National Center for Health Statistics (NCHS) is using data from electronic health records (EHRs) to automate survey reporting across various health care settings. The National Post-acute and Long-term Care Study (NPALS) is a biennial study of major sectors of paid, regulated LTC services in the United States and includes primary data collection from adult day services centers (ADSCs) and residential care communities (RCCs). To address reduced response rates and increased respondent burden primarily created by the COVID-19 pandemic, the 2022 NPALS included a study to assess the availability and extent of EHR data for ADSCs and RCCs. The goal of the study was to assess whether select NPALS RCC or ADSC data elements could be extracted from state or commercial EHR platforms or systems. This presentation will provide results of the study which included an environmental scan of EHR literature, investigation of state EHR regulations, subject matter expert interviews with provider or trade associations and EHR vendors, and a high-level analysis of vendor-collected EHR data that could potentially support data extraction or reporting for RCCs and ADSCs in NPALS.

Using Administrative Records to Calculate More Accurate COVID-19 Infection Rates in State Prisons

Elizabeth Ann Carson, *Bureau of Justice Statistics*

Throughout the COVID-19 pandemic, state governments, academics, advocacy organizations, and local and national media published a range of estimates for the rates of prisoners infected with COVID-19. These estimates varied primarily because each entity used a different population base to calculate their rates. Most researchers chose to use either the largest single-day count of prisoners in in 2020 or the mean monthly population as rate denominators to represent the number of persons exposed to COVID-19 in state prisons. Both measures, however, failed to capture the high turnover of state prisoners during the first year of the pandemic. During the COVID-19 pandemic, several states allowed persons who met specific criteria to be released early from prison. Using individual-level administrative records of persons in the custody of 36 state departments of corrections, BJS was able to calculate the exact number of days each prisoner was at risk for exposure to COVID-19 during the first year of the pandemic. A denominator that incorporates both the population size and the time of exposure allowed BJS to show the incidence density of COVID-19 infection in these state prison systems during the study period, an improvement over static cumulative incidence rates for this dynamic population.

Noncitizen Coverage and its Effects on U.S. Statistics

J. David Brown, *U.S. Census Bureau*
Mlsty Heggeness, *University of Kansas*
Marta Murray-Close, *U.S. Census Bureau*

We combine 31 types of administrative record (AR) and third-party sources to produce 2020 population estimates with the same reference date, April 1, 2020, and within the same time frame, as the 2020 Census of Population and Housing. The sources and methodology are designed to improve coverage of historically undercounted population groups. The study documents differences between the AR census, 2020 Census, and vintage-2020 Population Estimates Program estimates. We undertake several linkage exercises to illuminate what can account for the overall differences. Our analysis suggests that differences in coverage of non-U.S. citizens are an important explanatory factor. We explore the plausibility of the noncitizen estimates and reasons why their coverage differs across sources.

Combining Data Sources to Produce Nationally Representative Estimates of Hospital Encounter Characteristics

Jay Breidt, *NORC at the University of Chicago*
Dean Resnick, *NORC at the University of Chicago*
Geoff Jackson, *National Center for Health Statistics*
Donielle White, *National Center for Health Statistics*

The 2020 National Hospital Care Survey (NHCS) is a stratified random sample of US hospitals, conducted by the Centers for Disease Control and Prevention's National Center for Health Statistics (NCHS). Hospitals responding to NHCS provide nearly complete records of patient encounters over the entire 2020 calendar year, making the data extraordinarily valuable for understanding US hospital care utilization and informing health care policy. NHCS is subject to hospital-level nonresponse that reduces available sample sizes and potentially biases results, due to differential response rates across hospital types. Accordingly, NCHS and NORC at the University of Chicago have collaborated to enhance NHCS encounter data with additional data sources, reducing potential biases. The additional data sources include a proprietary commercial hospital encounter data source, treated as a nonprobability sample with unknown hospital participation propensities, as well as nationally representative hospital care benchmarks from the Healthcare Cost and Utilization Project (HCUP). The enhanced data can be used to create nationally representative estimates of hospital encounter characteristics. The enhanced data will also serve as the basis for additional data products, including a weighted public use file and experimental synthetic data products. We will describe our data enhancement approach along with methodological challenges and preliminary results.

Tuesday, October 24, 2023

9:00 AM

Session A-3: Perspectives from the COVID-19 Pandemic

Organizer: Tamara Lee, *Department of Veterans Affairs*

Chair: Tamara Lee, *Department of Veterans Affairs*

Location: Patuxent

Using Small Area Estimation to Create New Tools for Community Resilience

Bethany DeSalvo, *U.S. Census Bureau*

The Census Bureau launched a new tool for national agencies and local communities, the Community Resilience Estimates (CRE). The CRE tracks how vulnerable neighborhoods are to the impacts of COVID-19 and other local disasters, by measuring the capacity of individuals and households to cope with the external stresses of the impacts of a disaster. From the beginning of the pandemic, the negative effects of COVID-19 have been strongly related to certain individual and household characteristics. With access to granular microdata from the Census Bureau, the CRE maps the risk assessment of local populations down to the

neighborhood level and allows national and community leaders to more efficiently respond to emergencies. As the nation's leading provider of quality data about its people and economy, the Census Bureau is uniquely positioned to provide the most accurate and timely measures for an individually focused community resilience indicator. We use detailed demographic and economic data about individuals to build these estimates. Having the richest data sources, the Bureau can produce estimates with the most granularity, highest statistical quality, and broadest coverage, while still protecting privacy.

CDC COVID-19 Mortality Surveillance

Diba Khan, *National Center for Health Statistics*
Lauren Rossen, *National Center for Health Statistics*
Peter Grillo, *Centers for Disease Control and Prevention*
Meeyoung Park, *Centers for Disease Control and Prevention*
Casey Lyons, *Centers for Disease Control and Prevention*
Sarah Witter, *Centers for Disease Control and Prevention*
Matthew Ritchey, *Centers for Disease Control and Prevention*
Farida Ahmad, *National Center for Health Statistics*
Adi Gundlapalli, *Centers for Disease Control and Prevention*
Lauren Rossen, *National Center for Health Statistics*

Accurate real-time data are critical for informed public health decision-making during a public health emergency. In December 2019, a novel coronavirus was reported in China. Subsequently, in January 2020, the first person in the United States was diagnosed with Severe Acute Respiratory Syndrome due to the novel coronavirus, followed by the first confirmed COVID-19 death reported in the United States on February 29, 2020. This presentation will outline the methods used to track COVID-19 death data for 60 jurisdictions by CDC via three mortality surveillance systems: National Vital Statistics System (NVSS), Aggregate Case and Death Surveillance System (ACS) and line-level case surveillance. The NVSS collects death certificate data for all deaths in the US and 5 territories. ACS tracks aggregate COVID-19 case and death data using a variety of data collection methods: web-scraping, Epi Info survey and CSV file transmissions submissions, and Application Programming Interfaces. In the line-level case surveillance, jurisdictions may voluntarily report COVID-19 data to the CDC via electronic messaging to the Nationally Notifiable Disease Surveillance System (NNDSS) or by direct CSV file transmissions. This presentation will also discuss strengths and weaknesses of each surveillance system, along with recommendations for surveillance of deaths during future public health responses.

Validating the Results of a Probability-based Panel Survey with Administrative Data

Divya Vohra, *Mathematica*
Holly Matulewicz, *Mathematica*
Andy Weiss, *Mathematica*
Willow Crawford-Crudell, *Mathematica*
Chandra Couzens, *Mathematica*
Nancy Clusen, *Mathematica*

Although the Centers for Disease Control and Prevention (CDC) received administrative data on both the number of individuals testing positive for SARS-CoV-2 and the number reached through case investigation and contact tracing efforts, little was known nationally about the actions taken by cases and contacts that could mitigate the spread of disease, such as isolation, quarantine, and notifying close contacts. We conducted an online probability-based panel survey to fill these information gaps, providing timely evidence to assess the impact of public health interventions and inform planning for future public health emergencies. We compared the monthly number of SARS-CoV-2 cases estimated by our survey data with CDC's administrative case data between March 2020 and March 2022. We found high correlation between monthly case counts reported by CDC and the survey data estimates over those two years. This correlation greatly reduces our methodological concerns about long-term recall and the relatively low response rate inherent in national panels, giving us confidence in our post-exposure behavioral findings on isolation and quarantining compliance from 2020 to 2022. This presentation will cover key findings of interest to both public health officials as well as survey methodologists interested in the quality and efficacy of panel survey data.

A Multi-Factored Mechanism to Distribute Health Emergency Funds

Stacy Chen, *The MITRE Corporation*
Hongxun Qin, *The MITRE Corporation*
Timothy Champney, *The MITRE Corporation*
Rob Hartman, *The MITRE Corporation*
Surya Menon, *The MITRE Corporation*
Mary Munro, *The MITRE Corporation*
Yueh Quach, *The MITRE Corporation*

When COVID-19 struck in 2020, health providers needed urgent financial support to cope with the strains of the pandemic. Congress passed the Coronavirus Aid, Relief, and Economic Security Act (CARES) in 2020 where COVID-19 Provider Relief Fund was established, allocating funding primarily based on providers' past revenues and case rates. However, studies suggest that the fund was not effectively allocated, disproportionately benefitting hospitals with preexisting higher levels of funding, such as academic medical centers, rather than hospitals at high levels of financial strain, such as those in rural areas and safety-net systems. In this research, we conduct analysis on hospitals and the communities they serve, suggesting a novel multi-factor model to distribute funding. This not only will clarify current Relief Fund allocations, but also could inform rapid grant distribution in the future. Hospital data from Centers for Medicare & Medicaid Services (CMS) and sociodemographic data from Census American Community Survey (ACS) are used, and clustering models are applied to find natural groupings of hospitals based on community characteristics. The groupings, together with revenue, form the basis for the funding mechanism. The new models are then compared against the Provider Relief Fund to determine whether they provide equitable aid to critical access hospitals.

Tuesday, October 24, 2023

9:00 AM

Session A-4: Statistical Products First – A Paradigm Shift for Federal Statistics

Organizer: Stephanie Shipp, *University of Virginia*
Chair: Audrey Kindlon, *National Center for Science and Engineering Statistics*
Discussant: Camille Busette, *Brookings*

Location: Room 2110

Evolving a Data Enterprise to Support Relevant, Timely, and Equitable Statistical Products

Erika Becker-Medina, *U.S. Census Bureau*
Emily Molfino, *U.S. Census Bureau*
Nathan Ramsey, *U.S. Census Bureau*
Sallie Ann Keller, *U.S. Census Bureau*

Stakeholders in federal statistics want more timely, accurate, granular, and customizable information about people, places, and the economy than ever before. The U.S. Census Bureau's traditional survey-focused approach cannot meet this growing demand. But new data sources, data science innovations, and a re-focused commitment to meeting data user needs positions us to meet these challenges. In today's digital era, massive amounts of data are generated as we go about our daily lives. This provides an incredible opportunity to revolutionize how we capture and use data to develop relevant products. Instead of limiting ourselves to the data our surveys produce, we can flip the paradigm to design products based on what data users need. To do this we must integrate our survey data with other data sources. This involves eliciting the purposes and uses our data are to support, collaborating with internal and external data users to develop the products using ALL our data assets, and then embracing varying access modes for statistical product dissemination to support stakeholder needs at all levels of data acumen. This presentation will share how the Census Bureau plans to

re-envision its data enterprise based on a statistical product first approach.

A New Measure of Food Insecurity: A Curated Data Enterprise Demonstration Use Case

Cesar Montalvo, *University of Virginia*

Vicki Lancaster, *University of Virginia*

We propose and present a new measure of food insecurity at the census tract level for the independent cities and counties in the Washington DC Metropolitan Area. This model includes three critical risk factors that affect food insecurity: household size and composition, household income, and small geographies. We built a model based on a financial evaluation of living costs, including food, housing, childcare, healthcare, internet, transportation, taxes, and other expenses, using multiple sources of publicly available information. We use a statistical procedure to estimate joint distributions of the population according to income category and household size at the census tract level. Using information about living costs and income, we categorize households as food insecure, at risk of food insecurity, and food secure. The computations and technologies developed have applicability across other social and policy topics. This demonstration Use Case is important for defining and designing the capabilities for the 21st Century Curated Data Enterprise (CDE). The living cost computation, development of sub-county estimates, and data sources used are examples of reusable components that can be used to develop the CDE use case by use case.

Use Case: Climate Resilience of Skilled Nursing Facilities

Vicki Lancaster, *University of Virginia*

Climate Resilience is the capacity of an individual, community, or institution to respond dynamically and effectively to shifting climate impact circumstances while continuing to function at an acceptable level. Extreme climate events in Florida and Louisiana have exposed how the lack of climate resilience both within nursing facilities and the communities within which they reside can have tragic consequences. This Use Case will answer the question - can we quantify Skilled Nursing Facility resilience to a severe climate event by conducting a realistic appraisal of the facility and the community in which the facility resides? With data from the Centers for Medicare and Medicaid Services we constructed an Emergency Preparedness Deficiency index and data from the Brown University Center for Gerontology and Healthcare Research provided a description of the mental and physical acuity of residents, information that is essential in the event of an evacuation. We used FEMA's National Risk Index to estimate the probability of the nursing staff being able to reach the facility and from Homeland Infrastructure Foundation-Level Data and Census, data to quantify community resilience. We will discuss the issues found integrating data across the various platforms and the challenges these pose for a Curated Data Environment.

Art of the Possible: The Critical Role of Feasibility Use Cases for Census Modernization Efforts

Stephanie Shipp, *University of Virginia*

Edward Wu, *University of Virginia*

The proposed Curated Data Enterprise (CDE) opens new opportunities for measurement associated with changing economic and social conditions. It addresses the nation's increasing appetite for more timely and geographically granular data. Because the CDE is focused on the purpose and uses of the Census Bureau's data, practical applications must play a central role in developing and planning the implementation of the CDE. We discuss several of these applications, called "Use Cases." These Use Cases are of interest to many Census data users, as they were selected based on feedback from various Census stakeholder communities. For each research topic, a Use Case provides a starting point by exploring the current state of the literature, providing a comprehensive data inventory, and identifying gaps in existing data products. It then provides a roadmap for future research by identifying potential ways to integrate and use these data, while exploring capabilities to be built within the CDE and anticipating technical challenges to be addressed. Use Cases form the building blocks for developing the CDE. They provide a starting point for addressing questions of interest, while providing lessons on the skills and capabilities necessary to build the CDE.

Session A-5: Exploring Machine Learning Techniques

Organizer: Gavin Corral, *National Agricultural Statistics Service*

Chair: Valbona Bejleri, *National Agricultural Statistics Service*

Location: Room 2100

Ensemble Modeling Techniques for NAICS Classification in the Economic Census

Daniel Whitehead, *U.S. Census Bureau*

Brian Dumbacher, *U.S. Census Bureau*

The Business Establishment Automated Classification of NAICS (BEACON) is a machine learning tool developed by the U.S. Census Bureau to help Economic Census respondents select their establishment's North American Industry Classification System (NAICS) code. BEACON uses the respondent-provided text, in real time, to predict the respondent's most likely NAICS code. BEACON utilizes past Economic Census responses in conjunction with other data sources such as NAICS manual descriptions and Internal Revenue Service data to create a data dictionary for training and testing purposes. Through an ensemble method, BEACON hierarchically predicts a respondent's NAICS code, first at the 2-digit level and then at the 6-digit level. As a potential means of improving BEACON's current prediction method, we are exploring the use of model stacking to incorporate predictions from alternative models. This research paper details the ensemble modeling behind BEACON and explores this application of model stacking to improve predictions.

A Semi-Automated Nonresponse Detector (SANDS) Model for Open-response Data

Kristen Cibelli Hibben, *National Center for Health Statistics*

Zachary Smith, *National Center for Health Statistics*

Benjamin Rogers, *National Center for Health Statistics*

Valerie Ryan, *National Center for Health Statistics*

Paul Scanlon, *National Center for Health Statistics*

Travis Hoppe, *National Center for Health Statistics*

Open-ended survey questions, particularly embedded construct and error follow-up questions or "probes," allow respondents to provide additional information with few constraints and provide advantages to question designers when little is known about a topic. However, these items, especially in the context of online surveys, are susceptible to insufficient or irrelevant responses that can be burdensome and time-consuming for analysts to identify and remove. To address these challenges, we developed the Semi-Automated Nonresponse Detection for Survey (SANDS) model. Previous evaluation has shown that SANDS categorizes responses into types of nonresponse and valid responses with high sensitivity and specificity. As a pretrained language model, this approach is more accurate than rule-based approaches, and unlike existing machine learning approaches, it can be directly applied to open-ended responses without the need for model retraining or substantial text preprocessing. Building on our prior work, we present results of further evaluation compared with human review on an expanded set of topics. We additionally compare SANDS to rule-based approaches, including word count and response latency. Data are from 15 open-text questions from seven rounds of the NCHS Research and Development Survey (RANDS). We conclude with updated guidance and instructions on model use and access.

A Score Function to Prioritize Editing in Household Survey Data: A Machine Learning Approach

Nicolás Forteza, *Banco de España*

Sandra García-Urbe, *Banco de España*

The Spanish Survey of Household Finances (EFF) is a large-scale survey and a complex statistical operation. Data editing is a major task in the production process of survey data where the revision team manually checks the consistency among questions and considers the help of interviewer comments and audio records

to edit the data if necessary. Household interviews are sometimes fled with data ommisions and inconsistencies. When this occurs, households are recontacted and are re-asked certain parts of the questionnaire. In essence, the manual revision process enteails several costs, namely, time and measurement error. In this paper, using structured and unstructured surgy-generated data, we examine the use of machine learning techniques that allow to classify interviews that require the need to carefully analyze its questionnaire and potentially recontact the interviewed household. We find an algorithm or score function that predicts with relative high accuracy such kind of household interviews. Our contribution to the survey data production literature is twofold. First, we provide a way to shorten revision and data production time. Second, we propose a methodology to reduce the time between first and second contact for recontacted households, potentially also reducing measurement error.

From Policy to Practice: Insights from U.S. Army Administrative Data into Soldier Performance

Skylar Haskiell, *University of Virginia*

Joanna Schroeder, *University of Virginia*

Joel Thurston, *University of Virginia*

We examine how changes in U.S. Army policy impact the description of behaviors related to individual and team performance. In a mixed method design, we compare the output of machine learning against a qualitative review of the documents. The study uses Bidirectional Encoder Representations from Transformers (BERT), a neural network architecture specifically designed to handle large amounts of text, to analyze a corpus of Army documents (e.g., doctrine publications, field manuals) from 1946-2021. BERT captures the contextual relationships between words in a sentence and across documents, which we use to identify emergent and unobservable constructs across disparate data sources. We link these latent concepts to specific prescriptive and descriptive performance behaviors, as well as the characteristics Soldier are expected to manifest through their behavior (e.g., loyalty, bravery, leadership). Incorporating historical records allows us to track changes in the description of the concepts and behaviors across time. Results from BERT are compared against human coded qualitative analysis, including editorial notes from some of the documents' proponents. These analyses capture intentionally introduced thematic changes (e.g., updated definitions, new performance behaviors). By linking the emergent conceptual information together with behavioral descriptions, we assess how changes in policy may impact Soldier performance through the manifestation of specific behavior.

Assessing Survey Questions through a Machine Learning Pipeline: Emotions and Paralinguistic Behaviors

Hanyu Sun, *Westat*

Ting Yan, *Westat*

Anil Battalahalli, *Westat*

The literature suggests that certain acoustic behaviors can be indicative of confusion and frustration, for instance, high pitch variability. In survey research, previous studies have shown that the use of paralinguistic behaviors signals respondent confusion, such as, use of fillers, hesitation and uncertainty in voices, and long pauses (e.g., Conrad and Schober 2021). The occurrence of respondent confusion can be utilized to identify poorly worded questionnaire items. Conventionally, the analysis of paralinguistic behaviors often requires time-consuming and labor-intensive behavior coding of audio- or video-recorded interviews. Consequently, often a small sample of the recorded interviews can be coded due to resource constraints. In this study, we will present the use of a machine-learning pipeline we developed at Westat to detect questionnaire items with poor performance by emotions and paralinguistic behaviors from audio-recorded interviews. We will evaluate the performance of the pipeline using both mock interviews produced in a laboratory setting and field interviews from a nationally representative survey. We will also explore if and how question types affect the performance of the pipeline. Additionally, we will discuss the time and cost implications of using the pipeline as compared to the conventional behavior coding in detecting problematic questions.

Session A-6: Application of Retail Scanner Data to Calculate Monthly Area Price Measures for Foods

Organizer: Mary K. Muth, *RTI International*

Chair: Patrick McLaughlin, *Economic Research Service*

Discussant: Ana Aizcorbe, *Bureau of Economic Analysis*

Location: Room 0105

Nonprobability Weighting for Retail Scanner Data

Saki Kinney, *RTI International*

The USDA Economic Research Service purchases proprietary retail scanner data, known as IRI InfoScan (recently renamed OmniMarket Core Outlets), that is used to enable research on a variety of food-related topics such as nutrition and food security. These data contain detailed barcode-level information on all food and beverage products sold in a large sample of retail food stores in the United States; however, the sample is a nonprobability sample dominated by large chain stores. This presentation will describe the development of annual store-level weights that allow for population estimates by store type and geographic region that account for differences between the sample and population. We will provide background for the application and describe the construction of weights including computation of population totals using data from Nielsen's TDLinX retail store database using data linkage and imputation.

Using Rolling-window Multilateral Price Indexes to Track Food Costs Across Space and Over Time

Chen Zhen, *University of Georgia*

The measurement of food cost changes is important to policymakers and researchers. Although the Consumer Price Index (CPI) is available for food overall and by food category, the CPI does not allow for spatial price comparison and account for very little of the item substitution that consumers employ to mitigate the effect of inflation on household budget. In this presentation, we discuss a method of indexing food costs that 1) accounts for consumer optimization, 2) compares cost of living across space and time, 3) eliminates the chain drift bias commonly found in chained indexes of goods with strong seasonality in demand, 4) allows continuous updating as data for more recent periods become available without needing to revise the published index numbers, and 5) remains spatially transitive. A price index is spatially transitive if the ratio of index numbers between two locations is the same whether prices at the two locations are compared directly or indirectly through a third location. Food price indexes constructed using this method are not only useful to the government for routinely adjusting the Supplemental Nutrition Assistance Program benefits, but also to researchers interested in the association of food prices with food demand, diet, and health outcomes.

Comparing Food-at-Home Monthly Area Prices (F-MAP) with the Consumer Price Index

Megan Sweitzer, *Economic Research Service*

Abigail Okrent, *Economic Research Service*

The U.S. Bureau of Labor Statistics (BLS) Consumer Price Index (CPI) is the principal federal data source for tracking food price changes. Although the CPI is a timely source of price information, it only reflects temporal price variation. Because spatial price differences are important in food policy analysis, price information that varies both temporarily and spatially is useful for analyzing trends in food prices and in econometric models of the effects of food policies. To address this need, we created the Food-at-Home Monthly Area Prices data product (F-MAP) using IRI InfoScan store scanner data. The F-MAP provides food price measures that vary temporally and spatially. The dataset includes monthly mean unit values (\$/100 grams) and six price indices (Laspeyres, Paasche, Törnqvist, Fisher Ideal, GEKS, and CCD) for 90 food categories across the four Census

Regions and 10 major metropolitan areas. We compare the variation between and long-term trends in the F-MAP and the CPI and examine how the underlying source data and index formula contribute to differences between the price measurements for several food categories.

Plenary Session

Organizer: Linda J. Young, *National Agricultural Statistics Service*

Chair: Karin Orvis, *Office of Management and Budget*

Location: *Chesapeake Ballroom*

Opening Remarks

Karin Orvis, *Office of Management and Budget*

New Challenges and New Opportunities for the Federal Statistical System

Katharine G. Abraham, *University of Maryland*

Approaching the five-year anniversary of the signing of the Evidence Act, the Federal statistical system faces many challenges but also an exciting new set of opportunities. Under the Act, statistical agency leaders are being asked to play an expanded role as stewards of the data held by the Federal government. Working together and with the government’s program agencies—and, in some cases, with holders of private sector data—the agencies can help to better meet the growing demand for reliable information about our society and economy and about the effectiveness of policy interventions to address societal concerns. Embracing these opportunities will increase the agencies’ relevance in a rapidly changing world and position the federal statistical system well for the coming decades.

Katharine G. Abraham is a Distinguished University Professor at the University of Maryland. Her published research includes papers on the work and retirement decisions of older Americans; how government policies affect employers’ choices concerning employment and hours over the business cycle; the effects of financial aid on the decision to attend college; discrepancies in alternative measures of employment, wages and hours; and the measurement of economic activity. She served as Commissioner of the Bureau of Labor Statistics from 1993 through 2001 and as a Member of the President’s Council of Economic Advisers from 2011 through 2013. Abraham currently serves on standing academic advisory committees convened by the Congressional Budget Office, Bureau of Economic Analysis, and Federal Reserve Bank of Chicago. She is a Research Associate of the National Bureau of Economic Research and a Research Fellow of the IZA, a member of the National Academy of Sciences and of the American Academy of Arts and Sciences, a Distinguished Fellow of the American Economic Association, and an elected Fellow of the American Statistical Association and of the Society of Labor Economists. Abraham received her Ph.D. in economics from Harvard University in 1982 and her B.S. in economics from Iowa State University in 1976.

Session B-1: Naturally of Course – Strengthening Statistical Programs Through Natural Language Processing

Organizer: Ed Strocko, *Bureau of Transportation Statistics*

Chair: Mehdi Hashemipour, *Bureau of Transportation Statistics*

Location: Vessey 1

Incorporating Survey Weights into Structural Topic Models

Brandon Sepulvado, *NORC at the University of Chicago*

Caroline Lancaster, *NORC at the University of Chicago*

Joshua Y. Lerner, *NORC at the University of Chicago*

Evan Herring-Nathan, *NORC at the University of Chicago*

Stas Kolenikov, *NORC at the University of Chicago*

Surveys often include open-ended questions that elicit text responses. When there are few responses, researchers can manually review them, but manual review becomes infeasible as the number of responses increases, for example, into the thousands or tens of thousands. When researchers want to know the thematic composition of such text responses, topic modeling—an unsupervised natural language processing (NLP) method that identifies topics from a set of texts—offers a potential solution. Unfortunately, most topic modeling software does not allow the incorporation of survey weights, which can bias resulting statistics based upon the open-ended response text. This presentation will describe our team’s efforts to incorporate survey weights into the structural topic model (STM). STMs are a popular approach to topic models that—in addition to identifying topics within a set of texts—allows the estimation of how the prevalence and content of topics differ by covariates, such as gender and race/ethnicity. We will provide an overview of three ways survey weights can be incorporated into STMs and the advantages of each. Finally, we will present a case study comparing analyses of open-ended survey items using traditional (unweighted) STMs and the weighted STMs that we propose.

An NLP-Based Approach to Record Linkage

Lilian Huang, *NORC at the University of Chicago*

Brandon Sepulvado, *NORC at the University of Chicago*

Dean Resnick, *NORC at the University of Chicago*

Jennifer Taub, *NORC at the University of Chicago*

Brenda Betancourt, *NORC at the University of Chicago*

Record linkage is an essential activity when combining datasets; one pressing question is how to leverage text fields in identifying and linking entities. One popular approach to comparing text fields in different datasets is based upon edit distances (e.g. how many characters differ between two strings), but this fails to reflect conceptual distance. For example, “teacher” and “professor” are similar conceptually, but contain very different characters; thus, an edit-based approach such as the Fellegi-Sunter model would be unlikely to capture their similarity. We aimed to determine if word and sentence embeddings – representations of text which capture conceptual similarity – could improve record linkage results. We tested three popular models for generating embeddings: fastText, Universal Sentence Encoder, and Sentence Transformers. We then performed record linkage between the Database on Ideology, Money in Politics, and Elections (DIME) and a Federal Election Commission dataset of campaign contributions. The occupation variable in both datasets was encoded using word embeddings; we then assessed the agreement of the occupation variable and incorporated this into the overall linkage process. We found that word embeddings helped detect additional correct matches, compared to edit distances, but results included incorrect matches. Ongoing research seeks to reduce the number of false positives.

Applying Machine Learning Language Models to Link Similar Text Documents Together to Evaluate FAQ Coverage

Monica Puerto, *U.S. Census Bureau*

Kevin J Zajac, *U.S. Census Bureau*

Elizabeth May Nichols, *U.S. Census Bureau*

Shaun S Genter, *U.S. Census Bureau*

Brian Francis Sadacca, *Accenture Federal*

The last five years, researchers, and practitioners in Natural Language Processing (NLP) have made significant strides to improve the accuracy and abilities of machine learning models. Language models using some of the newest deep learning frameworks are trained on large bodies of unstructured text, like books, Wikipedia, and internet forums. These models have a more nuanced understanding of language across contexts, allowing researchers/users to group documents based on similar content or to facilitate document linkage. This presentation will showcase how NLP machine learning models can be used in a call center context to evaluate reference material coverage, focusing on the 2020 Census Questionnaire Assistance (CQA) operation which assisted respondents over the telephone by answering questions and completing the 2020 Census questionnaire. Over 5 million CQA calls were recorded with the consent of the caller and agents utilized over 300 different FAQs to answer all these types of questions. We will show how using machine learning such as topic modeling and document linkage can help evaluate FAQ coverage.

A Generic and Automated Staff Scraping Tool for School Webpages

Sara Alaoui, Haley Hunter-Zinck, *U.S. Census Bureau*

Ugo Etudo, *U.S. Census Bureau & Virginia Commonwealth University*

Louis Avenilla, *U.S. Census Bureau*

Allison Zotti, *U.S. Census Bureau*

Yathish Kolli, *U.S. Census Bureau*

Patrick Campanello, *U.S. Census Bureau*

Anup Mathur, *U.S. Census Bureau*

The National Teacher and Principal Survey (NTPS) collects information on elementary to high school principals and teachers across the United States. The survey occurs every two years and consists of an initial survey and a follow-up survey. Due to declining response rates, NTPS is investigating alternative data sources to supplement survey responses. To address this need, we used web scraping and a custom pipeline to identify, download, and extract information from school staff roster webpages. The pipeline consists of several steps. First, we submit salted samples of school addresses to the Google Places API to gather school websites and then crawl the returned sites to identify staff directory pages. We then use pretrained and custom developed named entity taggers in combination with the automated identification of HTML tag motifs, to consistently extract teacher information, like names and subjects, from a page. Finally, we represent each page's HTML elements as a graph and conduct graph traversals to link a teacher's name with their corresponding subject and email. We report on the volume of information scraped and validation against manually curated datasets to demonstrate the successful development of a generalizable web scraping program that provides data to augment the NTPS survey responses.

Tuesday, October 24, 2023

2:00 PM

Session B-2: Making Less Work: Impact and Innovations in Monitoring and Adjusting for Nonresponse

Organizer: Doug Williams, *Bureau of Labor Statistics*

Chair: Doug Williams, *Bureau of Labor Statistics*

Location: Vessey 2

Examining Response Rates in the Survey of Occupational Injuries and Illnesses (SOII)

Brittany Cheadle, *Bureau of Labor Statistics*

Erin Huband, *Bureau of Labor Statistics*

Dee Zamora, *Bureau of Labor Statistics*

The Survey of Occupational Injuries and Illnesses (SOII), administered by the Bureau of Labor Statistics (BLS), provides annual information on the counts and rates of work-related injuries and illnesses by incident, industry, geography, occupation, and other characteristics. Each year, BLS selects approximately 230,000 establishments having 11 or more employees in agricultural production and employers of all sizes in all other industries. Starting with survey year 2008, BLS added collection from state and local government establishments to provide estimates of occupational injuries and illnesses among government workers. Participation in the SOII by private sector employers is mandated by OSHA and participation for state and local government employers is determined by state laws. Data for the SOII are collected through a federal-state cooperative program. Although the SOII response rate may be considered high compared to other surveys, there has been a downward trend in response. The response rate for the SOII for survey year 2021 was 85.2% compared to a response rate of 89.7% for survey year 2011. As part of this response rate analysis, we investigate the variations for non-response by establishment size, ownership, industry and geography and assess the impact from the COVID-19 pandemic.

Validating the Census' Low Response Score

Stas Kolenikov, *NORC at the University of Chicago*

Patrick Coyle, *NORC at the University of Chicago*

David Dutwin, *NORC at the University of Chicago*

The U.S. Census Bureau provides an important service to the U.S. survey research community by calculating and distributing the Low Response Score (LRS). LRS is a measure of how much nonresponse Census Bureau faces in the initial self-response phases of data collection of their mandated response operations (Decennial Census and American Community Survey). LRS is provided at the level of Census tracts and Census block groups as a part of the Census Bureau Planning Database (PDB). While informative about the spatial patterns of nonresponse driven by the population demographics, the LRS need to be adapted to other surveys that face lower response rates. In this presentation, we utilize NORC surveys to establish relations between LRS and response rates that we observe in our field effort. We provide simple functional forms of the transformations needed, and discuss how survey statisticians can adopt LRS to fine-tune their general population address-based sampling designs.

Give it Time?: Sample Composition by Completion Date

Megan A. Hendrich, *Ipsos US Public Affairs*

Randall K. Thomas, *Ipsos US Public Affairs*

In survey research, early responders can differ from later responders (e.g., Fricker & Tourangeau, 2010). For example, older, female, urban, and White respondents are more likely to respond earlier in a field period. Some research suggests that younger, male, rural, Black, and Hispanic respondents respond more after more time in field. We conducted a study in Wisconsin to investigate how early responders differ from the overall sample in terms of sample composition and bias (with benchmark comparisons). We fielded two samples in parallel—the first sample included 968 web-based completes from an online probability-based panel, and the

second sample included 1,610 completes from an address-based sample using a mail-back survey. For each sample type, we compared the total sample with two different subsamples: the first 50% of each sample and the first 75% of each sample. We assessed the unweighted distributions for the weighting variables (e.g., age-gender, race-ethnicity, education) to benchmarks for each sample and found mixed results—some variables became more representative while others did not. We then examined the average bias from 15 benchmarks for each weighted sample and found that about half of the variables had reduced bias with increased time in field.

Nonresponse Bias Monitoring and Interventions in the American Community Survey

Jonathan Eggleston, *U.S. Census Bureau*

Stephanie Coffey, *U.S. Census Bureau*

Kendall Houghton, *U.S. Census Bureau*

Carl Lieberman, *U.S. Census Bureau*

This talk will discuss nonresponse bias monitoring and intervention in the American Community Survey (ACS), the largest household survey in the United States. First, we match administrative data to sampled addresses to analyze how respondents and nonrespondents differ with respect to income, program participation, housing, and demographic characteristics. Importantly, we examine how these differences vary by state and sub-state geographies. This fills a gap in the existing literature on nonresponse bias, as most prior nonresponse bias studies examine a national sample with a limited sample size. Second, we discuss how we use these administrative data in an adaptive design framework to help ensure the representativeness of the ACS. Quality models utilizing the administrative data can identify nonresponding households with characteristics already well covered by respondents, allowing for work stoppage on these cases to enable focusing collection efforts on others not yet well represented. This research adds to our knowledge on successfully conducting adaptive design interventions in cross-sectional surveys. Much prior work relies on limited information about nonrespondents, while our use of administrative microdata allows for finer targeting of households.

Improving the Collapsing Criteria and Nonresponse Adjustment of the Consumer Expenditures Survey

Stephen Ash, *Bureau of Labor Statistis*

Reducing nonresponse bias with the nonresponse adjustment of the survey weights is a critical component of producing survey estimates. Like many surveys, the Consumer Expenditure Survey uses the cell-adjustment method when making its nonresponse adjustments. This method involves collapsing cells together when there are not enough sample units in a cell. Improvements to the collapsing criteria are needed because they were developed when response rates were higher than the current level. It is important that we not only improve the outdated collapsing criteria but also consider other improvements due to decreasing response rates and evolving nonresponse adjustment methods, despite past studies of nonresponse bias for the Consumer Expenditures surveys that have not found evidence of nonresponse bias. Our research suggests revising the collapsing criteria and several other aspects of the nonresponse adjustment that will better reduce the mean squared error of expenditures estimates.

**Session B-3: Advanced Statistical Techniques for Survey Inference:
Variance, Editing, and Weighting**

Organizer: Alexia Cooper, *Bureau of Justice Statistics*

Chair: Alexia Cooper, *Bureau of Justice Statistics*

Location: *Patuxent*

Composite Weighting for Hybrid Samples

Mansour Fahimi, *Marketing Systems Group*

Increasingly, survey researchers are compelled to rely on hybrid samples to improve coverage or secure the needed number of respondents in a cost-effective manner by combining two or more independent samples from different sampling frames. In particular, expensive probability-based samples are often supplemented with those from online panels that are substantially less costly. Traditionally, the method of Composite Estimation has been used to blend estimates from different surveys to improve their robustness. This means individual point estimates from different surveys are produced separately and then pooled together, one estimate at a time. However, this piecemeal approach is computationally arduous, and its component point estimates are subject to the inferential limitations of the individual surveys that are used in this process. During this presentation the author will start with a quick review of the composite estimation methodology and then introduces the method of Composite Weighting that is significantly more efficient, both computationally and inferentially when pooling data from multiple surveys. For empirical illustrations, results from three surveys will be presented with each survey relying on hybrid samples comprised of probability-based components from the USPS address database and supplemental samples from online panels.

Developing and Evaluating Alternative Editing Strategies in the Survey of Income and Program Participation

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

Lindsay M. Monte, *U.S. Census Bureau*

Adrienne R. Brown, *U.S. Census Bureau*

The COVID-19 pandemic resulted in substantial changes in the landscape of certain social safety net programs in the United States. Analyses of the 2021 Survey of Income and Program Participation (SIPP) reveal that the data did not fully capture expected changes in the Supplemental Nutrition Assistance Program (SNAP) and unemployment insurance (UI), particularly at the onset of the pandemic. In response to these findings, we developed and assessed additional data editing and imputation processes to produce alternative SNAP and UI variables in SIPP to better reflect expected patterns. We focused attention on monthly benefit receipt and amounts, relying on information already available in the survey and standard data editing strategies. In this paper, we document the updated editing and imputation procedures, describe how they were implemented within the existing data processing framework and timeline, and present results about their effectiveness based on comparisons to available aggregate benchmarks and by linking SIPP respondents to administrative tax and program participation records. Initial examinations of these methods suggest mixed results. We end with a discussion of lessons learned, how SIPP and other surveys may benefit from such procedures in the future, and alternative editing strategies to explore in future work.

A Simple Solution for Adjusting Weights in a Complex Sample

Quatracia Lucky, *U.S. Census Bureau*

Yeng Xiong, *U.S. Census Bureau*

The Annual Integrated Economic Survey (AIES) is an economy-wide survey that will replace seven current annual surveys conducted by the Economic Directorate of the U.S. Census Bureau. The AIES sampling unit is an entire company, which can operate in more than one industry or more than one state. Consequently, the

national and subnational estimates produced by AIES are domain estimates. This paper describes the post-stratification procedures employed by AIES that reduce the sampling variability incurred by domain estimation. These procedures ensure additivity requirements for national and subnational tabulations and ensure that all post-stratified weights are strictly positive. After outlining our considered ratio, raking, and calibration methods, we present our empirical application, highlighting findings that ultimately led to the recommendation of a combined ratio estimation adjustment procedure.

Tuesday, October 24, 2023

2:00 PM

Session B-4: Equity in STEM, Health, Poverty, and Race and Ethnicity Identification

Organizer: Valerie Testa, *Statistics of Income (SOI)*

Chair: Valerie Testa, *Statistics of Income (SOI)*

Location: Room 2110

Measuring Poverty Sub-annually in the United States

Ani Rudra Silwal, *U.S. Census Bureau*

Rapid changes in the U.S. economy and COVID-19 have increased demand for more timely and frequent estimates of well-being. There is a lag between the reference period and publication of annual poverty statistics by the U.S. Census Bureau. Existing high-frequency measures of economic health such as GDP growth and unemployment rates do not give a complete picture of the resources available to households. Therefore, it is important to exploit existing data to produce more timely and frequent estimates of poverty. I build on existing studies combining the monthly Current Population Survey (CPS) with the Annual Social and Economic Supplement (CPS ASEC) to create a sub-annual measure of poverty with reference periods of 1, 3, and 4 months. I present sub-annual estimates of the Official Poverty Measure (OPM) and the Supplemental Poverty Measure (SPM) for 2009-2022. I discuss methodological issues around the design of this measure and present corroborating results from the Survey of Income and Program Participation (SIPP) and the Household Pulse Survey (HPS). The results suggest that a research series on sub-annual poverty estimates could complement existing annual estimates and could be a useful addition to the body of evidence upon which short-run economic policy decisions are based.

In Their Own Words: Race and Ethnicity Self-identification in the Faculty, Academic Careers and Environments (FACE) Survey

Laura Burns Fritch, *RTI International*

Caren Arbeit, *RTI International*

KC Culver, *University of Alabama*

Adrianna Kezar, *University of Southern California*

Emily Koren, *University of Southern California*

John Curtis, *JWC Consulting*

Using OMB's proposed new minimum categories for collecting race and ethnicity, we fielded an experiment on a large survey to better understand the detailed descriptors that respondents attach to the categories. We asked respondents to select the racial and ethnic categories that applied to them, using just the category title, and in the next item to describe their ethnic and racial identity in their own words. These open-ended responses were compared to OMB's proposed detailed categories with the goal of better understanding interpretation of this question and to suggest possible modifications to OMB's proposed categories. For a subset of respondents, the self-reported data were compared data provided by their employers using the IPEDS reporting standard for race and ethnicity. These data were collected as part of the Faculty survey pilot of the Faculty, Academic Careers and Environments (FACE) project targeting faculty, researchers, and academic staff. FACE is modeled after the National Survey of Postsecondary Faculty which was last collected

by NCES in 2004. Thus, FACE aims to understand who faculty are and how their environments shape their ability to thrive as instructors, researchers and/or public scholars in the community. FACE is funded through a National Science Foundation Grant (2200769).

Race and Ethnicity in SNAP Administrative Records and Census Data

James Noon, *U.S. Census Bureau*

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

Renuka Bhaskar, *U.S. Census Bureau*

Mehrgol Tiv, *U.S. Census Bureau*

This work evaluates race and ethnicity information available in SNAP administrative records (ARs) from a selected group of states by leveraging data linkage infrastructure available at the U.S. Census Bureau. We linked 2019 SNAP ARs to Census 2000, 2010 Census, and 2001-2019 American Community Survey data using unique identifiers (PIK) and compared race and ethnicity information across sources. For 2019 SNAP participants who linked to a previous census race response, we found higher consistency in race reporting among Blacks and Asians (80-90%+), while American Indian and Alaskan Natives and Native Hawaiian and other Pacific Islanders had the lowest agreement rates, alongside those with multi-race reporting (<50%). Missing race and ethnicity information in ARs was very low, and those with missing data were more likely to be White or Some Other Race. Multivariate analyses will be used to identify individual characteristics associated with missing and inconsistent reporting of race and ethnicity.

Measuring Sexual Harassment in the Science and Engineering Enterprise

Jennifer Beck, *National Center for Science and Engineering Statistics*

Elizabeth Richards, *Fors Marsh Group*

Anna Scolese, *Fors Marsh Group*

Marcus Maher, *Fors Marsh Group*

Rachel Walker-Kulzick, *Fors Marsh Group*

In 2018, the National Center for Science and Engineering Statistics (NCSES) asked the National Academies of Science, Engineering and Medicine (NAEM) to convene a panel and make recommendations for measuring the Science and Engineering workforce. The panel identified a specific need to collect data on the access and retention of historically underrepresented groups. The panel specifically highlighted the dearth of high-quality data on the barriers of participation for women and minorities in science and engineering and encouraged NCSES to develop survey questions and a survey module to measure harassment and discrimination. In response to this recommendation, NCSES recently completed initial exploratory work needed to begin developing survey questions to measure sexual harassment and discrimination in its portfolio of nationally-representative workforce surveys. This presentation will share the initial findings of that effort, including an implementation plan for moving forward toward survey question development. It also will recommend next steps for the collection of policy-relevant data measuring sexual harassment and discrimination in STEM.

Examining Health Equity Measurement and Representation through the Lens of Total Survey Error

Morgan Earp, *National Center for Health Statistics*

John R. Pleis, *National Center for Health Statistics*

Lauren Rossen, *National Center for Health Statistics*

In the face of emerging health challenges and rapid technological evolutions, NCHS strives to remain at the forefront of health policy guidance and research advancement. One of the goals set by NCHS is to strengthen NCHS' role in informing policies that promote health equity. Using NCHS data, policymakers, researchers, and public health professionals aim to identify health disparities and track progress toward greater health equity in the United States. NCHS' health equity priorities include 1) expanding data collection and analyses, 2) innovating through methodological work, and 3) assessing cross-sector disparities using linked data. Currently NCHS is not only examining overall health equity but is also looking closely at the relationship between equity and the various sources of total survey error to ensure accuracy of health equity findings.

This presentation will include an overview of the different types of methods being used by NCHS to assess health equity measurement and representation through the lens of the total survey error paradigm leading up to dissemination.

Techniques Tested to Improve Representation of Racial and Ethnic Minority, Younger, and Maternity Patients in a National Hospital Experience of Care Survey

William Lehrman, *Centers for Medicare and Medicaid Services*

Marc N. Elliott, *RAND*

Julie A. Brown, *RAND*

Katrin Hamborsoomian, *RAND*

Layla Parast, *University of Texas*

Megan K. Beckett, *RAND*

Surveys often underrepresent racial-and-ethnic minority patients. Methods that improve response rates (RRs) can more fully capture their experiences. HCAHPS, a Centers for Medicare & Medicaid Services survey of patients' experience of hospital care, began in 2006. Scores have been publicly reported since 2008 and used in hospital payment since 2012. Hospitals choose their survey mode. In a 2021 experiment we randomized 36,001 patients within 46 hospitals to one of 6 survey modes: currently implemented Mail Only, Phone Only, and Mail-Phone, and new Web-Mail, Web-Phone, Web-Mail-Phone modes. The Web-Mail-Phone response rate (AAPOR RR1) was highest (36%), followed by Mail-Phone (31%), Web-Phone (30%), and Web-Mail (29%). While race-and-ethnicity is known only for respondents, randomization permits comparison across survey arms. Web-Mail-Phone produced the highest yield for 3 of 5 racial-and-ethnic groups. Otherwise a two-phase mode almost always produced the highest yield. Yield of maternity and patients ages 18-84 was significantly higher in web-first modes. A longer data collection period (49 vs. the current 42 days) resulted in a 3 percentage point gain overall; the largest gains were among under-represented groups. Multi-mode approaches, especially web-first, and longer fielding periods can improve representation of underserved racial-and-ethnic minorities and overall RR in patient experience surveys.

Tuesday, October 24, 2023

2:00 PM

Session B-5: What do We Mean by "For Statistical Purposes Only"

Organizer: Michael Hawes and Sallie Keller, *U.S. Census Bureau*

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

Location: Room 2100

Data subjects are often told their information will be used for "statistical purposes," and statistical agencies are legally required to use these data "for statistical purposes only," but what does this actually mean? In today's data-driven world, statistics is a far reaching and expansive discipline, actively used across virtually all scientific fields, extensively leveraged by financial firms, social media platforms, and public agencies, with myriad daily implications, both large and small, for the average person. With statistics, as a concept, being so broad a discipline, one might expect that the term "statistical purposes" (intuitively, those actions taken in pursuit of the generation, use, or interpretation of statistics), would be similarly expansive. To the contrary, "statistical purposes" has become a term of art in law, ethics, and practice that defines and constrains the legal, ethical, and appropriate uses of certain data collections. This panel will explore how the term "statistical purposes" has been variously defined in federal laws and regulations, and how it has been interpreted in practice by the U.S. federal statistical system. In particular, the panelists will discuss how the legal and ethical

guardrails of “statistical purposes” align with the core objectives and mission of a statistical agency, and how some of the inherent ambiguities of what may constitute a statistical or non-statistical purpose get at the heart of some of the most vexing challenges facing statistical agencies today.

Panelists:

- Sallie Ann Keller, *U.S. Census Bureau*
- Bob Sivinski, *Office of Management and Budget*
- Gail Mulligan, *National Center for Education Statistics*

Tuesday, October 24, 2023

2:00 PM

Session B-6: How the Commodity Flow Survey Made its First 100 Million (Shipment Records)

Organizer: Christian Moscardi, *U.S. Census Bureau*

Chair: Cha-Chi Fan, *Bureau of Transportation Statistics*

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

Location: Room 0105

Modernizing 2022 CFS Data Collection

Berin Linfors, *U.S. Census Bureau*

With the goals of collecting more data while lowering respondent burden and program cost, the 2022 CFS incorporates several collection design innovations. Eliminating paper form submissions, allowing companies with multiple locations to consolidate their reporting, giving the respondent the option to provide all location shipments for the reference week, and using machine learning to code product data have all contributed to achieving these desired goals. Starting with respondent debriefings during the 2017 CFS, modernization initiatives were identified and tested during a 2020-2021 survey pilot. Another round of respondent debriefings was held during the pilot to further guide the 2022 CFS instrument design. The resulting 2022 CFS is on pace to collect 100 million shipment records compared with 6.5 million during the last survey cycle. The 2022 CFS debriefings have shown most respondents like these changes to how they provided the data.

Updating Data Editing and Review for Large-scale CFS Response Data

Gritiya Tanner, *U.S. Census Bureau*

Patrick Nguyen, *U.S. Census Bureau*

The CFS data processing pipeline involves critical steps such as data editing, imputations, and review, which are essential for ensuring the accuracy and completeness of data. This new cycle, we face challenges associated with scaling up the data processing pipeline, including increased computational costs, storage requirements, and data review. In this presentation, we will discuss the need for modernizing data processing pipelines to address scaling challenges, and highlight the changes implemented from the previous cycle to the current cycle. Specifically, we will focus on the strategies and tools used to manage the review of higher volumes of data with limited staffing resources and compare the previous and current approaches.

Monitoring Machine Learning Performance on Large Scale Shipment Data for the 2022 Commodity Flow Survey

Cecile Murray, *U.S. Census Bureau*

Federal statistical agencies are increasingly adopting machine learning models to increase data processing efficiency and improve data quality, but maintaining these models in production is key to ensuring that data quality remains high. The CFS has implemented a machine learning process to improve data quality as well as reduce operational costs and respondent burden. In the 2022 production cycle, we deployed and further

integrated this model to replace a burdensome survey question that required respondents to categorize their shipments, enabling us to collect over 10 times as much shipment data from respondents. We implemented validation and monitoring tools to ensure the ML process continued to produce high-quality results, despite the larger scale of data and the lack of respondent-provided data labels. This talk will describe our approach to validation and monitoring and describe how we addressed these data challenges.

The Freight Analysis Framework: Leveraging Federal Statistical Programs to Understand US Goods Movement

Monique Stinson, *Bureau of Transportation Statistics*

Stephanie Lawrence, *Bureau of Transportation Statistics*

The Freight Analysis Framework (FAF) provides a comprehensive picture of annual freight flows throughout the US. The FAF commodity flow database contains flow estimates (in terms of tonnage and value) by commodity type and mode between origin and destination regions. The FAF network assignment is a network-based representation of flows by mode (truck, rail, etc.) throughout the US. Both domestic and international flows are included. In this presentation, we will discuss the data inputs for FAF and the methods that we use to develop the commodity flow origin-destination estimates and, subsequently, transportation network flow estimates. The emphasis will be on the role of CFS data, which represents about two-thirds of total flows. We will briefly review methods for developing the remaining one-third of flows as well as methods for projecting future flow estimates. We will also discuss implications of CFS decisions, such as mode options in the questionnaire, for FAF development. Upcoming improvements to the FAF, and their relationship to the CFS, will also be discussed. Finally, we will describe how analysts have utilized FAF to evaluate transportation infrastructure, the role of goods movement in the economy, and other important questions.

Tuesday, October 24, 2023

3:45 PM

Session C-1: Building Diverse Leadership for the Statistical System

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

Chair: Kevin Scott, *Bureau of Justice Statistics*

Location: Vessey 1

Leadership Diversity Supports Underserved Communities

Hubert Hamer, *National Agricultural Statistics Service*

NASS's diverse leadership team provided the vision and motivation needed to expand its outreach to historically underserved farmers and ranchers as well as to producers with small, new, and beginning operations, since assuming responsibility for the Census of Agriculture. These producers were hard to reach, and as a result were often underrepresented in census results. While conducting the Census of Agriculture, it is imperative that NASS reach all agricultural populations, regions, and unique producers to paint the full picture of U.S. Agriculture. To that end, maintaining diversity in leadership is essential to continue building relationships with the many community-based organizations (CBOs) that represent the many farmers and ranchers at the grassroots level. The partnerships served both the community-based organizations' mission of providing service to every producer and NASS's goal of counting every producer.

Using Diverse Perspectives to Create Useful Data

Meghan Maury, *U.S. Census Bureau*

Ensuring that the data we create is relevant for the diverse communities we serve requires the input of team members with a diversity of lived experiences. When we recruit, hire, develop, and promote in federal government, we must actively value diverse perspectives alongside technical expertise.

Federal Statistical System - Coordinated Efforts on Guidance about Diversity, Equity, and Inclusion in our Data

Chris Chapman, *National Center for Education Statistics*

The Interagency Council on Statistical Policy has established a working group to develop guidance on how to strengthen federal data needed to understand diversity, equity, and inclusion across the gamut of indicators derived from federal data sources. The presentation will share information on where the work of the working group is and some of the issues that have been most challenging to address.

Diverse Agency Leadership vs. Diverse User Needs

Amy O'Hara, *Georgetown University*

Data consumers demand timely disaggregated data that reflect our society's incredible diversity. Yet agencies are taking stronger privacy stances, reducing the detail or accuracy in data products to reduce disclosure risks. How will "diverse agency leadership" navigate these trends? What if diverse leaders fail to deliver usable data to diverse user communities? What can incentivize agencies to propel each other to better serve all data users? Who will demand accountability to avoid performative "diverse leadership?"

Tuesday, October 24, 2023

3:45 PM

Session C-2: Research on the Initial Proposals for Updating OMB Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity

Organizer: Aleia Clark Fobia, *U.S. Census Bureau*

Chair: Bob Sivinski, *Office of Management and Budget*

Discussant: Rogelio Saenz, *University of Texas-San Antonio*

Location: Vessey 2

Content-Related Proposals from the Interagency Technical Working Group on Race and Ethnicity Standards: Question Format and "Middle Eastern or North African"

Susan Jenkins, *Department of Health and Human Services*

In January 2023, the U.S. Office of Management and Budget (OMB) released an initial set of recommended revisions proposed by the Working Group to revise OMB's Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity (SPD 15), last revised in 1997.

This presentation will focus the following two initial proposals:

- To collect race and ethnicity information using one combined question. The Working Group proposes that SPD 15 move from the two separate questions format to a single combined question as the required design for self-reported race and ethnicity information collections. Employing a new combined question design may take significant time and resources for some surveys and information collections to implement. Flexibilities should be allowed for agencies dependent on aggregate data, data that are not self-reported, or data from non-Federal providers.
- Add "Middle Eastern or North African" (MENA) as a new minimum category. The Working Group proposes that "Middle Eastern or North African" be added to SPD 15 as a new minimum reporting category distinct from all other reporting categories. The definition of the current "White" reporting category would be edited to remove MENA from its definition.

This presentation will discuss the background and evidence that motivated the Working Group on these initial proposals such as, for example, the previous 2014-2018 Federal Interagency Working Group for Research on Race and Ethnicity, existing Federal Government research, and experiences from the 2020 Census.

Implementation-Related Proposals from the Interagency Technical Working Group on Race and Ethnicity Standards

Amy Branum, *National Center for Health Statistics*

In January 2023, the U.S. Office of Management and Budget (OMB) released an initial set of recommended revisions proposed by the Working Group to revise OMB's Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity (SPD 15), last revised in 1997. This presentation will focus on its initial proposal related to the implementation of a revised SPD 15: Guidance is necessary to implement SPD 15 revisions on Federal information collections. In doing so, this presentation will discuss the Working Group's research-to-date on such items as, for example: statistical methods to connect data produced from previous and revised collection formats, procedures for collecting, processing, and reporting detailed racial and ethnic categories, approaches for collecting race and ethnicity information when self-identification is not possible, etc.

Communications and Outreach of Stakeholders and the Public from the Interagency Technical Working Group on Race and Ethnicity Standards

Jennifer Saindon, *Office of Management and Budget*

In 2022, the Chief Statistician of the United States identified updating Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity (SPD 15), last revised in 1997, as a top priority to ensure that SPD 15 better reflects the diversity of the American people. Much purposeful communications and outreach have been conducted with stakeholders, including the public, since Fall 2022 and since the January 2023 release of the initial set of recommended revisions to revise SPD 15 proposed by the Federal Interagency Technical Working Group on Race and Ethnicity Standards. This presentation will describe the Working Group's communications and outreach efforts such as public listening sessions, town halls, a tribal consultation, comments received through the Federal Register Notice, and additional opportunities

Tuesday, October 24, 2023

3:45 PM

Session C-3: Modernizing Data Dissemination at the National Center For Health Statistics: Evidence from Health, United States

Organizer: Renee Gindi, *National Center for Health Statistics*

Chair: Sheila Franco, *National Center for Health Statistics*

Discussant: Bryan Combs, *National Agricultural Statistics Service*

Location: Patuxent

The Health, United States Redesign Research and Findings

Sheila Franco, *National Center for Health Statistics*

Health, United States products are updated periodically to support new methods and preferences for how people access and use health statistics. As part of the latest major redesign, the Health, United States team completed a systematic evaluation of who uses Health, United States, how they use it, and how to modernize products to better meet user needs. As part of this evaluation, the team interviewed government, academic, and public health professionals; obtained feedback from users via a web survey; completed a literature review to understand who cites Health, United States and what sections are commonly cited. They also conducted a market analysis of other high-impact statistical reports and websites to identify new design elements and technology. This information was used to redesign Health, United States from the ground up. Changes included improving how data are stored and accessed, expanding the number of products to meet diverse user needs, improving the timeliness of data releases, modernizing product design to increase interest and usability, and revamping the website to improve accessibility.

Using Communication Science to Redesign *Health, United States* Products, Develop Dissemination Strategies, and Monitor User Engagement

Christine Jones, *National Center for Health Statistics*

As part of the National Center for Health Statistics' goal to modernize data availability and access, communication and program staff at NCHS collaborated on the development of a comprehensive, multifaceted communication strategy for *Health, United States*. This resulted in a complete redesign of many *Health, United States* products using evidence-based communication science techniques to improve data comprehension, engagement, discoverability, usability, and access. These best practices were applied to every *Health, United States* product, including the website, annual report, social media messaging, and email and print marketing materials.

Reporting on the Nation's Health using *Health, United States*

Renee Gindi, *National Center for Health Statistics*

Health, United States presents trends in health status and determinants, healthcare utilization, healthcare resources, and healthcare expenditures and payers with materials geared towards policymakers, public health professionals, and the public. Historically, program staff have focused on producing a single comprehensive report. However, guided by the findings of a systematic evaluation of data users and their needs, *Health, United States* transitioned to a digital first format in 2022. Among its suite of products, the newly designed *Health, United States* website includes topic pages with key findings, charts, and trend analyses. The website also includes trend tables, which are available on the topic pages and through a searchable data finder. A second product, *Health, United States, Annual Perspective* integrates selected analyses from the full compendium of online data. This information is synthesized in a condensed report with a narrower health-focused theme. This publication meets *Health, United States*' congressional mandate while also contextualizing key findings from multiple sources to highlight the complex, multifactorial nature of health and wellness outcomes. The newest edition, released on January 19, 2023, tracks national trends across different health areas with a focus on health disparities by sex, race and ethnicity, insurance status, poverty level, and education.

Tuesday, October 24, 2023

3:45 PM

Session C-4: How the Census Bureau is Using Administrative Records Across the American Community Survey Life Cycle

Organizer: Sandra Clark, *U.S. Census Bureau*

Chair: Sandra Clark, *U.S. Census Bureau*

Location: Room 2110

Using Administrative Records Data in the American Community Survey: Overview

Dorothy Barth, *U.S. Census Bureau*

The current design of the American Community Survey, with an annual sample of roughly 3.5 million housing unit addresses, allows the Census Bureau to collect and update demographic, social, economic, and housing data for the United States every year. The Census Bureau plans to innovate by using administrative records data throughout the survey life cycle to enhance the survey data and the resulting data products. This talk will provide a brief overview of the topics covered on the survey, the data collection strategy, and the general flow of data processing and editing operations to show how we plan to incorporate administrative records data in different ways throughout the survey life cycle.

Implementing an Adaptive Approach to Collect Acreage on the American Community Survey

Ariel Binder, *U.S. Census Bureau*

The U.S. Census Bureau has a long history of using administrative records (AR) to provide quality information about the U.S. population and economy. For example, AR data have been used for decades to produce population estimates and projections. With decreasing survey response rates, the Census Bureau is researching ways to leverage AR sources. Specifically, there are active efforts to examine ways to use AR data to enhance the American Community Survey (ACS) and reduce the burden placed on our respondents, improve data quality, and create blended data products to meet data user needs. One of the ACS items recently evaluated for AR data usage is the lot size for a house or mobile home item. We evaluated third-party acreage data sourced from property tax records, and determined that high coverage and agreement rates suggest it was fit for use in place of ACS survey data. After determining acceptable use of the AR data, we devised a method for using the data in place of survey responses. For CAPI and internet responses, if acreage data are available, the survey item will be skipped and AR data will be used in its place. The lot size item will remain on the paper questionnaire and via Telephone Questionnaire Assistance (TQA), but the survey response will be replaced with AR data when it is available. Using this method, we simulated what we may expect to see in production to compare the AR data-use method with current production method, which only uses survey responses.

Predicting Vacant Housing Units in the American Community Survey

Andrew Keller, *U.S. Census Bureau*

The American Community Survey (ACS) conducts field operations to determine if a nonresponding address is occupied or vacant. Building from the experience of modeling vacancy for nonresponding addresses in the 2020 Census, we develop a model for predicting ACS vacancy using survey paradata, information about the neighborhood, and administrative data. The goal is to use this vacancy information to make decisions regarding ACS operations. For example, field visits can be reduced for units modeled as vacant. After we develop the model, the quality of model predictions is analyzed retrospectively against outcomes from the field. The purpose is to understand both false positive and false negative outcomes in the context of the model predictions.

Adaptive and Responsive Designs with Multiple Competing Criteria: Adaptive Survey Design in the American Community Survey

Stephanie Coffey, *U.S. Census Bureau*

Jonathan Eggleston, *U.S. Census Bureau*

Kendall Houghton, *U.S. Census Bureau*

Carl Lieberman, *U.S. Census Bureau*

Adaptive Survey Designs (ASDs) provide a framework for balancing survey errors and costs, allowing survey managers to utilize limited resources to protect data quality. But ASDs become more complicated when applied in longitudinal or repeated cross-sectional surveys, as the interventions in one period influence interventions made in future data collection periods. The Census Bureau is implementing an ASD in the American Community Survey (ACS) to protect data quality while controlling data collection costs. The ACS is a repeated cross-sectional survey that combines many survey panels to create annual and 5-year estimates. As a result, stopping work on cases in a suboptimal way can have an impact on published estimates and their margins of error for years. Here, we investigated how to balance several competing criteria: the need to control costs while maintaining or increasing representativeness; the need to maintain representativeness without reducing effort repeatedly in the same geographies; and the need to reduce effort without decreasing sample size or inflating variances in geographic domains to unacceptable levels. This presentation discusses each of these criteria, our optimization strategy and objective function, and results of simulations that demonstrate the impact on cases identified for intervention, and the impact on cost and representativeness.

Tuesday, October 24, 2023

3:45 PM

Session C-5: Parsing Policy: Analyzing the Impacts, Sentiments, Utilities and Risks of Cloud-Based Large Language Models and Conversational AI

Organizer: Benjamin Rogers, *National Center for Health Statistics*

Chair: Benjamin Rogers, *National Center for Health Statistics*

Location: Room 2100

Cloud based large language models (LLMs) and Conversational AI, such as Chat GPT, have exploded in popularity in the past year. Their popularity has largely been based on their effectiveness with tasks such as text generation and programming assistance. This panel will discuss public sentiment, utility, risks, and potential economic impact of using cloud based LLMs in both the federal context and in the private sector. Attendees will learn about the potential risks of Conversational AI, how cloud based LLMs can be leveraged effectively and safely by users, what current guidance is being used by federal agencies for employees to follow as well as other considerations as the conversational AI space continues to grow and expand.

Panelists:

- Alex Farach, *Microsoft*
- Travis Hoppe, *National Center for Health Statistics*
- Tala Fakhouri, *Food and Drug Administration*
- Zachary Whitman, *U.S. Census Bureau*

Tuesday, October 24, 2023

3:45 PM

Session C-6: Differential Privacy Population Estimates and Implications for Public Health Policy and Surveillance

Organizer: Angela Werner, *Centers for Disease Control and Prevention*

Chair: Saki Kinney, *RTI International*

Discussant: G. David Williamson, *Centers for Disease Control and Prevention*

Location: Room 0105

Assessing the Impacts of Differential Privacy on Public Health Surveillance at Varying Geographic Resolutions

Angela K Werner, *Centers for Disease Control and Prevention*

Nicholas Skaff, *Centers for Disease Control and Prevention*

Adam S Vaughan, *Centers for Disease Control and Prevention*

Public health professionals rely on accurate and reliable census information for trends over time, calculating rates, and decision-making. With the 2020 Decennial Census, the Census Bureau introduced a computational procedure known as differential privacy. Differential privacy injects statistical noise into population and household data products. This work presents use cases to assess the impacts of differential privacy on public health surveillance. Three use cases were included: COVID-19, National Environmental Public Health Tracking Program asthma emergency department visits and acute myocardial infarction hospitalizations, and heart disease mortality. For each use case, rates were calculated using the Census 2010 Summary File 1 (SF1) denominator, as enumerated in the 2010 Census, and the Census 2010 Differential Privacy (DP) denominator with the final disclosure avoidance settings applied. The percent difference between the SF1 and DP rates was calculated for all rates to determine the magnitude of the difference. County-level data showed significant differences for counties with smaller populations and when stratifying age-adjusted rates. Data continued to

show differences for age-specific rates and census tract-level age-adjusted rates. Rates calculated using 2020 Decennial Census data may be under/overestimated due to differential privacy, and this may affect monitoring trends, finer spatial resolution data, and health equity goals.

Impact of Differential Privacy on Cancer Surveillance Statistics

Mandi Yu, *National Cancer Institute*

Jiming Jiang, *University of California, Davis*

Rates obtained from cancer registries are important measures for identifying determinates of cancer disparities, thus informing health policy. The estimation relies on populations derived from the decennial census as the denominators. When denominators involve random errors introduced by differential privacy (DP), accuracy and reliability of rates can be possibly compromised, which could result in mislead health policy. This presentation simulates county-level DP errors using a TopDown approach that is similar (in principle) to that used to perturb Census 2020 data and add them to the original Census 2010 data to approximate DP-protected Census 2010 data. The sizes of DP errors are based on the latest Census 2010 demonstration dataset. Age-adjusted rates (AAR) of cancer incidence estimated using the original and our simulated DP-protected Census 2010 denominators are compared to assess the impact of DP. A third set of AARs corrected for DP errors are also compared. The results suggest that the magnitude of the bias due to DP is small and not comparable to that of sampling error. However, for small counties, the bias can be nontrivial. Future efforts to develop suppression criteria in data release considering the impact of DP may be needed.

Model-based hybrid small area population estimates: Combining disparate sources of local population data

Lance A. Waller, *Emory University*

Emily Peterson, *Emory University*

Local and regional policy setting requires accurate and timely estimates of population sizes. Examples include assessments of numbers of school-aged children, local vaccination coverage rates, and incidence/prevalence of disease. Official population statistics often only appear after lengthy (but vital) assessments of data completeness and accuracy. Recent years have seen a rise of proprietary, near real-time estimates of local population sizes. We explore approaches to build model-based, hybrid estimates incorporating survey-based estimates, demographic projections, administrative data, and more recent machine-learning-based projections. We outline the different approaches and associated types of uncertainties (reported and unreported). From this catalog of disparate data sources, we outline a framework for combining information across data frameworks and from neighboring areas. We provide initial examples, descriptions and evaluations of assumptions, potential challenges in implementation and interpretation, proposed validity checks, and considerations for future expansion.

Abstract Listings for Wednesday, October 25

- **Concurrent Sessions D** 8:30 am – 10:00 am
- **Concurrent Sessions E** 10:30 am – 12:00 pm
- **Concurrent Sessions F** 1:45 pm – 3:15 pm
- **Concurrent Sessions G** 3:30 pm – 5:00 pm

Session D-1: Building Evidence by Linkage to Data at the U.S. Census Bureau

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

Chair: Barbara Downs, *U.S. Census Bureau*

Discussant: Benjamin Bolitzer, *Department of Commerce*

Location: Chesapeake A

The Demographics of the Recipients of the First Economic Impact Payment

John Voorheis, *U.S. Census Bureau*

Leah R. Clark, *U.S. Census Bureau*

Adam J. Cole, *Department of the Treasury*

Amanda Eng, *U.S. Census Bureau*

Ben S. Meiselman, *Department of the Treasury*

Kevin Pierce, *Internal Revenue Service*

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

Starting in April 2020, the federal government began to distribute Economic Impact Payments (EIPs) in response to the health and economic crisis caused by COVID-19. More than 160 million payments were disbursed. We produce statistics concerning the receipt of EIPs by individuals and households across key demographic subgroups. We find that payments went out particularly quickly to households with children and lower-income households, and the rate of receipt was quite high for individuals over age 60, likely due to a coordinated effort to issue payments automatically to social security recipients. We disaggregate statistics by race/ethnicity to document whether racial disparities arose in EIP disbursement. We provide a set of detailed counts in tables for use by the public.

Methodology on Creating the U.S. Linked Retail Health Clinic (LiRHC) Database

Alice Zawacki, *U.S. Census Bureau*

Joey Marshall, *U.S. Census Bureau*

Donald Cherry, *National Center for Health Statistics*

Xianghua Yin, *National Center for Health Statistics*

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

Retail health clinics (RHCs) are a relatively new type of health care setting and understanding the role they play as a source of ambulatory care in the United States is important. To better understand these settings, a joint project by the Census Bureau and National Center for Health Statistics used data science techniques to link together data on RHCs from Convenient Care Association, County Business Patterns Business Register, and National Plan and Provider Enumeration System to create the Linked RHC (LiRHC, pronounced "lyric") database of locations throughout the United States during the years 2018 to 2020. The matching methodology used to perform this linkage is described, as well as the benchmarking, match statistics, and manual review and quality checks used to assess the resulting matched data. The large majority (81%) of matches received quality scores at or above 75/100, and most matches were linked in the first two (of eight) matching passes, indicating high confidence in the final linked dataset. The LiRHC database contained 2,000 RHCs and found that 97% of these clinics were in metropolitan statistical areas and 950 were in the South region of the United States.

The Labor-Market Returns to Earning Industry Credentials

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

Caroline Walker, *U.S. Census Bureau*

An ongoing and expanding project at the U.S. Census Bureau stems from an agreement between the Bureau, the National Association of Manufacturers, and the National Student Clearinghouse. We examine the labor

market returns to earning industry-certified credentials in a variety of sectors, including manufacturing, health care, safety, and IT. In coordination with credential providers, we estimate the impact of a credential on wages, employment, and employment in the industry sector. To accomplish this, we link a set of students who earned credentials (provided by NAM) to their enrollment and completion records (provided by NSC), which we then link to IRS tax records for earnings and employment (W-2s and 1040s); the Business Register for industry of employment; and the ACS and decennial Census 2000 and 2010 for demographic information such as age, race/ethnicity, and gender. We present earnings trajectories for workers with credentials by type of credential, industry of employment, and demographic characteristics. The ultimate goal of the project is the provision of public-use statistics that can be used to highlight the benefits of credential receipt.

Wednesday, October 25, 2023

8:30 AM

Session D-2: NLSY97 COVID-19 Supplement

Organizer: Leah Christian, *NORC at the University of Chicago*

Chair: Holly Olson, *Bureau of Labor Statistics*

Discussant: Jason Fields, *U.S. Census Bureau*

Location: Chesapeake B

Instrument Programming for Web and Phone Multi-Mode Administration

Rosella Gardecki, *The Ohio State University*

Amanda Roose, *The Ohio State University*

Although NLS had previously used in-person, phone, and self administration, the COVID supplement was the first survey instrument offered for self-administration over the web and the first where all questions might be answered directly by respondents or be read by interviewers over the phone. CHRR exploited flexible, multi-modal survey authoring software to address several challenges raised by this requirement. Issues to consider in this situation include appropriate on-screen presentation of questions, variation of question text for self- vs. FI-administration, and the possibility of mode switches at any point within an interview. Due to BLS security limitations, a verification system was designed which did not collect or reveal PII but which could reliably confirm that the correct respondent had been reached. The successful administration of this mixed-mode survey provides a base on which to build future efforts that include web self-administration as a response option.

A Multi-Mode Survey Design for the NLSY97 COVID-19 Supplement

Quentin Brummet, *NORC at the University of Chicago*

Anthony Washburn, *NORC at the University of Chicago*

Leah Christian, *NORC at the University of Chicago*

Relative to other rounds of the NLSY97, the NLSY97 COVID supplement used an expanded multi-mode survey design with a large web component to quickly obtain information during the changing circumstances of the pandemic. Web was the primary mode of data collection (85.2% completed online) and some surveys were completed by phone, with outreach using a combination of postal letters, emails, texts, and telephone calls. A focus on designing a short 12-minute survey covering the most pressing topics was key to facilitate this design. Overall, this multi-mode strategy yielded a 66.1% completion rate (5,616 completed of 8,490 sampled) with sample representation similar to the overall NLSY97 fielded sample. Outreach by NORC field staff was especially critical to achieving a high response rate and a more representative sample with 78.2% of completes requiring some outreach by NORC field staff (outreach was made to 85.6% of all sample members). Taken as a whole, these results indicate that a carefully designed multi-mode survey approach with primarily web data collection can yield high quality information, especially when quick turnaround data collection is needed. Nonetheless, these experiences also highlight the importance of robust contact information and available interviewing staff to support the mixed-mode design.

NLSY97 COVID-19 Supplement: Initial Findings and Future Use

Donna Rothstein, *Bureau of Labor Statistics*
Alison Aughinbaugh, *Bureau of Labor Statistics*
Keenan Dworak-Fisher, *Bureau of Labor Statistics*

Data from the NLSY97 COVID supplement shed light on the work experiences and health of Americans during the COVID-19 pandemic. Descriptive statistics from the supplement show that a substantial percentage of working men and women were teleworking some hours (46 percent) or all hours (25 percent) in the week before the supplement interview. Having a prior job that was conducive to telework increased the likelihood of telework, and those with a bachelor's degree or higher were much more likely to telework than those with lower levels of education. Reduced county-level activity at the workplace had a positive effect on the likelihood of teleworking, as lessened activity at the workplace during the pandemic was generally associated with increased remote work since many individuals switched from working in the office to working at home. Individuals with poorer health experienced more labor market volatility due to the COVID-19 pandemic than those with excellent or very good health. As longitudinal NLSY97 data are released detailing employment during and after the pandemic, researchers will be able to trace the impact of the pandemic on labor market experiences over the shorter and (eventually) longer term—as the effects may reverberate over many aspects of peoples' lives.

Wednesday, October 25, 2023

8:30 AM

Session D-3: New Perspectives and Methods on Privacy and Disclosure Control

Organizer: Brian Cramer, *National Center for Education Statistics*

Chair: Brian Cramer, *National Center for Education Statistics*

Discussant: Tom Krenzke, *Westat*

Location: Chesapeake C

Estimating Preferences Over Data to Inform Statistical Disclosure Methods Decisions

Elan Segarra, *Bureau of Labor Statistics*

This project provides a framework and empirical example of how to estimate consumer demand for published statistics and incorporate these estimates into the process of statistical disclosure control (SDC). Data providers are tasked with balancing the benefits of published statistics with the risks of re-identification in the confidential micro-data. Rather than just maximizing the number of published statistics, the proposed approach allows providers to assess which statistics might be more useful to consumers. Consumer demand for statistics is modeled using a discrete choice model where individual statistics can vary by characteristics such as their conditioning variables (e.g. labor data sliced by occupation versus industry). Consumer preferences are estimated using standard approaches in the nested logit model literature and these estimates are mapped to an objective function so data providers can assess how potential SDC decisions impact consumers. To illustrate its feasibility, the framework is applied to tabulations of the Census of Fatal Occupational Injuries where privacy and secondary disclosure risk are of paramount concern. Preferences over statistics are estimated using pageview data of the numerous tables and figures published via the Bureau of Labor Statistics public website.

How do Stakeholders Understand Privacy? Findings from In-depth Interviews about Differential Privacy

Isabela Bertolini Coelho, *University of Maryland*

Samantha Chiu, *Independent Consultant*

Frauke Kreuter, *University of Munich and University of Maryland*

Jörg Drechsler, *University of Maryland & Institute for Employment Research*

Rolando Rodriguez, *U.S. Census Bureau*
Trent Buskirk, *Bowling Green State University*

Our research investigates stakeholders' demands and challenges with data privacy. In this study, "stakeholders" represent a variety of research participants, from individuals with little to no knowledge of privacy; to sophisticated or those with high knowledge of privacy. We are interested in public versus federal sector understanding of differential privacy and privacy concerns. Additional demographics on experience, knowledge, and confidence with differential privacy and other privacy techniques enable us to assess potential attitudinal differences between types of stakeholders. The primary research question is "What are the stakeholders' attitudes, perceptions, and participation levels toward privacy?" Secondary research interests focus on differential privacy. Beyond knowledge building, our research is motivated to provide information on how future programs and educational tools can increase the adoption of formal privacy tools. Our initial sample comprises ten in-depth interviews and two focus groups with US Census Bureau privacy experts. Our results provide a variety of perspectives, responsibilities, and trust in a need for formal or additional data privacy measures. Comparing prior knowledge of privacy protection methods and recommendations for future privacy course materials are also analyzed. The preliminary results are also compared to the general public from a Twitter sample.

Differential Privacy for Economic Statistics

Margaret Beckom, *U.S. Census Bureau*
Anthony Caruso, *U.S. Census Bureau*
William Sexton, *Tumult Labs*

The U.S. Census Bureau and Tumult Labs will present work on a new statistical disclosure control methodology being developed for the Census Bureau's County Business Patterns program. This new methodology is based on a novel formal privacy framework, Per-Record Differential Privacy (PRDP), that ensures establishments of all sizes receive a provable privacy guarantee. The presentation will include a discussion of the formal privacy roadmap in the Census Bureau's Economic Directorate, challenges to implementation of differential privacy for establishment statistics, an overview of PRDP, parameter tuning, advantages and limitations of PRDP, and future work.

Modernizing Education Data Systems through Privacy Enhancing Technologies (PETs)

Stephanie Straus, *Georgetown University*
Amy O'Hara, *Georgetown University*

This session explores how Privacy Enhancing Technologies (PETs) can be used in education agency and institutional data systems to drive evidence-based policies. PETs are a group of privacy technologies that allow for increased access to and sharing of extremely sensitive data, through a combination of encryption, noise infusion, and access controls. Through our research, we have found that PETs are currently utilized in K-12 and higher education industries to produce better and more equitable outcomes for students, specifically through improved service delivery, increased public transparency via tiered access, and the linkage of previously un-linkable datasets between distrusting parties. Despite the success of existing uses cases, however, PETs are not widely known by data owners nor fully scalable within education data systems yet. We will discuss the barriers to this scalability, such as legal, regulatory, and capacity considerations, while also framing PETs within today's education data infrastructure landscape. We will additionally highlight current pilot projects with education partners in which we are helping them to implement PETs in their organizations. Given the recent federal pushes for PET inclusion in government data systems, these findings and implications extend beyond education to all local, state, and federal government agencies.

Session D-4: How Can U.S. Professional Guidelines Inform Federal Statistical Policy?

Organizer: Jennifer Park, *National Academies of Sciences, Engineering, and Medicine*

Chair: Rochelle E. Tractenberg, *Georgetown University*

Discussant: Howard Hogan, *ASA Committee on Ethical Practice*

Location: *Vessey 1*

Alignment of ASA Ethical Guidelines with National Guidance

Rochelle E. Tractenberg, *Georgetown University*

Jennifer Park, *National Academies of Sciences, Engineering, and Medicine*

Using structural content analysis methods, we examined the alignment of ASA Ethical Guidelines with other national and federal guidance. We found considerable alignment across guidance, despite differences in the nature, purpose, and intended audience. However, notable differences—and potential for tensions—remain. Areas of strong alignment are well-suited to elevate guidance for federal statistical and data science practice. Areas of potential tension and gaps would benefit from clarifying language, providing examples of best practices, and further work. The ASA Ethical Guidelines are particularly well-suited in this regard given their focus on the practice of statistics and data science at both the individual and organizational levels.

Perspective on OMB Statistical Policy Directives

Kerrie Leslie, *Office of Management and Budget*

The Office of Management and Budget has a central role in establishing federal statistical policy and standards through its Office of Information and Regulatory Affairs and Office of the Chief Statistician. Federal statistical policy directives are one mechanism through which coordination of the federal statistical system occurs. There are thirteen principal federal statistical agencies whose mission is to produce official statistics; another three recognized statistical units share that mission within their respective program offices. Additionally, there are more than a hundred other federal agencies whose core mission is not the production of statistics but nonetheless are engaged in substantial statistical activities. The passage of the Foundations for Evidence-Based Policy Act in 2018 and subsequent guidance issued by OMB have increased opportunities—and expectations—for OMB to engage further with statistical programs, particularly in the areas of data quality, confidentiality, access, and governance. This talk will examine the alignment of OMB Statistical Policy Directives 1 and 2 with the ASA Ethical Guidelines and other national and federal guidance.

Perspective on NASEM Principles and Practices

Connie Citro, *National Academies of Sciences, Engineering, and Medicine*

The National Academies of Sciences, Engineering, and Medicine every four years publishes *Principles and Practices for a Federal Statistical Agency (P&P)*. The 7th edition has 5 principles: relevance, trust of users, respect for data providers, independence from political and undue external influence, and continuous innovation, and 10 practices. This talk will examine the alignment of the P&P principles and practices with the ASA Ethical Guidelines and other national and federal guidance. Tensions and gaps will be discussed, particularly regarding situations where relevance may conflict with ethical principles.

Perspectives on OMB Data Ethics Tenets

Michael Hawes, *U.S. Census Bureau*

The Federal Data Strategy was issued in 2019 to advance access, interoperability, and utility of federal data. Its 2020 Action Plan called for the development of a Data Ethics Framework to help agency employees, managers, and leaders make ethical decisions as they acquire, manage, and use data throughout the data life cycle. Accordingly, the intended audience of this framework extends beyond federal statistical agencies,

designated statistical units, and those with substantial statistical programs. The framework reflects input from the Chief Data Officer Council, the Interagency Council on Statistical Policy, and the Federal Privacy Council. This talk will examine the alignment of the OMB Data Ethics Tenets with the ASA Ethical Guidelines and other national and federal guidance.

Wednesday, October 25, 2023

8:30 AM

Session D-5: Approaches for Improving Survey Response: Perspectives on Incentives, Contact Strategies, and Refusal Conversion

Organizer: Sharon Stern, *U.S. Census Bureau*

Chair: Sharon Stern, *U.S. Census Bureau*

Location: Vessey 2

You Didn't Answer Our Survey, but What About this Text? Converting Hard-to-reach Respondents Through Text Messaging

Maura Spiegelman, *National Center for Education Statistics*

Allison Zotti, *U.S. Census Bureau*

As response rates decrease, new data collection methods may help persuade particularly hard-to-reach sample members to respond to surveys. In this experiment, we attempt to convert reluctant respondents by introducing text messages to other contact methods, and by offering a short, easy-to-answer survey. This experiment was conducted during the 2022 Teacher Follow-up Survey to the 2020-21 National Teacher Principal Survey, sponsored by the National Center for Education Statistics. For sample members who had not completed either a web survey or paper questionnaire after receiving 6 e-mails and 4 mailed packages, we randomly assigned them to either: receive an additional (5th) mailed package; receive a text message with a web survey URL and user ID; or receive a text message inviting them to answer a short two-way SMS text survey by responding to texted yes/no questions, in lieu of the full questionnaire. We compare whether introducing a new mode of contacting and surveying respondents, late in the data collection process, can persuade reluctant respondents to complete the questionnaire. In addition, we examine whether reluctant respondents who engage with the short two-way SMS survey can then be persuaded to complete the full survey.

Calling all Early Birds: Testing a Deadline-Limited Incentive in a Sequential Mixed-Mode Survey

Michelle McNamara, *National Center for Education Statistics*

Rebecca Medway, *American Institutes of Research*

Danielle Battle, *American Institutes of Research*

Ai Rene Ong, *American Institutes of Research*

This presentation will report the results of an “early bird incentive” experiment conducted as part of the 2023 National Household Education Survey (NHES). Sampled households assigned to this condition were offered an additional promised incentive (\$20 cash) for responding before a specific date. This deadline was selected not only to incentivize early response—but also to encourage response before the web-push survey’s eventual switch from web-only mailings to ones that included paper questionnaires. In addition to the typical cost savings associated with web response, the two-phase design of the NHES makes web response especially desirable. Web response to the first phase of the survey (a household screener) allows sampled respondents to move directly into the second phase of the survey (a detailed topical survey about one of the children’s care and education). By contrast, paper response to the first phase requires a break in the response process to allow time for the data collector to conduct within-household sampling and mail the second phase questionnaire. We will compare the early bird condition to a baseline web-push condition (identical except for the exclusion of the promised incentive offer) in terms of the response rate, response timing, mode of response, and nonresponse bias.

Refusal Reasons: Changes Over Time and Differences Between Household Types in the National Health and Nutrition Examination Survey

Steven Fink, *National Center for Health Statistics*

Matt Jans, *National Center for Health Statistics*

Kevin Chuang, *National Center for Health Statistics*

Denise Schaar, *National Center for Health Statistics*

Jill Fleming, *National Center for Health Statistics*

Andrew Caporaso, *Westat*

Jason Clark, *Westat*

George Dixon, *Westat*

Susan Genoversa, *Westat*

Understanding the reasons that people refuse to participate in surveys is an ongoing challenge for survey methodologists. Paradata collected at the doorstep can be very useful in determining people's thought processes in refusing participation. This study examines the refusal reason(s) that respondents provided when asked to participate in the National Health and Nutrition Examination Survey (NHANES) in-person interview. The study has three research questions: 1) which refusal reasons are most and least frequently reported, 2) does their rank order change over time (2017-2021), and 3) do refusal reasons differ between people with and without minors, and by other demographic characteristics observable by interviewers before interacting with the household? Initial results show that the most common refusal reason is "not interested" (40-80% over years) followed by time concerns (i.e., "too busy"; 5-20% over years). Results will be discussed in the context of social psychological nonresponse theories and interviewer best practices.

Can a Higher Incentive Increase Physical Examination Response? Evidence from the National Health and Nutrition Examination Survey

Te-Ching Chen, *National Center for Health Statistics*

Matt Jans, *National Center for Health Statistics*

Lara Akinbami, *National Center for Health Statistics*

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

Damon Ogburn, *National Center for Health Statistics*

David Woodwell, *National Center for Health Statistics*

Allan Uribe, *National Center for Health Statistics*

Jessica Graber, *National Center for Health Statistics*

Incentives increase survey participation. The National Health and Nutrition Examination Survey (NHANES) offers an incentive for participating in a physical examination that takes place at a Mobile Examination Center (MEC) after each respondents' in-home interview. This study investigates the impact of increasing that incentive from \$85 to \$125 for respondents 16 years old and older. Initial results show that the \$125 incentive produced a higher MEC examination rate (78.2% vs. 73.3%, $p=0.009$) and lower cancelled/no show among non-Hispanic White and non-Hispanic Black (78.4% vs. 69.1%, $p=0.013$) respondents. Male respondents had a higher MEC examination rate, and both male and female respondents had lower cancellation/no-show rates when offered the higher incentive (\$125). The higher incentive increased MEC participation among younger adults, and adults 20 years and older had a lower cancelled/no-show rate with the higher incentive. Further, the effect of the higher MEC incentive was seen among respondents from households with 1-6 members, those with health insurance, and those with general (self-rated) health rated as excellent, very good, or good. The higher incentive increased response whether it was offered face-to-face or by-phone. Results will be discussed in the context of survey participation motivators and nonresponse bias.

Experiments to Convert Partial to Completes

Vanessa Harrell, *SSRS*

Margie Engle Bauer, *SSRS*

Todd Hughes, *UCLA Center for Health Policy Research*

Jiangzhou Fu, *UCLA Center for Health Policy Research*

Survey respondents invited by mail to complete a web survey sometimes start a survey but fail to complete it for a variety of reasons, resulting in a breakoff. While some respondents may not be willing to resume due to disapproval of the survey content or a particular question, it is safe to assume that others break off for a variety of other reasons, such as lack of time or loss of interest. Since these respondents read our materials and started the survey initially, we believe that offering incentives may help convert these surveys to completes, at a lower cost than releasing additional sample. In this presentation, we will delve into experiments that we conducted on two health surveys – the California Health Interview Survey (CHIS) and the Massachusetts Health Insurance Survey (MHIS) – for converting partial surveys to completed surveys. Specifically, we will assess the impact of a letter, a pre-incentive, and a promised post-incentive in converting partials to completes.

Wednesday, October 25, 2023

8:30 AM

Session D-6: Application of Respondent-Centered Establishment Survey Design Principles

Organizer: Melissa Cidade, *U.S. Census Bureau*

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

Discussant: Melissa Cidade, *U.S. Census Bureau*

Location: Room 0105

Respondent-Centered Establishment Survey Design Principles: An Overview

Melissa Cidade, *U.S. Census Bureau*

Sarah Grady, *Energy Information Administration*

Kenneth Herrell, *National Agricultural Statistical Service*

Temika Holland, *U.S. Census Bureau*

Maura Spiegelman, *National Center for Education Statistics*

This presentation lays out the four pillars of respondent-centered establishment survey design principles, including best practices for design decisions. These best practices have been developed by an interagency team of survey methodologists specializing in establishment survey testing. At the conclusion, we will lay out the Annual Integrated Economic Survey – the bringing together of seven annual Census Bureau surveys into one instrument – and the respondent-centered research agenda that has informed design decisions.

Combining Content: Considerations for Impactful Change

Heidi St.Onge, *U.S. Census Bureau*

Early in the process of designing the AIES, we harmonized the content within these seven collections into one instrument. This process heavily considered units of collection and respondents' preferences as well as the individual questions and how to create a combined questionnaire. In this presentation, I first present the ways we have navigated the integration of the content across these seven surveys and then demonstrate how we evaluated the best ways to align the differing units of collection based on a respondent-centered approach.

Iterative Instrument Usability Testing and the AIES

Rebecca Keegan, *U.S. Census Bureau*

Several phases of respondent-centered research have been conducted to inform the design of the Annual Integrated Economic Survey. This research has aimed to keep respondent feedback and needs at the forefront of the instrument design decision-making process. The discussed approaches include a record keeping study, a data accessibility study, structural interviewing, and usability testing. In this presentation I will provide an overview of the multiple respondent-centered studies that have informed the design of the AIES instrument.

Respondent-centered Response Options and the AIES

Rebecca Hutchinson, *U.S. Census Bureau*

One of the design goals for the Annual Integrated Economic Survey is minimizing respondent burden while still ensuring high quality data collection. AIES's adaptive design approach to data collection will facilitate more efficient and easier reporting options for companies. In this presentation, I will highlight AIES's preparation for the use of alternate response options available to AIES respondents including more efficient data collection tools and the use of third-party data, direct company feeds, and administrative data.

Wednesday, October 25, 2023

10:30 AM

Session E-1: Advancing Federal Data Collections by Experimenting with Questionnaire Design and Content, Interview Mode, and Contact Strategy

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

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

Discussant: Darcy Miller, *National Agricultural Statistics Service*

Location: Chesapeake A

Count Me In: Challenges of Conducting a Cognitive Interview Study with Adults with an Intellectual and Developmental Disability

Ann MacFadyen, *National Center for Health Statistics*

Amanda Wilmot, *National Center for Health Statistics*

Amanda Titus, *National Center for Health Statistics*

As part of a project aiming to establish a national prevalence estimate of adults with Intellectual and Developmental Disabilities (IDD), the National Center for Health Statistics' Collaborating Center for Questionnaire Design and Evaluation Research (CCQDER) conducted an evaluation of questions that may be used on a national household survey to identify adults with IDD. The evaluation utilized cognitive interviewing methodology. Both respondents with IDD and, where appropriate, a potential survey proxy respondent, took part in one-on-one interviews, forming respondent-proxy dyads. From start to completion there were many challenges in setting up, conducting, and analyzing interviews from this evaluation. These included issues related to recruitment, scheduling and obtaining informed consent, as well as administering the interview and interpreting the findings, from a heterogeneous group of respondents with differing communication styles and cognitive ability. For instance, some respondents needed support when scheduling their interview appointment, or the analyst needed to bear in mind any cognitive delay when reviewing the interview recording. This presentation will describe the ways in which we attempted to overcome such challenges and lessons learned along the way.

A CARI Analysis of Interviewer and Respondent Behavior Within the CE Interview Survey

Ariana Welsh, *Bureau of Labor Statistics*

Victoria Narine, *Bureau of Labor Statistics*

Erica Yu, *Bureau of Labor Statistics*
Brett McBride, *Bureau of Labor Statistics*
Wendy Carlton, *Bureau of Labor Statistics*

To capture data on expenditures, income, and demographic characteristics of consumers in the United States, the Consumer Expenditure (CE) interview survey uses several different question formats designed to minimize both measurement error and respondent burden. For each question format, interviewers are tasked with following standardized survey protocols to read the questions as scripted; deviations from script can affect the meaning of the survey questions and the quality of the data collected. But paradata and interviewer feedback indicate that interviewers at times take shortcuts through these questions. The BLS introduced CARI (Computer Audio Recorded Interviewing) into the CE Interview Survey in 2022 to better understand interviewer-respondent dialogue and if it indicates respondent confusion or unexpected interviewer behavior. As part of the analysis, researchers listened to audio recordings of survey interactions to evaluate in what ways and under what circumstances interviewers deviate from scripted protocols, as well as the effect of those interviewer deviations on data quality. In this presentation, these findings will be discussed, with the goal of looking at how questionnaire designers can better account for interviewer behaviors.

Are Video Interviews for Everyone? A Comparison of CAVI 'Stayers' and 'Leavers' in a Longitudinal Household Survey

Jesus Arrue, *Westat*
Jennifer Kelley, *Westat*
Brad Edwards, *Westat*
Rick Dulaney, *Westat*
Lena Centeno, *Westat*

Starting in spring 2022, video interviews (CAVI) were established as an alternative method of data collection in the Medical Expenditure Panel Survey (MEPS). By fall 2022 and spring 2023, video interviews were fully implemented as either the primary or the secondary mode of the multi-mode data collection, and will be offered on a regular basis for subsequent rounds. The initial analysis of respondents' demographics indicated that video respondents differed significantly from CAPI and CATI regarding age, education, marital status, race/ethnicity, interview language, and household size composition. However, as a novel mode of data collection, it is still unclear if those initially adopting CAVI would use it again for subsequent rounds or if they would switch to a different, preferred mode. This presentation examines data from MEPS-HC's fall 2022 and spring 2023 to compare respondents who stay with CAVI for both data collection periods (Stayers) to those who switch (Leavers) to another mode in spring 2023. The findings will provide better insight into the type of respondents willing to adopt CAVI and those willing to stay with CAVI (for longitudinal surveys), and provide practitioners with strategies for sample management and allocating resources for a multimode data collection that includes CAVI.

Lessons Learned: A Direct Care Worker Pilot Study

Manisha Sengupta, *National Center for Health Statistics*
Christine Caffrey, *National Center for Health Statistics*
Jessica Lendon, *National Center for Health Statistics*
Marie Squillace, *ASPE*
Angela Greene, *RTI International*
Melissa Hobbs, *RTI International*

Direct care workers (DCWs), such as personal care aides, nursing assistants and aides play an essential role in the health and well-being of over 20 million Americans who receive long-term services and supports (LTSS). DCWs assist older adults and people with disabilities in activities of daily living and instrumental activities of daily living. The last national surveys of DCWs were conducted in 2004 (National Nursing Assistant Survey) and in 2007 (National Home Health Aide Survey). According to the 2020 National Post-acute and Long-term Care Study (NPALS), DCWs provided more hours of care in residential care communities, adult day service centers, and nursing homes than did any other staff type. This presentation is a methodological discussion of a pilot study of DCWs in these three settings. NCHS and RTI used the NPALS infrastructure to design and test

a protocol to sample DCWs and obtain their contact information from NPALS respondents. NCHS also designed and tested questionnaire items and contact strategies to survey both employed and contracted DCWs. Findings will inform the feasibility of conducting a future national DCW survey in 2024. Strengthening data infrastructure is critically important to allow for better workforce planning efforts and to inform policies to support DCWs.

Prenotification Experiment in a Survey of Law Enforcement Agencies

Ryan Weber, *RTI International*

Harley Rohloff, *RTI International*

Tim Smith, *RTI International*

Elizabeth Davis, *Bureau of Justice Statistics*

Alexia Cooper, *Bureau of Justice Statistics*

Sean Goodison, *Bureau of Justice Statistics*

The Bureau of Justice Statistics has administered the Census of State and Local Law Enforcement Agencies (CSLLEA) regularly since 1986. Approximately 20,000 state and local law enforcement agencies throughout the country are asked to participate in each data collection cycle (roughly every 4 years). For the most recent CSLLEA (2022), RTI conducted a prenotification experiment with a random sample of 1,000 local police agencies and sheriff's offices, stratified by agency size (FTE officer counts). Each agency in the experimental group received a letter prior to the start of data collection announcing the survey and asking them to login to confirm their contact information. Survey invitations and subsequent outreach attempts were the same for control and experimental cases based on survey response status. Although still in active data collection, early indications are that the prenotification letter had no effect on response speed or response rate, either overall or by agency size and type. Any effect observed is cancelled out when weighting to account for the large number of local police departments on the frame. We will discuss implications for future establishment surveys, particularly those involving law enforcement agencies.

Wednesday, October 25, 2023

10:30 AM

Session E-2: Leveraging Data: Using Administrative and Synthetic Information

Organizer: Gavin Corral, *National Agricultural Statistics Service*

Chair: Lu Chen, *National Agricultural Statistics Service*

Location: Chesapeake B

Can It Work for Employers? Evaluating the Expansion of Administrative-Records Use beyond Nonemployer Demographics Statistics

Adela Luque, *U.S. Census Bureau*

Valeska Araujo, *U.S. Census Bureau & George Mason University*

John Earle, *George Mason University & U.S. Census Bureau*

Vitaliy Novik, *U.S. Census Bureau & George Washington University*

Jared Wold, *U.S. Census Bureau & George Mason University*

Sammuel Young, *Arizona State University & U.S. Census Bureau*

Our ability to identify and research business demographic trends and performance disparities across demographic groups hinges upon the availability of reliable, frequent, and timely business demographics data. In response to declining response rates, and increasing imputation rates and costs, starting with reference year 2017, the Census Bureau began providing nonemployer demographics not through a survey, but a program that leverages administrative and census records to assign demographics to the universe of nonemployer firms: the annual Nonemployer Statistics by Demographics series (NES-D). Given NES-D's success, Census is now evaluating the feasibility of assigning demographic characteristics to U.S. employer

businesses using administrative records. In the presentation we will describe the background, methodology, ongoing challenges, current results and next steps of this ongoing effort. The use of administrative records in business demographics statistics should be viewed as a complement to surveys, and a vehicle to unburden respondents and allow the survey to measure issues that cannot appropriately be captured with administrative records or third-party data –hence strengthening and expanding the capacity of the Federal Statistical Ecosystem.

Who is a Veteran? Evaluating DoD Data for Use in the Census Business Demographics Program

Adji Fatou Diagne, *U.S. Census Bureau*

Adela Luque, *U.S. Census Bureau*

The Census' Business Demographics program provides business demographics estimates through a blended-data approach that combines administrative records (AR)-based estimates for nonemployer firms (through NES-D, the Nonemployer Statistics by Demographics series) and survey-based estimates for employer firms (through the Annual Business Survey, ABS). This program produces statistics by sex, race, ethnicity, and veteran status. However, for veteran status the AR-based concept is narrower because it leaves out some military service categories covered in ABS questions. In this project, we propose to use data from the Department of Defense (DoD) to supplement the Department of Veterans Affairs (VA) data to bring the AR-based concept closer to the survey-based concept. We perform several exercises linking the AR datasets and find the DoD data to be a complement of the VA data by increasing coverage. We also assess agreement between these data and survey responses from the ABS and find that administrative records mostly align with survey responses.

Imputing for Extraordinary Sample Events: A Story of Targeted Donor Pools and Administrative Data

Bradley Rhein, *Bureau of Labor Statistics*

The Current Employment Statistics (CES) State and Area program produces monthly employment estimates by subnational geography and industry through a survey of about 666,000 establishments. Occasionally a large-scale event, such as a hurricane, occurs near the survey reference period, significantly affecting data collection for that month, and challenging an implicit missing-at-random assumption in the estimator. Several methods have been activated to catch the ensuing employment drop, though none of these methods account for the non-response directly. This paper proposes the potential use of a random hot deck imputation approach, constructing a targeted donor pool from a similar circumstance occurring in past administrative data from the Quarterly Census of Employment and Wages (QCEW) program. The authors explore using a donor pool from an earlier hurricane to impute for missing reporters affected by hurricanes in 2017 and 2018.

Estimating the U.S. Citizen Voting-Age Population (CVAP) Using Blended Survey Data, Administrative Record data, and Modeling

J. David Brown, *U.S. Census Bureau*

Genevieve Denoeux, *George Washington University*

Misty L. Heggeness, *University of Kansas*

Carl Lieberman, *U.S. Census Bureau*

Lauren Medina, *U.S. Census Bureau*

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

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

Joseph L. Schafer, *U.S. Census Bureau*

Matthew Spence, *U.S. Census Bureau*

Lawrence Warren, *U.S. Census Bureau*

Moises Yi, *U.S. Census Bureau*

This study develops a method using administrative records (AR) to fill in responses for nonresponding

housing units rather than adjusting survey weights to account for nonresponse. The method also inserts AR and modeling in place of edits and imputations for survey citizenship item nonresponses. We produce Citizen Voting-Age Population (CVAP) tabulations using this enhanced CVAP method and compare them to estimates not using AR at all and ones that use AR to inform nonresponse bias weight adjustments. The competing methods and sources produce noticeably different estimates, especially among Hispanics. We explore reasons behind the differences in detail. We find that weight adjustments informed by administrative data may be less effective at addressing nonresponse bias than direct use of administrative data.

Wednesday, October 25, 2023

10:30 AM

Session E-3: Inclusive Insights: Assessing the Implementation and Effects of Adding SOGI Questions to Large-Scale Data Collections

Organizer: Julie Weeks, *National Center for Health Statistics*

Chair: Julie Weeks, *National Center for Health Statistics*

Discussant: Kristen S. Miller, *National Center for Health Statistics*

Location: *Chesapeake C*

Survey Measurement of Sex at Birth and Transgender Identity in the Behavioral Risk Factor Surveillance System

Kathryn O'Neill, *University of Pennsylvania*

Jody Herman, *The Williams Institute, UCLA*

Much research on the U.S. transgender population relies upon the Behavioral Risk Factor Surveillance System (BRFSS). The BRFSS is the world's largest continuously conducted health survey, and is one of few large-scale surveys to attempt to identify transgender respondents. The term transgender is most often defined as a difference between one's sex at birth and gender identity. While other surveys identify people using a two-step approach, which asks sex assigned at birth and gender identity separately, the BRFSS asks about respondents' identity with the question, "Do you consider yourself to be transgender?" This measure captures people who identify as transgender, but may not capture people who fit the definition but do not use that term. The recent addition of a sex assigned at birth module by some states allows us to assess how individuals who do or do not identify as transgender answer these questions and where the combination of sex at birth and other measures may produce unexpected results. Variation in states' implementation of the sex at birth module adds insight into response patterns. Findings suggest the absence of the sex at birth module has resulted in incorrect health questions being asked to many respondents, particularly for transgender respondents.

Asking about Gender as a Single, Non-binary Construct Versus a Two-Step Question Approach

Valerie Ryan, *National Center for Health Statistics*

This presentation describes an experiment using the National Center for Health Statistics' Research and Development Survey (RANDS) to understand how various question design elements impact the collection of gender. The RANDS is an ongoing set of surveys used for quantitative and mixed methodologies to assess measurement error, validity, and bias. Specifically, for this presentation, findings from an experiment in which a single item is compared to a two-question approach (whereby sex-assigned-at-birth and current gender identity are both asked) will be presented. Previous RANDS studies have shown that, when using a two-step approach for collecting gender identity, 22.9% of identified gender minority respondents were false positive cases. Importantly, this error significantly skewed estimates on key health characteristics, overestimating the health and wellbeing of gender minorities. Findings from this experiment will indicate how the two approaches (one versus two questions) differ in terms of error and estimates of key health outcomes.

Order Effects and Question Wording of Gender and Sexuality Blocks on Surveys: Experimental Evidence in the US and UK

Christina Pao, *Princeton University*

Considering increased disclosures of LGBTQ+ and non-binary/trans identities, many statistical bodies have added and/or changed questions relating to sex, gender, and sexuality on their surveys and censuses. Nonetheless, question wording, ordering, and framing for gender and sexuality measures have differed—leaving concerns as to whether these differences significantly affect response and disclosure. Using a randomized survey experiment fielded in the US and UK (N=2,518) in 2022, I test whether i) question wording of two-step gender questions and ii) block ordering of gender and sexuality change the disclosure (or lack thereof) of non-cisgender and non-straight identities. I identify causal effects with difference-in-means tests using Bonferroni corrections. There is indication of greater nonresponse on the two-step gender measure modeled off the UK census question than the measure modeled after the US General Social Survey (GSS). Further, asking a sexuality block before a gender identity block increases disclosure of non-cisgender identities. No other significant differences were found in terms of question ordering and nonresponse. There are some significant differences in disclosure and nonresponse based on different respondent pathways for a sociodemographic survey section. These effects, particularly at a population level, could heavily skew estimates of the LGBTIQ+ population and have downstream effects on policy recommendations.

Adding Sexual Orientation and Gender Identity Measures to the Medicare Current Beneficiary Survey (MCBS)

Andrea Mayfield, *NORC at the University of Chicago*

Marisa Wishart, *NORC at the University of Chicago*

Melissa Heim Viox, *NORC at the University of Chicago*

Hannah Murrow, *NORC at the University of Chicago*

Marina Vornovitsky, *Centers for Medicare and Medicaid Services*

In January 2021, the White House released Executive Order 13988 on Preventing and Combating Discrimination on the Basis of Gender Identity or Sexual Orientation, which calls upon agencies to identify existing and new policies to promote equal treatment under the law and ensure that all persons can access healthcare and other essential services without being subjected to sex discrimination. After cognitive testing and a careful review of sexual orientation and gender identity measures currently used in other surveys, the Medicare Current Beneficiary Survey (MCBS) will introduce socio-demographic items sourced from the National Health Interview Survey and the U.S. Census Bureau's Household Pulse Survey about sexual orientation and gender identity into the Fall 2023 questionnaire to improve measures of health equity. The MCBS is a continuous survey of a nationally representative sample of the Medicare population, conducted by the Centers for Medicare & Medicaid Services (CMS) through a contract with NORC at the University of Chicago. This presentation describes methods used to incorporate sexual orientation and gender identity into the questionnaire, including item selection, placement within the questionnaire, impact on item universe for sex-specific health status questions, and consideration for proxy interviews.

Session E-4: Synthetic Data Generation for Survey Data

Organizer: Terrance D. Savitsky, *Bureau of Labor Statistics*

Chair: Terrance D. Savitsky, *Bureau of Labor Statistics*

Discussant: Monika Hu, *Vassar College*

Location: *Vessey 1*

Synthesis for Complex Surveys

Yajuan Si, *University of Michigan*

Jerry Reiter, *Duke University*

Synthetic data comprise simulated records with values generated from statistical models estimated with the confidential data. However, there is a knowledge gap about how statistical agencies should generate synthetic data to respect complex survey designs. The practice here is purely ad hoc, with minimal theoretical or empirical evidence to support best practices. We aim to develop methods and recommend best practices for incorporating complex survey designs in synthetic data. We leverage the framework of Bayesian finite population inference for synthetic data generation, deriving new inferential procedures that account for the synthetic data generation process. We will evaluate these new techniques and existing techniques theoretically and empirically to develop a set of recommended best practices.

Synthetic Population Generation for Nested Data using Differentially Private Posteriors

Matthew R. Williams, *RTI International*

Hang Kim, *University of Cincinnati*

Terrance D. Savitsky, *Bureau of Labor Statistics*

Monika Hu, *Vassar College*

When working with restricted use data with a nested structure (e.g. students within schools, or patients within hospitals), both the individual and the group participation may be considered sensitive. Direct application of differential privacy methods by setting a privacy budget for individuals quickly compounds when considering the joint case of the aggregated group. The privacy weighted pseudo-posterior is a form of the popular exponential mechanism and can be used to generate differentially private parameter estimates as well as synthetic populations for non-nested data. We explore the challenges and trade-offs of extending the privacy weighted pseudo-posterior approach to a two-level hierarchical model such as an analysis of variance (ANOVA) model. Challenges include (i) establishing a group-level neighborhood definition for databases analogous to the usual Hamming - 1 definition for neighboring databases that differ by one record and (ii) interogating the privacy risk of latent variables (e.g. random group means). Trade-offs include the usual competing privacy and utility goals as well as the competing priorities between individual and group level objectives.

Private Tabular Survey Data Products through Synthetic Microdata Generation

Terrance D. Savitsky, *Bureau of Labor Statistics*

Monika Hu, *Vassar College*

Matthew R. Williams, *RTI International*

We propose two synthetic microdata approaches to generate private tabular survey data products for public release. We adapt a pseudo posterior mechanism that downweights by-record likelihood contributions with weights $\in [0,1]$ based on their identification disclosure risks to producing tabular products for survey data. Our method applied to an observed survey database achieves an asymptotic global probabilistic differential privacy guarantee. Our two approaches synthesize the observed sample distribution of the outcome and survey weights, jointly, such that both quantities together possess a privacy guarantee. The privacy-protected outcome and survey weights are used to construct tabular cell estimates (where the cell

inclusion indicators are treated as known and public) and associated standard errors to correct for survey sampling bias. Through a real data application to the Survey of Doctorate Recipients public use file and simulation studies motivated by the application, we demonstrate that our two microdata synthesis approaches to construct tabular products provide superior utility preservation as compared to the additive-noise approach of the Laplace Mechanism. Moreover, our approaches allow the release of microdata to the public, enabling additional analyses at no extra privacy cost.

Calibration procedure for estimates obtained from posterior approximation algorithms, with application to domain-level modeling

Julie Gershunskaya, *Bureau of Labor Statistics*

Terrance D. Savitsky, *Bureau of Labor Statistics*

The mean field Variational Bayes (VB) algorithm implemented in Stan is relatively fast and efficient, making it feasible to produce model-estimated official statistics on a rapid timeline. Yet, while consistent point estimates of parameters are achieved, the mean field approximation often produces inaccurate uncertainty quantification to the extent that parameters are correlated a posteriori. In this paper, we propose a simulation procedure that corrects the coverage of the credibility intervals for model parameters. Our procedure detects and corrects biased estimation of both first and second moments of a marginal parameter posterior distribution for any solver algorithm that consistently propagates such bias over multiple datasets. The method generates replicate datasets using parameters estimated in an initial model run. The model is subsequently re-estimated on each replicate dataset, and we use the re-estimated posterior distributions to measure and correct bias using the parameter values in the initial model run. We demonstrate the performance of our procedure in Monte Carlo simulation study and apply it to real data from the Current Employment Statistics survey.

Wednesday, October 25, 2023

10:30 AM

Session E-5: Ensuring Scientific Integrity in the Federal Statistical System

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

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

Discussant: Michael Hawes, *U.S. Census Bureau*

Location: Vessey 2

Official statistics are an important public good, but to have value these data must be, and must be perceived to be, credible, accurate, and objective. Any erosion in public trust in the integrity of federal statistics would impair the nation's ability to rely on those statistics for critical policymaking, public and private sector decision-making, research, and other important societal uses. This session will discuss the importance of scientific integrity for the federal statistical system and various factors that could threaten or undermine it. Panelists will also discuss concrete steps that statistical agencies can take to protect and enhance the integrity of their scientific and statistical methods, operations, and products.

Panelists:

- Chris Marcum, *White House*
- Barry Johnson, *Internal Revenue Service*
- Melissa Chiu, *CNSTAT*
- Monique Eleby, *U.S. Census Bureau*
- Michael Walsh, *U.S. Census Bureau*

Session E-6: The Missing Data Puzzle: Exploring Imputation Methods

Organizer: Diba Khan, *National Center for Health Statistics*

Chair: Diba Khan, *National Center for Health Statistics*

Discussant: Phil Kott, *RTI International*

Location: Room 0105

That's a Long Survey! Using Split-Questionnaire Design to Reduce Respondent Burden in a State Health Survey

Cameron McPhee, *SSRS*

YuChing Yang, *UCLA*

Jiangzhou Fu, *UCLA*

Cordelia Horch, *SSRS*

Declining survey participation rates often lead researchers to ask as much as they can from cooperating respondents, leading to longer questionnaires and increased burden. One method for combatting this is the use of split-questionnaire designs (SQD) where a questionnaire is divided into a set of modules each comprised of a subset of the survey questions. Research has shown SQDs lead to minimal loss of accuracy in estimates when the designed thoughtfully. Most research exploring SQDs employs multiple imputation (MI) to produce full-sample estimates. MI is favored because it is least likely to attenuate estimated variances. However, MI poses a challenge for surveys that release micro-data since MI yields multiple plausible values for every item, which can become unwieldy and difficult to analyze. This paper examines the feasibility of using hotdeck imputation to recover data missing due to the use of SQDs. The analysis will use data from the 2021 and 2022 California Health Interview Survey. Different SQD allocations and hotdeck methods will be simulated. Resulting univariate and multivariate estimates will be compared to those achieved with the full sample. Analyses will also examine the effect of the imputation procedure on precision and whether different imputation models can improve estimates of uncertainty.

Effect of Improving Data Imputation and Processing Procedures on ERS Farm Household Income Forecasts

David Williams, *Economic Research Service*

Wei-Yin Loh, *University of Wisconsin-Madison*

Daniel Ayasse, *Economic Research Service*

Katherine Lim, *Economic Research Service*

Christine Whitt, *Economic Research Service*

We develop new imputations for missing observations in the Agricultural Resource Management Survey (ARMS) using the machine learning algorithm GUIDE. Then we evaluate the statistical performance of these imputations compared to existing ERS procedures and consider the effects of alternative imputation methods on one of our ERS products the farm household income forecast. GUIDE is a multi-purpose machine learning algorithm for constructing classification and regression trees and forests. We utilize GUIDE to impute missing observations and compare those new imputations to the existing imputation procedures by testing either's ability to maintain the data's original statistical relationships (means, variances, covariances). In addition to the statistical improvements, we demonstrate the economically sensible and intuitive imputations obtained by GUIDE as well as the ability of GUIDE to impute missing values in a variable without having to impute missing values in the predictor variables. The latter property implies that GUIDE does not require assumptions on missingness mechanisms in the predictor variables. We then demonstrate the practical implications of our data processing and imputation improvements on a popular ERS product, the farm household income forecast.

Editing Tools for Reconciling Household Energy Characteristics with Administrative Data from Energy Suppliers

L. Kaili Diamond, *Energy Information Administration*

S. Grace Deng, *Energy Information Administration*

The Residential Energy Consumption Survey (RECS) is a periodic study conducted by the U.S. Energy Information Administration (EIA) that provides detailed estimates of energy usage in U.S. homes. RECS collects data through two separate surveys: the Household Survey that collects housing characteristics from occupied primary residences, and the Energy Supplier Survey (ESS) that collects energy billing data from energy suppliers. This paper discusses editing methods and tools for identifying and resolving inconsistencies between the Household Survey responses and ESS data. These methodologies include implementing a phased editing process, the utilization of modeled energy consumption to identify potential inconsistencies, and the creation of a dashboard to compare the modeled consumption to the actual consumption reported by suppliers. We then explore the impact of the changes from reconciliation edits on key estimates such as space heating fuel. Finally, we will discuss how lessons learned from this editing process will inform decisions for future RECS cycles.

Improved Longitudinal Imputation Method in Survey of Doctorate Recipients

Minsun Riddles, *Westat*

Jean Opsomer, *Westat*

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

Laura Alvarez-Rojas, *Westat*

Shelley Brock Roth, *Westat*

Medha Uppala, *Westat*

Longitudinal surveys offer insights not possible with cross-sectional surveys, such as providing visibility into potential impacts of changes in policy by following outcomes of interest over time. However, longitudinal data are more complex by design. Particularly, in imputation for longitudinal surveys, respecting longitudinal relationships is crucial to preserve temporal patterns in addition to accurately predicting missing data at a given time. The National Center for Science and Engineering Statistics' Survey of Doctorate Recipients, which produces both cross-sectional and longitudinal data, currently imputes missing data through a hierarchical hot-deck method in two phases: imputing missing items for respondents to the current cycle without incorporating longitudinal patterns consistently, then, imputing missing years of data for longitudinal cohort cases using longitudinal patterns while taking the cross-sectional imputations as fixed. We investigate the effects of the systematic usage of information from past cycles. The goal is to improve donor matching in the cross-sectional imputation and, in turn, enhance the data quality of not only the cross-sectional data by strengthening the imputation accuracy given the stable population but also the longitudinal datasets by reducing the chance of creating erroneous longitudinal patterns. Cross-sectional and longitudinal estimates using the current and proposed imputation methods will be compared.

Evaluating Imputation Models for Totals Considering Missing Patterns

Young-Jun Kweon, *Bureau of Transportation Statistics*

Imputation is performed to keep as many observations as possible by replacing missing values with estimated values and the resulting complete data can be used by standard statistical methods. In some cases, the goal of imputation is not so much about adequate estimates at an individual unit but the total across all units. The Bureau of Transportation Statistics (BTS) conducted imputation analysis on annual boarding counts in the National Ferry Database (NFD). About 35% of the reported ferry segments have missing values in passenger and/or vehicle boardings and are not missing completely at random. Several imputation models were developed and evaluated for estimating the national total boarding considering missing patterns. The presentation introduces the NFD, discusses the evaluation setup, and present the results.

Session F-1: Who's Keeping the Score . . . on Data Quality

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

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

Location: *Chesapeake A*

An Introduction to the FCSM Framework for Data Quality

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

Data quality is an important component of any study. Effective understanding of data quality is essential for public officials, private businesses, and the public to make data-driven decisions. Inferior data quality can result in misleading information and poor decisions. A clear, documented understanding of the strengths and weaknesses of data assures attention to ameliorating the weaknesses, enable appropriate users of the data, and reinforces the credibility of the data and their use. The Federal Committee on Statistical Methodology's Framework for Data Quality can be used to assess the data quality of a variety of different data products and sources. This presentation will provide an overview of the framework and introduce a companion piece to the framework, a handbook of case studies that describe the process of implementing FCSM's framework into their survey processes.

Rule-Based Data Validation and Reconciliation of Survey Responses

Gunnar Ingle, *Summit Consulting, LLC*

Albert Lee, *Summit Consulting, LLC*

Ryan Kling, *Summit Consulting, LLC*

Balint Peto, *Summit Consulting, LLC*

Each year the US Department of Agriculture's National Agricultural Statistics Service (NASS) conducts more than a hundred surveys to understand and enumerate every aspect of agriculture in the United States. The quality of survey responses varies with survey and respondent. Ensuring that survey responses are valid, reliable, and internally consistent is vital to publishing accurate official statistics. NASS is undertaking modernization efforts to detect and edit survey responses through rule validation. These innovations include (1) a review and reconciliation of documented (e.g., written in business rules) and undocumented (e.g., only appearing in programming code) validation specifications, (2) distinguishing validation rules whose errors might be correctable with logical programming code or numeric methods, (3) using numeric methods, such as the Fellegi-Holt algorithm, and R software packages developed by Statistics Netherlands to automate response-level validation checks and error corrections, and (4) flagging instances of automated correction or validation errors for NASS analysts. This paper will describe the processes and procedures used for each step and highlight challenges and solutions to issues commonly encountered.

A Data Quality Scorecard to Assess a Data Source's Fitness for Use

Elizabeth Mannshardt, *National Center for Science and Engineering Statistics*

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

Lisa B. Mirel, *National Center for Science and Engineering Statistics*

Julie Banks, *NORC at the University of Chicago*

F. Jay Breidt, *NORC at the University of Chicago*

Benjamin R. Peck, *NORC at the University of Chicago*

Kiegan Rice, *NORC at the University of Chicago*

Zachary H. Seeskin, *NORC at the University of Chicago*

Lance A. Selfa, *NORC at the University of Chicago*

Grace Xie, *NORC at the University of Chicago*

Assessing data quality is critical to determine a data source's fitness for use. The Federal Committee on Statistical Methodology (FCSM) has developed a data quality framework that provides a common language

for federal agencies and researchers to make decisions about data quality. The FCSM framework is based on three core domains - utility, objectivity, and integrity – each having varying dimensions. Recently, the National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation and NORC at the University of Chicago collaboratively developed an approach to generate scorecards based on the eleven dimensions outlined in the FCSM Data Quality Framework. Our “Federal Data Quality Assessment Framework” (FDQAF) scorecard generates a score in each of the eleven FCSM data quality dimensions for a specific use case via binary questions about the use case, its data sources, and relevant metadata. The overall score sums the rescaled scores across the eleven dimensions. Higher scores reflect overall higher data quality with dimension scores giving insight on strengths and weaknesses of a data source for that use case. We illustrate the FDQAF scorecard approach with NCSES use cases and describe some of the challenges in developing a multipurpose assessment system.

Wednesday, October 25, 2023

1:45 PM

Session F-2: Dynamics of Data Linkage: Techniques and Challenges

Organizer: Tamara Lee, *Department of Veterans Affairs*

Chair: Tamara Lee, *Department of Veterans Affairs*

Location: Chesapeake B

Measuring Innovation in Multinational Enterprises

Allison Derrick, *Bureau of Economic Analysis*

Innovation is an essential activity of successful multinational enterprises (MNEs). They innovate to develop new goods and services, improve production processes, increase productivity, and tailor products for local markets. To manage the resulting intellectual property (IP) throughout their global operations, MNEs must balance competing considerations such as tax liabilities, skilled labor requirements, and IP protections. Yet, innovation within these MNEs is difficult to observe and measure for statistical and research purposes. The main challenge of measuring innovation is that the resulting knowledge, expertise, or idea is intangible; quantifying and valuing it is more complicated than quantifying and valuing physical assets like machinery or real estate. Then, each stage of innovation—creation, ownership, and use—often takes place in different locations by different units of the MNE. While most R&D is still conducted by the parent, increasingly, the resulting IP products are legally held elsewhere for tax purposes. Because intangible assets like IP products can be used by more than one entity in the MNE at a time, which unit owns the IP product and which units are using it in production is opaque. Transfer pricing and long time horizons further complicate measurement attempts. This paper combines multiple MNE survey data sources from the U.S. Bureau of Economic Analysis and patent application and acquisition data to improve measurement of innovation-related activities in U.S. MNEs.

Use of Self-Generated Identification Codes for Data Linkages in Highly Clustered Samples Involving Youth Populations

Chidinma Nwankwo, *Mathematica*

Tiffany Waits, *Mathematica*

Avery Hennigar, *Mathematica*

Lourdes Fernandez, *Mathematica*

Amy Harris, *Mathematica*

William Waddell, *Mathematica*

Sam Boneli, *Mathematica*

Elizabeth Cavadel, *Mathematica*

Researchers have used self-generated identification codes (SGICs) as a reliable method to link responses between various time points (Calatrava et al., 2022). These codes consist of elements unique to the

respondent that can be used to link data from measures used at two different points in time when no other identification method is possible or when anonymity is critical. This is especially important for adolescent populations, who may be hesitant to respond truthfully to survey questions on sensitive topics without the assurance of anonymity (Calatrava et al., 2022). Our research team conducted a pilot study to measure adolescent self-regulation between two time points in a highly clustered sample of high-school age youth. We used eight unique self-identification items, five for each survey, using different combinations of questions to test the effectiveness of the data linkage. This study shares preliminary findings from two research questions. First, what is our overall SGIC matching rate in a highly clustered sample of youth in schools? Second, how do these matching rates compare to other previously used SGIC matching rates? Presenters discuss how questions were selected and refined, pre-tested, and analyzed and share implications for use of SGICs in highly clustered samples of youth.

Enabling Economic Statistics Modernization

Jessica Wellwood, *U.S. Census Bureau*

Erica Marquette, *U.S. Census Bureau*

Ali Obaidi, *MITRE*

Adrienne Chen-Young, *MITRE*

The U.S. Census Bureau is prototyping a vision of an integrated infrastructure containing a comprehensive list of government and non-government businesses, with their core attributes, in a Business Frame that will enable linkages to respondent data, administrative data, and other enterprise demographic, geospatial, and job data. The Business Frame will assess challenges with merging records from multiple sources using identity matching algorithms to disambiguate data and capitalize on the similarities in data currently distributed across the enterprise, to create a robust centralized repository defining the complex population and relationships among businesses. Existing statistical programs can leverage this integrated business environment as they look to meet the demand for cross-domain data and increased granularity of attributes needed for economic data products. This presentation will describe the motivation behind the vision to strengthen research capabilities and promote innovation via the Business Frame. Further it will discuss the approach and findings in designing the data architecture for this modernized environment that will expand the capacity to answer critical questions about the nation's economy.

Linking the 1980 Post-Enumeration Program Samples to the Decennial Census

Kelsey Drotning, *U.S. Census Bureau*

Katie Genadek, *U.S. Census Bureau*

The Decennial Census Digitization and Linkage (DCDL) project is a large-scale effort to digitize and link the 1960-1990 Censuses and include them in the Census Bureau's data linkage infrastructure. When these Censuses were collected, names were not digitized preventing linkage to other datasets. Existing record linkage methods rely on either identification keys or fuzzy matching of string variables like names to uniquely identify records and evaluate the best match. In this paper, we present a case where these techniques have limited use for our record linkage goal because the selected datasets lack identification keys and string variables, providing a case study for DCDL. We link three sample datasets which contain names to the 1980 Decennial Census: the 1980 April Current Population Survey, the 1980 Enumeration Sample, and the 1980 August Current Population Survey. We use and assess both deterministic and probabilistic record linkage algorithms. We find that record linkage using limited identifying information presents challenges when using standard linkage techniques, resulting in low linkage rates. This paper will inform future attempts to link historical datasets with limited identifying information.

Agree to Disagree? Comparing IRS, NCOA, and Census Bureau Survey Migration Measures

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

Mark Ellis, *University of Washington*

Lee Fiorio, *University of Washington*

Why do migration estimates from the American Community Survey (ACS) and the Current Population

Survey's Annual Social and Economic Supplement (CPS-ASEC) differ? And why do these survey estimates differ from measures derived from Internal Revenue Service (IRS) tax records and the United States Postal Service's (USPS) National Change of Address (NCOA) registry? While differences in aggregate migration rates stem in part from variation in populations covered and definitions of key concepts, we show that they also reflect significant disagreement across sources in migration reporting at the individual level. We use IRS and NCOA administrative records linked with ACS and CPS-ASEC responses to quantify the extent to which survey migration reports (dis)agree with address changes voluntarily submitted to the USPS via the NCOA registry and/or captured in IRS 1040s, 1099s, and W-2s. Among those changing addresses in IRS or NCOA records in the one-year period referenced in a survey response, roughly two-thirds of ACS householders report having lived at a different residence one year ago, but fewer than half of CPS-ASEC householders do so. We document demographic and socioeconomic variation in migration (dis)agreement and discuss the implications of our findings for understanding long-term migration trends in the United States.

Wednesday, October 25, 2023

1:45 PM

Session F-3: Advancements in Sexual Orientation and Gender Identity Measurement

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

Chair: Christina Dragon, *National Institutes of Health*

Location: Chesapeake C

Sex and Gender Identity Measurement in Criminal Justice System Administrative Data: Evidence from the Instrument Redesign Efforts for the Bureau of Justice Statistics' Survey of Sexual Victimization

Emily Buehler, *Bureau of Justice Statistics*

Greta B. Clark, *U.S. Census Bureau*

Krysten Mesner, *U.S. Census Bureau*

Jessica A.R. White, *U.S. Census Bureau*

The Bureau of Justice Statistics (BJS) conducts the Survey of Sexual Victimization (SSV) annually, in part to fulfil the agency's mandate under the Prison Rape Elimination Act (PREA; P.L. 108-79). The SSV collects data about allegations and substantiated incidents of sexual abuse and harassment from adult correctional and juvenile justice facilities. In 2023, BJS and its data collection agent, the U.S. Census Bureau, cognitively tested revised survey instruments with facility administrators. Specifically, the revised instrument included a two-item approach to the measurement of sex and gender identity of victims and perpetrators of incidents of sexual victimization. While there are recommendations of best practices for the collection of self-reported sexual orientation and gender identity data on federal statistical surveys, further research is necessary into how administrative records capture these data. The findings from this research will detail the challenges and opportunities of collecting both sex and gender identity data from administrative responders in the criminal justice system and implications for other types of establishment surveys will be discussed.

Collecting SOGI Data for the Federal Workforce

Renee Ellis, *U.S. Census Bureau*

Christina Dragon, *National Institutes of Health*

Numerous factors drive the need for sexual orientation and gender identity (SOGI) workforce demographic data collection, including analyzing potential barriers to equal employment opportunity as well as assessing and informing diversity, equity, inclusion, and accessibility efforts. In FY21, the Equal Employment Opportunity Commission received 1,968 charges alleging discrimination based on sexual orientation or gender identity. Recently, the U.S. Supreme Court ruled that Title VII of the Civil Rights Act of 1964 prohibits employment discrimination on the basis sexual orientation or gender identity. However, despite Federal

protections, 46% of U.S. LGBTQ+ employees are closeted at work. The Federal interagency LGBTQ+ employee resource group, Pride in Federal Service (PFS), provides a rich environment for exploring Federal employees' attitudes and perceptions around disclosing their SOGI demographic information. We developed a study to measure (1) Federal employee potential willingness to disclose SOGI status and (2) the impact of state-level anti-LGBTQ+ laws and proposed bills on Federal employee potential willingness to disclose SOGI. This presentation will illuminate findings about LGBTQ+ federal workers attitudes and perceptions about SOGI data collection in the workplace and highlight strategies to allay potential concerns.

Updates on A Research Agenda for the Sexual Orientation and Gender Identity Questions on the Experimental Household Pulse Survey

Zachary Scherer, *U.S. Census Bureau*

Jacqueline Banks, *U.S. Census Bureau*

The U.S. Census Bureau added sexual orientation and gender identity (SOGI) questions to its experimental Household Pulse Survey (HPS) in July 2021. The Census Bureau is undertaking a research agenda for the SOGI questions on the HPS to better assess the quality of the already-collected data and explore how the question sequence could be refined to further enhance data quality. First, the Census Bureau is evaluating breakoff data to determine whether respondents are more likely to exit the survey on the SOGI questions. Second, the Census Bureau conducted a split-panel test of a write-in response line for the sexual orientation question on the HPS beginning with Phase 3.7 of the survey and will be evaluating these data. Finally, the Census Bureau is carrying out a record linkage project to connect HPS responses to the Decennial Census and the Social Security Administration's Numident file to answer questions such as whether sex responses from the Decennial Census align more closely with HPS sex assigned at birth or current gender identity responses, for example. This presentation will highlight these projects and present results to the extent that they are available.

Just to Confirm: Evaluating the Reliability and Validity of Survey Questions on Sex and Gender

Elise Christopher, *National Center for Education Statistics*

David Richards, *National Center for Education Statistics*

Maura Spiegelman, *National Center for Education Statistics*

Harper Haynes, *RTI International*

Surveys may ask respondents to confirm responses to sex assigned at birth and gender to minimize response errors, which can lead cisgender respondents to be falsely categorized as gender minorities, inflate population estimates, and yield inaccurate conclusions about transgender individuals. However, probing individuals who report a gender that is different from their sex assigned at birth to confirm this pair of answers can make transgender respondents feel singled out or "othered." Asking the same confirmation of all respondents, including those who report the same answer for their sex and gender, does not necessarily reduce that stigma for transgender individuals, since they would not know that all respondents were asked to confirm their answers. We tested two separate confirmation questions, asking respondents about their sex assigned at birth, to confirm their answer, then about their gender, and to again confirm their answer. Respondents were re-interviewed 3 weeks later to confirm their responses. We assess how frequently respondents changed an answer, whether their responses indicate they are cisgender or transgender, and whether they provide the same responses during re-interview. These findings will inform whether asking two confirmation questions of all respondents can increase data quality without causing respondents to feel singled out.

Testing and Implementing SOGI Questions in the American Community Survey

Elizabeth A. Poehler, *U.S. Census Bureau*

Jessica Holzberg, *U.S. Census Bureau*

Amy Symens Smith, *U.S. Census Bureau*

The U.S. Census Bureau is testing the inclusion of sexual orientation and gender identity (SOGI) questions in the American Community Survey (ACS). As a multi-mode survey, the ACS test will include internet, paper, and

personal-interview responses. A split-panel test will be conducted that will also include testing of question wording and categories, along with alternatives to write-in collection in the internet mode. In particular, the ACS will focus on evaluating the feasibility and quality of within household proxy response for the SOGI questions. This presentation will discuss the plans for the test and future implementation of SOGI content on the ACS.

Wednesday, October 25, 2023

1:45 PM

Session F-4: Democratizing Data: A Search and Discovery Platform for Public Data Assets

Organizer: Spiro Stefanou, *Economic Research Service*

Chair: Nancy Potok, *NAPx Consulting*

Discussant: Vipin Arora, *Bureau of Economic Analysis*

Location: Vessey 1

Overview of the Search and Discovery Platform and the Data Science Behind It

Nancy Potok, *NAPx Consulting*

Theory of Change Showing Value of ERS Statistical Data

Spiro Stefanou, *Economic Research Service*

Tackling Declining Response Rates: Increasing Response Rates by Farmers Because They Can See How Their Responses Can be Used to Benefit Their Own Businesses as Well as the Public

Hubert Hamer, *National Agricultural Statistics Service*

Forging New Partnerships: A vision for Education Statistics

Josh DeLaRosa, *National Center for Education Statistics*

NCSES: Making Connections Across the Data Ecosystem

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

Session F-5: Testing a New Combined Race and Ethnicity Question in Household Surveys: Agency Methods and Results

Organizer: Rebecca L. Morrison, *Bureau of Labor Statistics*

Chair: Rebecca L. Morrison, *Bureau of Labor Statistics*

Discussant: Darby Steiger, *SSRS*

Location: Vessey 2

Quantitative Testing Results from a New Race and Ethnicity Question from the Bureau of Labor Statistics

Tywanquila Walker, *Bureau of Labor Statistics*

Rebecca L. Morrison, *Bureau of Labor Statistics*

Robin Kaplan, *Bureau of Labor Statistics*

Victoria Narine, *Bureau of Labor Statistics*

Erica Yu, *Bureau of Labor Statistics*

In a full factorial, between-subjects experimental design implemented in two web surveys, respondents were randomly assigned to experimental conditions. In the first condition, respondents received one of two questions (“How would you describe your race or ethnicity?” vs. “How would you describe your race and/or ethnicity?”). In the second condition, respondents received either longer or shorter instructions (“Select all that apply” vs. “Select all that apply AND enter additional details in the spaces below. Note, you may report more than one group.”). In addition, all participants received a short set of probes on how well the combined race and ethnicity question reflected the respondents’ identification, and a few questions concerning the relationship between cultural origins and identity. The authors will summarize the findings from the methodological experiments and follow-up probes.

Results from Qualitative Testing of a New Race and Ethnicity Question from the National Agricultural Statistics Service

Kathy Ott, *National Agricultural Statistics Service*

Struther Van Horn, *National Agricultural Statistics Service*

Two types of testing were conducted with the farm and ranch population, cognitive interviews to mimic paper data collection and cognitive interviews to mimic telephone data collection. Participants in both types of testing were exposed to either a single race/ethnicity question or a set of two race/ethnicity questions that asked more details. In addition, participants were exposed to either longer or shorter instructions (“Select all that apply” vs. “Select all that apply AND enter additional details in the spaces below. Note, you may report more than one group.”). After answering the test questions, all participants received a short set of probes that asked about what they thought the question(s) were asking, how they chose their answer, whether the categories presented accurately described their race and/or ethnicity, and how easy or difficult it was for them to report for another person involved in decisions for their farm or ranch population. The authors will summarize the methods used and the findings from the qualitative testing.

Reflecting a Diverse America: Qualitative Testing of Changes to Race and Ethnicity Collection in the United States

Victoria Dounoucos, *RTI International*

Y. Patrick Hsieh, *RTI International*

Daniela Glusberg, *RSS*

Alisú Schoua-Glusberg, *RSS*

Aleia Clark Fobia, *U.S. Census Bureau*

Jennifer Childs, *U.S. Census Bureau*

RTI International and RSS, under the guidance of the U.S. Census Bureau, conducted approximately 100 cognitive interviews in English and Spanish to evaluate the proposed revisions to OMB's 1997 Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity (SPD 15). The cognitive testing thoroughly explored the updates to the race and ethnicity question format, terminology and wording of questions, as well as the instructions for respondents and associated guidance. Cognitive testing examined how participants interacted with and responded to two versions of the combined race and ethnicity question. The primary goals of cognitive testing were to (1) understand how the combined race and ethnicity question format affects response distributions, respondent interpretation, self-identification, and understanding; and (2) assess whether changes to the question stem and simplification of instructions affects respondent understanding and response. Recruitment of participants was spread across the broad racial and ethnic groups defined in the SPD15 Revision and to specifically include both respondents who identify with the detailed categories listed on the form and respondents who do not identify with those particular categories. This presentation will provide a broad overview of findings.

A Cognitive Interview Evaluation of the Proposed Combined Race and Ethnicity Question for Statistical Policy Directive 15, Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity

Stephanie Willson, *National Center for Health Statistics*

The National Center for Health Statistics conducted a cognitive interview evaluation of the combined race and ethnicity question as proposed in the Federal Register Notice. One hundred interviews were conducted in English and 50 in Spanish. Interviews were conducted in a virtual, one-on-one setting using the Zoom meeting platform. The question was tested in both interviewer- and self-administered formats.

Respondents' race and ethnic identities were explored in relation to how they chose to answer the question. In addition, several research questions guided follow-up probing and will be addressed in this presentation, including 1) clarity of question instructions, specifically the mark-all-that-apply option and the combination of race and ethnicity, 2) mode effects, 3) the adequacy of (and any confusion over) the subcategories associated with each main category, 4) interpretive differences by language, 5) interpretive differences by group, 6) the functioning of the new MENA category, and 7) the adequacy of the main group/race options in the absence of an 'other' category.

Wednesday, October 25, 2023

1:45 PM

Session F-6: Applications of Modeling and Post Randomization to Improve Estimation

Organizer: Jennifer Rhorer, *National Agricultural Statistics Service*

Chair: Jennifer Rhorer, *National Agricultural Statistics Service*

Location: Room 0105

On Response Model Selection and its Use for Testing Bias of Mean Estimates in Case of Not Missing at Random Nonresponse

Michael Sverchkov, *Bureau of Labor Statistics*

Sverchkov and Pfeffermann (S-P, 2008, 2018) consider estimation under informative sampling and not missing at random nonresponse. To account for the nonresponse, S-P assume a given response model and estimate the corresponding response probabilities by application of the Missing Information Principle, which consists of defining the likelihood as if there was complete response and then integrating out the unobserved outcomes from the likelihood employing the relationship between the sample and sample-complement distributions. A key condition for the success of this approach is the specification of the response model. In this presentation we consider information criteria based on the above likelihood, and show how they can be used for the selection of the response model. We apply this approach for testing potential bias in Consumer

Expenditure mean total estimates.

Two Not-So-Different Approaches for Dealing with Not-Missing-at-Random Unit Nonresponse

Phillip S Kott, *RTI International*

Unit nonresponse can cause bias in survey-derived estimates when the probability of unit response is correlated with the variable(s) being estimated. If a variable being estimated can be modeled, and unit nonresponse is ignorable or, equivalently, the unit is missing at random, then traditional methods can be used to remove (asymptotically) the bias from the estimate of the variable's population mean. When unit nonresponse is not ignorable or, equivalently, the unit is not missing at random, then pattern-mixture modeling and SUDAAN's WTADJX provide two different approaches for dealing with unit nonresponse. We will examine these approaches and see that they are more similar than they appear to be even though pattern-mixture modeling focuses on the model for the variable being estimated while WTADJX implicitly models the probability of unit response. Each approach has its advantages and disadvantages. We will then discuss whether these approaches can, in the absence of measurement error, satisfactorily measure the (potential for) bias in more traditional estimates.

Predicting the Effect of Business Births and Deaths on the Current Employment Statistics Survey: Using Sample Information to Minimize Coverage Error

Steve Mance, *Bureau of Labor Statistics*

Chris Grieves, *Bureau of Labor Statistics*

Collin Witt, *Bureau of Labor Statistics*

The Current Employment Statistics (CES) program uses a model to account for the bias in monthly payroll employment estimates that arises from establishment births and deaths that fall outside the survey frame. Actual birth-death values calculated from administrative counts are available with a substantial lag but need to be predicted for the current month's estimation. The existing CES net birth-death model relies on the bias from business births and deaths to follow a consistent, seasonal pattern, characterized by an ARIMA process. This has broken down during extreme changes in the labor market and necessitated interventions in the model during the COVID-19 recession and recovery. Previous research showed the ARIMA models can be improved by adding covariates available coincident with the survey, and this paper explores several modelling frameworks that use information from the CES survey to predict birth-death values, substantially outperforming previously examined models over the period from 2007–2021, covering the Great Recession through the pandemic recovery. Forecast combination techniques are also examined and compared with predictions from the individual frameworks.

Maximum Likelihood Estimation of Response Propensity to a Nonprobability Survey

Vladislav Beresovsky, *Bureau of Labor Statistics*

Terrance Savitsky, *Bureau of Labor Statistics*

Matthew R. Williams, *RTI International*

Julie Gershunskaya, *Bureau of Labor Statistics*

Non-probability, or convenience, surveys attract significant attention as a less expensive and more expedient alternative to probability surveys for estimating population characteristics. For design-based inferences a convenience sample may be treated as selected by Poisson sampling from the population of interest. While the response indicator to a non-probability survey is unobserved, response propensity may be implicitly estimated by maximizing Bernoulli likelihood of the observed indicator of the convenience sample on the combined set of convenience and reference probability samples from the same population. We utilize the direct relationship between the Bernoulli likelihood parameter mentioned above, response propensity to the non-probability survey, and selection probability of the reference probability sample. We prove consistency of the proposed likelihood-based methods and compare their efficiency to the pseudo-likelihood methods, theoretically and in simulations, for different sample sizes, sample fractions and degree of overlap between convenience and reference samples in covariate space. We demonstrate higher efficiency of the AIC model selection criteria for the proposed response propensity estimation methods.

Small area estimation of forest carbon using carbon maps and model-assisted regression in Maryland

Andrew Lister, *USDA Forest Service*

George Hurtt, *University of Maryland*

Lei Ma, *University of Maryland*

Ty Wilson, *USDA Forest Service*

High resolution carbon maps are becoming ubiquitous as interest in terrestrial carbon dynamics and modeling capabilities increase. The state of Maryland, USA, is among several that have invested heavily in developing a carbon monitoring system that relies on carbon maps. These maps, which are developed with various types of remotely sensed data like LiDAR and satellite imagery, provide information on carbon stock and dynamics at a fine scale. However, these maps are rarely used in a probabilistic sampling framework that allows for the creation of traditional sample-based uncertainty estimates such as confidence intervals of estimates. The USDA Forest Service's Forest Inventory and Analysis (FIA) program, on the other hand, is based on a probability sample of ground plots. Traditional probability sampling and post-stratified estimation, using canopy cover stratum maps, are used to generate acceptably-precise estimates for larger areas such as groups of counties or the state as a whole. There exists a need to improve FIA estimation precision at finer geographic scales. In the current study, we present results of an assessment of the benefit of using model-assisted regression with Maryland's carbon maps compared to post-stratified estimation using FIA's standard methods. The goal of the study is to develop an estimation framework that fits into FIA's production system and allows for improved precision by leveraging wall-to-wall high resolution carbon maps. The framework could also be used by states like Maryland that use carbon maps in their monitoring systems.

Wednesday, October 25, 2023

3:30 PM

Session G-1: If We Modernize the Survey, Will They Respond?

Organizer: Doug Williams, *Bureau of Labor Statistics*

Chair: Robin Kaplan, *Bureau of Labor Statistics*

Location: Chesapeake A

Rapid Response Surveys: Household Pulse Survey 2.0

Hyon B. Shin, *U.S. Census Bureau*

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

Cassandra Logan, *U.S. Census Bureau*

Jason Fields, *U.S. Census Bureau*

The U.S. Census Bureau has been conducting the experimental Household Pulse Survey (HPS) since April 2020 to provide near real-time information on household impacts due to the COVID-19 pandemic. As a proof-of-concept, the HPS demonstrated many strengths but some weaknesses in collecting and disseminating data in a rapid way. In the three years since its inception, the HPS has gone through many changes. It also garnered accolades for its innovative way of providing near real-time data that traditional surveys are not designed to provide. The HPS provided timely and relevant data but also had low response rates and some nonresponse bias. There is room for improvement so building on the successes of HPS and looking forward to making improvements to representativeness, the Census Bureau created a dedicated team to manage the Demographic High Frequency Surveys Program (DHFS). The next iteration of the DHFS is the Military Panel which utilizes a dedicated panel of active-duty, and spouses of, military personnel. Another iteration is the Census Household Panel which will collect timely data from a dedicated, nationally representative household panel. This presentation will cover the evolution in the DHFS and highlight both of these new surveys.

National Hospital Care Survey: Redesigning the Annual Hospital Interview

Catherine Rappole, *National Center for Health Statistics*

Geoffrey Jackson, *National Center for Health Statistics*

The National Hospital Care Survey (NHCS) Annual Hospital Interview (AHI) collects summary administrative hospital data, such as annual visit counts, to supplement the NHCS's electronic collection of inpatient and emergency department encounter data. Results from the NHCS AHI are used to inform the development of national weights for the NHCS. Due to budget restrictions, the NHCS AHI was pulled out of the field in 2016; however, it was re-instated in 2020 with the original interviewer-administered questionnaire now implemented as a self-response web questionnaire. The 2020 AHI questionnaire was not updated to take advantage of the change in the mode of administration and had a lower than expected response rate of 14%. Prior to the 2021 NHCS data collection, the AHI questionnaire and electronic survey user interface underwent a significant redesign to simplify questions, reduce respondent burden, and add new topics of interest. The 2021 redesign had an immediate impact with a 74% increase in the number of hospitals that completed the survey compared to the 2020 AHI. This presentation will discuss the redesign of the AHI and highlight selected results from the 2020-2022 AHI.

Data Quality in the Online Panel Self-Administered Diary Test of the Consumer Expenditure Survey

Graham Jones, *Bureau of Labor Statistics*

Nikki Graf, *Bureau of Labor Statistics*

Wendy Carlton, *Bureau of Labor Statistics*

Tucker Miller, *Bureau of Labor Statistics*

The U.S. Bureau of Labor Statistics' Consumer Expenditure (CE) Survey program recently partnered with Ipsos to conduct a test of a fully self-administered expenditure diary using a probability-based online panel. In CE's Online Probability Self-Administered Diary (OPSAD) test, panel respondents were invited to complete a self-administered Household Characteristics Survey and a subsequent two-week Spending Diary, without the aid of an interviewer. To better understand the data quality implications of the OPSAD test design, the results were compared to CE diary production data from the corresponding months (November 2021 to January 2022). This presentation will cover analytical findings from the OPSAD test, including relevant comparisons to CE production data. Along with comparisons of diary completion rates and household demographic composition, this presentation will focus on analyzing data quality between the OPSAD and CE production diary cases. Data quality measures will include, but are not limited to, item nonresponse, expenditure entry counts, expenditure amounts, and non-itemization/bundling of entries. Preliminary findings indicate that expenditure counts and total reported expenditure amounts were generally higher in CE production diary cases than in OPSAD.

Comparison of Paper-and-pencil Versus Tablet Administration of the 2021 National Youth Risk Behavior Survey (YRBS)

Jingjing Li, PhD, MD, MPH, *Centers for Disease Control and Prevention*

Adriana Rico, MPH, CPH, *Centers for Disease Control and Prevention*

Nancy Brener, PhD, *Centers for Disease Control and Prevention*

Alice Roberts, MS, *ICF International*

Jonetta Mpofo, PhD, *Centers for Disease Control and Prevention*

Mike Underwood, PhD, *Centers for Disease Control and Prevention*

BACKGROUND: Aligning with CDC's Public Health Data Modernization Initiative, the national Youth Risk Behavior Survey (YRBS) is transitioning from paper-and-pencil instrument (PAPI) to electronic administration using tablets. This study examines the reporting of demographic characteristics and 89 health behaviors and experiences between the PAPI- and tablet-administered 2021 national YRBS questionnaires. **METHODS:** The 2021 national YRBS applied a three-stage, cluster sample design to obtain a nationally representative sample of U.S. students in grades 9–12 in the 50 states and the District of Columbia. During 2021, 57 participating high schools randomly assigned students to complete YRBS using PAPI (n=4684) or tablets (n=3645). Prevalence and missingness in demographics characteristics (sex, race/ethnicity, grade,

and sexual identity) as well as in 89 health behavior and experience items were compared by administration mode (PAPI vs. tablet).

RESULTS: Demographic characteristics did not differ by mode. The majority (93.2%) of health behavior and experience items showed no difference in prevalence by mode, adjusting for demographics. However, 29.2% of items exhibited significant differences in missingness by mode.

CONCLUSIONS: Survey administration mode was not significantly associated with behavior and experience reporting among high school students. Findings provide evidence to support electronic (tablet) survey administration for the national YRBS.

Changes in Survey Methods during the COVID-19 Crisis and Data Quality: the Case of the Spanish Survey of Household Finances (EFF) 2020

Elena Vozmediano, *Banco de España*

Laura Crespo, *Banco de España*

Sandra García-Urbe, *Banco de España*

The Spanish Survey of Household Finances provides detailed information on the income, assets, debt and spending of Spanish households every three years since 2002. Owing to the complexity and the length of its questionnaire, the data are collected through personal interviews with the households, conducted by interviewers with specific training and computer-assisted (CAPI). In this sense, the EFF 2020 represents a unique wave as a result of its concurrence with the COVID-19 pandemic which forced the revision of some methodological dimensions, mainly the interviewing mode, which had to be switched from CAPI to CATI. In this paper, we provide a detailed description of the methodological changes implemented in the EFF2020 as a result of the COVID-19 crisis and we estimate the overall impact on different aspects of data quality. This evaluation is based on the comparison of several data quality indicators between the EFF2020 and its preceding wave, the EFF2017, which is its nearest and most comparable edition. Our results show that these changes did not seem to have an important impact on particular dimensions of the data such as the length of the interviews, the item non-response for monetary questions, respondents' attitudes and interest during the interview and potential biases when answering multiple response questions.

Wednesday, October 25, 2023

3:30 PM

Session G-2: Directions in Measuring Disability: Use and Evaluation of the Washington Group Short Set on Functioning

Organizer: Julie D. Weeks, *National Center for Health Statistics*

Chair: Laryssa Mykyta, *National Center for Health Statistics*

Discussant: Jennifer Madans, *Center for Inclusive Policy*

Location: Chesapeake B

An Introduction to the Washington Group Short Set on Functioning

Cordell Golden, *National Center for Health Statistics*

This paper describes advances in the definition of disability and disability measurement, introducing and discussing the development of the WG-SS and highlighting its use in international and federal surveys. In addition, the presentation will highlight how the WG-SS differs from the current set of questions currently used to measure disability in the ACS. One key difference between the WG-SS and the ACS question set is the response choices: the ACS questions have dichotomous Yes/No responses while the WG-SS questions allow respondents to select among four answer categories reflecting a continuum of disability (no difficulty, some difficulty, a lot of difficulty, cannot do at all) to better assess the extent to which an individual can participate in society.

Census Bureau Evaluations of Proposed Changes to the ACS Disability Questions

Natalie Young, *U.S. Census Bureau*

In 2020-2022, the U.S. Census Bureau conducted cognitive testing and a field test on proposed changes to the American Community Survey (ACS), including testing a disability measure based on the Washington Group Short Set on Functioning (WG-SS). The new measure would allow the ACS to produce internationally comparable estimates that are more consistent with current conceptualizations of disability. The cognitive testing was designed to 1) evaluate variations in the wording of items in the WG-SS, and 2) assess how respondents decide between the four response categories in the WG-SS. Results confirm findings regarding the WG-SS tested in other environments. Most respondents did not have a problem understanding or responding to the questions, regardless of variations in wording. Consistent with prior evaluations, only respondents who perceived themselves as having an impairment that severely curtailed their ability to complete activities on a daily basis selected response categories indicative of disability. The second part of the presentation will discuss further evaluation of the new measure within a split-ballot field test. The field test was designed to assess how the question set performs in the multi-modal ACS context. Evaluation metrics to be discussed include item missing rates, response distributions, and response reliability.

Comparing Survey Reports of Disability Using the WG-SS and ACS Question Sets

Julie D. Weeks, *National Center for Health Statistics*

Using data from the 2011-2012 National Health Interview Survey, this paper examines differences in survey reports of disability between two sets of disability questions, the Washington Group Short Set on Functioning (WG-SS) and the 6 questions currently used to measure disability in the American Community Survey (ACS). Results show that although estimated disability prevalence was higher in the ACS measure compared with the WG-SS measure, overall agreement is high given the high percentage of respondents reporting no difficulty on both question sets. A variety of sociodemographic and health factors contributed to the observed discordance. The population with disability defined by the ACS questions is more heterogenous in functional level than that defined by the WG-SS questions. The WG-SS ordinal response categories allow for an examination of disability severity, which is useful in describing the full continuum of functioning.

Wednesday, October 25, 2023

3:30 PM

Session G-3: Recent Advances in Data Privacy and Public Policy

Organizer: Harrison Quick, *Drexel University*

Chair: Jingchen Hu, *Vassar College*

Location: Chesapeake C

Recent Efforts in Statistical Privacy and Public Policy

Jingchen Hu, *Vassar College*

In the past year, a series of workshops on the theme of statistical privacy and public policy was held to improve and strengthen the relationships among the privacy experts, researchers, data stewards, data practitioners, and public policymakers. This is a collaboration between the National Institute of Statistical Sciences, Penn State University, Vassar College, and Urban Institute, and generously supported by the Alfred P. Sloan Foundation. In this talk, we would like to report key findings from the workshop series. In addition, we would like to share ongoing efforts from various working groups and call for those interested to get involved.

To Share, or Not to Share, That is the Question: Use of Synthetic Data Approach in Criminal Justice

KiDeuk Kim, *Urban Institute*

Having a criminal history, or being required to disclose it, can result in significant and long-lasting consequences, such as being denied voting rights and facing employment background checks. Decades of criminological research also indicate that merely being labeled as having a criminal record can have adverse effects on the individual and their family, as they may internalize this label. As such, criminal justice agencies are increasingly concerned with protecting the privacy of criminal records. Ironically, these agencies also have a responsibility to inform the public of potential risks to their safety by releasing identifying information about "high-risk" individuals. Sex offender registries are one such example, where both privacy loss and data utility can be maximized without having to make a trade-off. Recognizing this complexity of data privacy issues and responsibilities of criminal justice agencies, the current presentation discusses a conceptual framework for data protection in criminal justice. Findings from a case study on felony criminal cases are discussed, followed by a discussion on research and policy implications.

Do No Harm Guide: Applying Equity Awareness on Data Privacy Methods

Claire Bowen, *Urban Institute*

Researchers and organizations can increase privacy in datasets through methods such as aggregating, suppressing, or substituting random values. But these means of protecting individuals' information do not always equally affect the groups of people represented in the data. In this talk, I will introduce a guide that contains a literature review of equity-focused work in statistical data privacy (SDP) and interviews with nine experts on privacy-preserving methods and data sharing. These experts include researchers and practitioners from academia, government, and industry sectors with diverse technical backgrounds, where we sought to understand both how and to what extent they consider the questions of equity in their work. We also created an illustrative example to highlight potential disparities in public policy decision-making that can result from applying SDP methods without an equitable workflow.

Generating Differentially Private, Tract-level Synthetic Data for Births in Pennsylvania

Harrison Quick, *Drexel University*

Measures such as the percentage of births with low birthweight or the percent of births that were preterm are important indicators of the health of a community. While birth data may be available from individual states, data at the national level are limited to counties with population sizes greater than 100,000 on the Centers for Disease Control and Prevention's web-based WONDER tool, ostensibly to preserve privacy. Moreover, the data that are available are generally released at the county-level, which prohibits inference on neighborhood-level disparities in highly populated urban areas. To improve the utility of publicly available birth data while simultaneously providing improved privacy protections, we propose an approach for generating differentially private synthetic tract-level birth data using the recently developed Poisson-gamma framework, stratified by factors such as the age and race of the mother, the duration of the pregnancy, and the baby's birthweight. To illustrate our approach, we will consider tract-level birth data from Pennsylvania and compare our results to those from other, more conventional approaches for generating differentially private synthetic data (e.g., the Laplace mechanism). The talk will end with a brief discussion of some of the challenges encountered in this analysis and potential next steps.

***Session G-4: Advances in Implementing Title III of the Evidence Act:
Enhancing Public Trust in Official Statistics and Expanding Data Access for
Evidence Building***

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

Chair: Karin Orvis, *Office of Management and Budget*

Location: Vessey 1

Building Trust in Federal Statistics

Kerrie Leslie, *Office of Management and Budget*

The goal of improving the availability and use of evidence in policymaking requires trust; trust from data providers that the information that they provide for evidence purposes won't be used in a way that could violate their privacy and trust from the public and decisionmakers that the evidence is accurate and objective. To that end, the Evidence Act requires OMB to produce regulations designed to strengthen trust in the Federal statistical agencies' ability to protect the data they collect and to produce accurate and objective statistics. The regulatory requirement in the statute echoes the fundamental responsibilities of Federal statistical agencies outlined in OMB's Statistical Policy Directive No. 1 and specifies that every other Federal agency must enable and support the statistical agencies in order for them to meet these responsibilities. This presentation will cover the content of the Notice of Proposed Rulemaking for this regulation, referred to as OMB's Trust Regulation, as well as current progress towards finalizing the regulation.

Expanding Secure Access to Confidential Data Assets

Spiro Stefanou, *Economic Research Service*

The Evidence Act responds to the increased demand for data for evidence-building by policymakers, statistical agencies, and the public. The challenge is to balance the utility of providing access to high quality evidence to inform policy making against the need to protect the privacy and confidentiality of data providers. The risk of reidentification has increased for two primary reasons: 1) advancements in computing power; and 2) the availability of massive amounts of increasingly granular public and private sector data on individuals and organizations. A thorough and standardized assessment will allow agencies to determine the level of risk and, consequently, the appropriate mode of access to maintain adequate data security. The emerging regulatory goals focus on addressing the level of risk against the utility of access. To this end, not all data should be treated as equally sensitive leading to disclosure risk being treated as a continuum. Further, there is a shared responsibility between statistical agency and data users for the protecting data. Going forward, these regulatory goals should address both current and future data holdings of the Federal Statistical System.

Standard Application Process: Portal, Process, Governance

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

The Evidence Act set forth an ambitious goal of creating a standardized process for applying to use confidential federal data for evidence-building purposes. In response to this requirement, the federal statistical system launched the Standard Application Process (SAP) portal in 2022 as a single location where agencies, state, local, and tribal governments, researchers, and other individuals can discover and apply for secure access to confidential data. The SAP portal launch resulted from a coordinated effort among the federal statistical agencies and reflects a major milestone in the federal government's efforts to improve data access and strengthen evidence-building capacity in a privacy-protective manner. This presentation begins with an overview of the legislative requirements and the policy guidance that enabled the creation of technical specifications to build an SAP portal, discusses the SAP portal development including the critical role of stakeholder engagement, and concludes by describing the planned governance structure to guide the

strategic direction for the SAP.

The ACDEB's Vision of a National Service for Evidence Building

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

The Advisory Committee on Data for Evidence Building (ACDEB) articulated a vision for a National Secure Data Service (NSDS) as a philosophy, a service, and a place for the future of data sharing, linking, and application of privacy enhancing techniques across data providers and users including federal statistical and programmatic agencies, state and local government, the private-sector, and the public. This vision for the NSDS fits within a rapidly evolving data and evidence ecosystem with an infrastructure grounded by the principles and practices of a broad and highly decentralized, federal statistical system (FSS). With the ecosystem evolving rapidly, it is important to focus on opportunities that exist right now to lay the foundation for the future. NSF will lead a NSDS demonstration project designed to determine not only the feasibility of some of ACDEB's recommendations but also to further the vision of a seamless FSS that facilitates access to data for evidence building while ensuring privacy and transparency.

The Role of the Statistical Official in Advancing Evidence Building

Alex Marten, *Environmental Protection Agency*

Statistical expertise allows organizations to ensure that data are gathered, processed, and curated so as to produce statistical products with the highest standards of data quality while protecting confidentiality, privacy, and security. The Evidence Act required that each of the 24 CFO Act agencies must designate a Statistical Official to advise on statistical policy, techniques, and procedures. The Act ensures that each agency will have champion to help ensure that agencies can produce and disseminate relevant and timely statistical information, conduct credible and accurate statistical activities, conduct objective statistical activities, and protect the trust of information providers. This new role provides tremendous opportunity to advance the availability and quality of statistical evidence to inform decision-making, though not without challenges in identifying and implementing new policies and procedures within complex organizations. This presentation will discuss the new role of the Statistical Official, the opportunities it provides to strengthen evidence building, potential hurdles to success, and ways in which it could strengthen the Federal Statistical System.

Wednesday, October 25, 2023

3:30 PM

Session G-5: Innovations from the Criminal Justice Administrative Records System (CJARS) Project

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

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

Location: Vessey 2

An Overview of the Criminal Justice Administrative Records System

Keith Finlay, *U.S. Census Bureau*

The Criminal Justice Administrative Records System (CJARS) is a nationally integrated data repository designed to transform research and policymaking on the United States criminal justice system. At the University of Michigan, CJARS collects longitudinal electronic records from criminal justice agencies and harmonizes these records to track a criminal episode across all stages of the system. At the U.S. Census Bureau, harmonized criminal justice records can be linked anonymously at the person-level with extensive social, demographic, and economic information from national survey and administrative records. The CJARS project is a partnership between the Census Bureau and University of Michigan with the goal of expanding research and statistical reporting on the criminal justice system.

Presenting the Justice Outcomes Explorer (JOE)

Michael Mueller-Smith, *University of Michigan*

Jordan Papp, *University of Michigan*

The CJARS team is developing a public-facing data dashboard named the Justice Outcomes Explorer (JOE) with statistics about the criminal justice system and socioeconomic characteristics and outcomes of U.S. criminal justice caseloads. These statistics leverage the strengths of CJARS data and the data linkage opportunities at the U.S. Census Bureau. The statistics cover conviction, probation, incarceration, and parole caseloads. Outcomes are measured from longitudinal administrative data covering recidivism, labor market outcomes, and public program participation. This presentation will provide a description of methods of statistical production and a general description of JOE. Through this data dashboard, CJARS will shed new light on the socioeconomic lives of the people who interact with the justice system.

Socioeconomic Impact of Adverse Childhood Experiences in Young Adulthood

Shawn Ratcliff, *U.S. Census Bureau*

Keith Finlay, *U.S. Census Bureau*

Jordan Papp, *University of Michigan*

Megan C. Kearns, *Centers for Disease Control and Prevention*

Cora Peterson, *Centers for Disease Control and Prevention*

While research on adverse childhood experiences (ACEs) continues to flourish and evolve, there is limited systematic evidence on the long-term impacts of ACEs, particularly on how exposure during childhood/adolescence shapes socioeconomic well-being in (young) adulthood. Here, we explore how the exposure to various ACEs before the age of 18 (household member substance use disorder, incarceration, or mental illness; parental separation; parent death; witnessing intimate partner violence; physical/emotional abuse/neglect; or sexual abuse) shape multiple socioeconomic outcomes in young adulthood (18-23), including criminal justice involvement, education, employment, teen birth, and public assistance. Utilizing the U.S. Census Bureau Data Linkage Infrastructure, we measure ACEs exposure before age 18 among a sample of individuals born between 1999-2003. Specifically, we combine data from the Criminal Justice Administrative Records System, Census Household Composition Key, Decennial Census (2000, 2010), American Community Survey, and other survey and administrative data. Given we focus on experiences in young adulthood, all of our measures (excluding teen pregnancy) are observed between 2017 and 2021 and after the child is at least 18 years old. To examine these trends, we report both descriptive statistics and a series of logistic regressions to assess the net effect of each type of ACE on socioeconomic status in young adulthood, controlling for birth year, birth state, race/ethnicity, family structure, and parental income. The results presented here provide two major takeaways: First, we demonstrate the ability to measure ACEs using administrative data available using the U.S. Census Bureau's data linkage infrastructure – expanding work that has traditionally relied on self-reports in surveys. Second, our research demonstrates the longitudinal effect of ACEs on individuals' socioeconomic status, including the effect ACEs have above and beyond economic contexts (e.g., household income, poverty).

Wednesday, October 25, 2023

3:30 PM

Session G-6: America's DataHub Consortium: Building Capacity for the Future

Organizer: May Aydin, *National Center for Science and Engineering Statistics*

Chair: May Aydin, *National Center for Science and Engineering Statistics*

Location: Room 0105

America's DataHub Consortium Support of NCSES Priorities

May Aydin, *National Center for Science and Engineering Statistics*

This presentation will provide an overview of America's DataHub Consortium (ADC), how the ADC supports NCSES research goals and priorities, and a brief overview of NCSES research interests.

National Secure Data Service Demonstration Project

Heather Madray, *National Center for Science and Engineering Statistics*

This presentation will provide an overview of the National Secure Data Service Demonstration (NSDS-D) project including authorizing legislation, overarching goals, the relationship to a potential future NSDS, and activities to date.

Opportunities to Engage in ADC/NSDS-D Activities

Mike Atkinson, *ATI*

This presentation will provide information on opportunities to get involved with ADC and/or NSDS-D activities and end with a listening session where audience members will have an opportunity to provide input on planned activities.

Abstract Listings for Thursday, October 26

- **Concurrent Sessions H** 8:30 am – 10:00 am
- **Concurrent Sessions I** 10:30 am – 12:00 pm
- **Concurrent Sessions J** 1:45 pm – 3:15 pm
- **Concurrent Sessions K** 3:30 pm – 5:00 pm

Session H-1: Business Revenue and Employment Growth and Decline During the COVID-19 Pandemic Period: Evidence from the Business Dynamics Statistics

Organizer: Martha Stinson, *U.S. Census Bureau*

Chair: Cheryl Grim, *U.S. Census Bureau*

Location: Chesapeake A

The Business Dynamics Statistics: Describing the Evolution of the U.S. Economy from 1978-2021

Martha Stinson, *U.S. Census Bureau*

Christopher Goetz, *U.S. Census Bureau*

Spencer Knoll, *U.S. Census Bureau*

The Business Dynamics Statistics (BDS) is a statistical product of the U.S. Census Bureau that measures the entry and exit of businesses and the associated job creation and destruction. This product is unique due to its long time series (1978-2021), level of granularity (county by sector tables), and high-quality source data (universe of private-sector, non-farm businesses from administrative tax data). We use these data to investigate the differences in business start-up and exit rates across sectors, geographies, and time. We compare job growth from new businesses to that of existing businesses and document a long-term decline in the share of economic activity at young and small businesses and the increasing concentration of activity at old and large firms. Finally, we highlight how these trends changed during the COVID-19 pandemic period. We consider the impact of historically high business closure rates in 2020 on subsequent business openings and job creation and destruction in 2021.

Business Dynamics Statistics for Single-Unit Firms

Christopher Goetz, *U.S. Census Bureau*

Richard Beem, *U.S. Census Bureau*

Martha Stinson, *U.S. Census Bureau*

Sean Wang, *U.S. Census Bureau*

The Business Dynamics Statistics of Single Unit Firms (BDS-SU) is an experimental data product that provides information on employment and payroll dynamics for each quarter of the year at businesses that operate in one physical location. This paper describes the creation of the data tables and the value they add to the existing Business Dynamics Statistics (BDS) product. We then present some analysis of the published statistics to provide context for the numbers and demonstrate how they can be used to understand both national and local business conditions. First, we examine the features of the data from the COVID-19 recession in 2020, which are not evident in the core BDS. Specifically, we observe the heterogeneous effects on various industries, highlighting which sorts of businesses were particularly hard hit. We next examine business exit rates in detail and consider why different metro areas experienced the pandemic in different ways. We also consider entry rates and look for evidence of a surge in new businesses as seen in other data sources. Finally, we contrast the business dynamics from this recession with those from the Great Recession, as well as examine wider trends that occurred during the 2007-2021 time span.

Employment and Revenue Dynamics of SBA Covid Relief Recipients

Richard Beem, *U.S. Census Bureau*

Shawn Klimek, *Department of Commerce*

In a new experimental data product produced by the U.S. Census Bureau's Center for Economic Studies, the Business Dynamics Statistics (BDS) will be tabulated for firms that received SBA assistance during the

pandemic in the form of (1) Paycheck Protection Program loans, (2) COVID Economic Injury Disaster Loans, (3) Shuttered Venue Operators Grants, and/or (4) loans from the Restaurant Revitalization Fund. Loan-level microdata from the SBA will be matched to the universe of employer firms and their establishments in the Longitudinal Business Database (LBD). Like the standard BDS, this product will report statistics on firm and establishment entry/exit as well as job creation and destruction for recipient and non-recipient firms. Were firms that received SBA pandemic assistance less likely to shut down their operations in 2020 and 2021? Were annual revenues impacted differentially between recipient and non-recipient firms? Were firms that received larger loans and/or grants less susceptible to a slowdown in business dynamism? Was there a discernable difference in business dynamism trends for recipient firms in rural versus urban areas? Data users will be able to explore these questions and many more as subsequent annual updates to the data are released.

Thursday, October 26, 2023

8:30 AM

Session H-2: Linking NCES Data to Other Data Sets: Research, Opportunities, and Challenges Linking Records from NCES Data Sets to Other Data Sets

Organizer: Brian Cramer, *National Center for Education Statistics*

Chair: Brian Cramer, *National Center for Education Statistics*

Location: Chesapeake B

Linking NAEP and COVID Data Hub

Katherine Castellano, *Educational Testing Service*

Carol Eckerly, *Educational Testing Service*

Helena Jia, *Educational Testing Service*

The COVID-19 pandemic's unprecedented national disruption to student schooling since March 2020 has resulted in widespread concern about student learning, particularly about differential impact by student group. In the 2020-2021 school year, schools used a variety of instructional modes (e.g., in-person, remote, and hybrid). These modes often changed week-to-week or month-to-month within the same school in response to real-time shifts in the public health crisis (e.g., surges/outbreaks in communities could result in schools turning from in-person or hybrid schedules to fully remote). To understand better the impact of the instructional mode on student learning, various agencies collected data on instructional mode. Given the various modes used by schools throughout the school year, it was often difficult to capture the specific practices of the school. In this study, we relate the COVID-19 public school hub dataset (hub) by researchers at Brown University on percentage-of-the-year-virtual with the Monthly School Survey (MSS) collected by the National Center for Education Statistics and the Age 9 NAEP Long-Term-Trend (LTT) school administrator survey. We show how correspondence between these different data sources supported critical analyses using the MSS and LTT data.

Linking NAEP 2022 to the 2021 Monthly School Survey

Carol Eckerly, *Educational Testing Service*

Robert Finnegan, *Educational Testing Service*

Helena Jia, *Educational Testing Service*

Laura Jerry, *Educational Testing Service*

In response to President Biden's Executive Order on supporting the reopening and continuing operation of schools and early childhood education providers, the National Center for Education Statistics (NCES) administered the Monthly School Survey (MSS) to a nationally representative sample of schools in Spring 2021. The MSS questions included what learning modalities were offered (remote only, full-time in-person, hybrid), the numbers of students enrolled in each learning mode, and the amount of synchronous instruction

students had during remote instruction in 2021. In this presentation, we show a study linking the 2022 NAEP mathematics and reading grade 4 and 8 data to the MSS questionnaire data, based on the common set of schools that were sampled for the 2021 MSS and the 2022 NAEP operational assessments. The following overarching questions will be addressed via the linking study:

- How does student performance in 2022 differ by schools' responses to instruction or learning mode offered in 2021?
- How does student performance in 2022 differ by schools' responses to enrollment in instruction modes in 2021?
- For those students engaged in remote instruction (fully remote and/or hybrid), how does student performance in 2022 relate to the amount of synchronous instruction students had during remote instruction in 2021?

College Enrollment Benchmarks for the Grade 12 NAEP Mathematics Assessment

Burhan Ogut, *American Institutes of Research*

George Bohrnstedt, *American Institutes of Research*

Markus Broer, *American Institutes of Research*

This study linked ninth-grade student background data and school-reported high school and postsecondary transcript data from the national High School Longitudinal Study of 2009 (HSL:09) to student item responses on the 2013 National Assessment of Educational Progress (NAEP) mathematics assessment to examine the relationship between college preparedness and NAEP mathematics achievement. NAEP scale scores imputed for all HSL:09 participants via marginal maximum likelihood regression analyses were used to predict college preparedness. This study extended the earlier investigations of college preparedness on NAEP from the prediction of college enrollment alone to the prediction of additional postsecondary outcomes including remedial coursetaking and first-year college GPA based on postsecondary transcript data. Defining college preparedness as enrolling in a 2- or 4-year college without remedial coursetaking and acquiring a first-year college GPA of 2.7 or higher, as did the National Assessment Governing Board, this study showed that the probability of college preparedness increased from 26 percent for students performing at the NAEP Basic level to 58 percent at the NAEP Proficient level and 87 percent at the NAEP Advanced level.

Early Reading Skill Development and Characteristics of Reading Skill Profiles

B. Jasmine Park, *American Institutes of Research*

Xiaying Zheng, *American Institutes of Research*

Yuan Zhang, *American Institutes of Research*

Using two nationally representative datasets (Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 [ECLS-K:2011] and NAEP 2015 grade 4 reading assessment data), early reading development trajectories were explored. Results from the study suggest five groups of students with distinct early reading development patterns, highlighting the importance of adequate reading growth in kindergarten and first grade. Additionally, entering kindergarten with mastery of alphabetic principles as well as finishing first grade with mastery of phonological awareness is highly correlated with performing NAEP Basic or above on grade 4 NAEP reading assessments. Findings from this study suggest three important practices to support students' adequate reading growth in early grades and prevent students from falling behind: (a) providing high-quality early literacy instruction prior to kindergarten to help students master basic reading skills, (b) monitoring the development progress of all foundational reading skills such as phonological awareness, even in kindergarten and first grade, and (c) providing focused, evidence-based instructions to students as early as possible.

Session H-3: Data Integration Techniques for Improving Economic Statistics on Multinational Enterprises

Organizer: Kyle Hood, *Bureau of Economic Analysis*

Chair: Ricardo Limes, *Bureau of Economic Analysis*

Discussant: Allison Derrick, *Bureau of Economic Analysis*

Location: Chesapeake C

Multinational Firms in the U.S. Economy: Insights from Newly Integrated Microdata

Jessica McCloskey, *Bureau of Economic Analysis*

Fariha Kamal, *U.S. Census Bureau*

Wei Ouyang, *U.S. Census Bureau*

This paper describes the construction of two confidential crosswalk files enabling a comprehensive identification of multinational firms in the U.S. economy. The effort combines firm-level surveys on direct investment conducted by the U.S. Bureau of Economic Analysis (BEA) and the U.S. Census Bureau's Business Register (BR) spanning the universe of employer businesses from 1997 to 2017. First, the parent crosswalk links BEA firm-level surveys on U.S. direct investment abroad and the BR. Second, the affiliate crosswalk links BEA firm-level surveys on foreign direct investment in the United States and the BR. Using these newly available links, we distinguish between U.S.- and foreign-owned multinational firms and describe their prevalence and economic activities in the national economy, by sector, and by geography.

Linking Data Across International Survey Programs

Jennifer Bruner, *Bureau of Economic Analysis*

Alexis Grimm, *Bureau of Economic Analysis*

The U.S. Bureau of Economic Analysis (BEA) administers an international survey program, which is composed of two areas: 1) trade in services, and 2) foreign direct investment and activities of multinational enterprises (AMNE). The trade in services and direct investment/AMNE program areas are separately maintained, and each has its own system of unique reporter identifiers. To use the data together, a link must be established. The linked data enables greater insight into the operations of a large subset of U.S. international services traders. In 2019, a comprehensive link of 2009-2017 data from these surveys was created. The link greatly expanded on previous efforts to combine the data and automated a large part of the process via programmatic validity checks to evaluate matches on common variables. Research has allowed further refinement of the linking process. For example, the 2023 link update expanded the use of fuzzy matching in the link verification stage. The use of the techniques developed for this internal linking have also been successfully applied to small-scale efforts to link BEA data with external sources, while research is underway to expand use to larger scale links that could enable new layers of data quality cross-checks.

Using Entity Resolution to Improve Inward FDI—QCEW Estimates

Lowell Mason, *Bureau of Labor Statistics*

Statistical agencies increasingly rely on integrating existing data to create new products. We integrate the Bureau of Economic Analysis enterprise-level data on Foreign Direct Investment with establishment data from the Bureau of Labor Statistic's Quarterly Census of Wages and Employment: <https://www.bls.gov/fdi/>. This is challenging because the two data sources lack a consistent common identifier. Previous linking efforts have involved a great deal of manual intervention. In linking 2012 data, a preliminary link using common identifiers yielded an initial error rate of 87.7%, reduced to 19.0% after 1,510.5 hours of analyst labor. We reduce initial linkage error and subsequent labor costs for 2017 data using entity resolution. These methods augment identifiers by making use of other features the data sources share. These features, such as business names and addresses, industrial classification, and employment levels, are similar for true linkages. After indexing the two data sources, we form candidate pairs and use supervised machine learning techniques to

classify whether candidate pairs represent a true linkage or an incorrect linkage based on how similar the common features of the two data sources are. We also implement post-classification clustering methods that adjust the classified candidate pairs to ensure a coherent output.

Improving FDI State-level Employment Statistics Using the QCEW

Ricardo Limes, *Bureau of Economic Analysis*

Ryan Smith, *Bureau of Economic Analysis*

The Bureau of Economic Analysis' (BEA) statistics on new foreign direct investment (FDI) in the United States provide information on the acquisition, establishment, and expansion of U.S. business enterprises by foreign direct investors. These annual statistics provide information on the amount and characteristics of new FDI, including the employment associated with these investments, and are published with country, industry, and state detail. Beginning with statistics for 2018, BEA now uses the Bureau of Labor Statistics' (BLS) Quarterly Census of Employment and Wages (QCEW) establishment-level data (for participating states) to distribute the state level employment of the largest acquisitions of U.S. businesses by foreign investors. This provides a better representation of where the employees of the larger acquired U.S. businesses, which are more likely to operate in multiple states, are employed. The presentation will cover the motivations, processes, and results of this effort through which administrative data were used to improve official economic statistics without increasing public burden.

Thursday, October 26, 2023

8:30 AM

Session H-4: Data Quality: Addressing Disclosure, Deidentification, and the Utility of Publicly Available Data

Organizer: Suzanne Strong, *Bureau of Justice Statistics*

Chair: David Johnson, *National Academies of Science, Engineering, and Medicine*

Location: *Vessey 1*

Leveraging the FCSM Data Quality Framework to Better Communicate Data Quality

Catherine Lamoreaux, *National Center for Health Statistics*

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

In 2020, the FCSM launched the Framework for Data Quality, structured with three domains of data quality: integrity, objectivity, and utility. This framework provides a useful lens to evaluate data quality and can also be leveraged to better communicate data quality, a task complicated by differing types of data, uses, and users. In this presentation, we will follow a case study of high-quality federal data from its release through various interpretations by news outlets and social media users on public platforms. This example comes from an exploratory study examining examples of data quality communication across a variety of web sources. Target data sources were NCHS federal and web-based surveys, intended to provide national information to the public. Examples were selected based on audience reach, engagement, and the discussion of data quality, if present. We studied how the various dimensions of data quality were communicated, identifying where communication was effective and where it fell short. These findings are not generalizable, and users will interpret data quality differently. This case study illustrates the evaluation capabilities of the FCSM Framework for Data Quality in data quality communications.

Data Deidentification Research and Resources from the NIST Collaborative Research Cycle

Christine Task, *Knexus Research*

Gary Howarth, *National Institute of Standards and Technology*

Karan Bhagat, *Knexus Research*

The National Institute of Standards and Technology's Collaborative Research Cycle (CRC) presents results and introduces indispensable tools (data and metrology) for investigating deidentification techniques. Governments collect extensive information from citizens, such as tax information, health data, and many other records. These data are rich and have far reaching utility. Deidentified records, especially those with mathematically provable privacy guarantees, are a promising avenue for governments to release information while protecting privacy. Yet, deidentification techniques that compromise utility pose demonstrable risks. Methods introducing bias, artifacts, and other problematic data may lead to errors in policy making, for example. NIST is confronting these challenges and accelerating research into deidentification techniques through the CRC. The program invited the public to submit deidentified benchmark data (from the American Community Survey), receiving dozens of submissions using diverse techniques. The curated data focus on highly diverse communities with difficult to model characteristics. We evaluate the fidelity of deidentified data using a novel suite of evaluation and visualization techniques drawn from recommendations of staff from national statistic agencies worldwide. We released the deidentified data and evaluation results in a machine-readable 'research acceleration bundle' and are disseminating research performed with these resources.

Refining Disclosure Controls for the Census of Fatal Occupational Injuries

Danny Friel, *Bureau of Labor Statistics*

Alyssa Gillen, *Bureau of Labor Statistics*

Julie Krautter, *Bureau of Labor Statistics*

Yvangelista Saastamoinen, *Bureau of Labor Statistics*

The Census of Fatal Occupational Injuries (CFOI) provides a complete count of workplace fatal injuries in the United States. Selecting a disclosure control method for the CFOI is challenging for several reasons. First, because CFOI is a census, there is no uncertainty when considering the identity of reporting units. Second, many CFOI cell counts are small, so confidentiality rules are frequently violated. Finally, the large number of cells provide many paths to backing out information via table differencing.

The CFOI hypercube is a disclosure control algorithm that applies suppressions to cells defined by all combinations of levels of CFOI variables, across up to four variables at a time. As more variables are considered, the number of cells that must be screened for confidentiality increases exponentially.

The hypercube significantly improved the confidentiality protections of CFOI data: it flags up to 85% of cells for suppression. The current work proposes three refinements to the hypercube that could reconfigure and reduce the number of suppressions without compromising confidentiality.

We demonstrate that up to 90% of cells flagged by the hypercube may be publishable after refinements are made. We discuss lessons learned and identify considerations for other researchers developing custom techniques for disclosure control.

Private Funding of 'Free' Data: A Theoretical Framework

Rachel Soloveichik, *Bureau of Economic Analysis*

This paper develops a theoretical framework in which business buy both data services and other production inputs. Because data is only partially excludable, data sellers who charge high prices are vulnerable to piracy and may not be able to fund the fixed cost of creating data. And even if data can be sold at a high price, social welfare and productivity are still lower than they would be if data was 'free' (Coyle 2022). This paper demonstrates a potential solution to the theoretical framework which both avoids piracy and maximizes social welfare. If data is complimentary to a particular capital asset, then 'free' data can increase the market value of that particular asset by enough that the capital owner benefits from funding 'free' data creation. The paper then argues that 'free' data funded by private capital owners is a large investment category. To start out, the paper presents two empirical case studies of privately funded 'free' data: individual credit files (Soloveichik 2023a) and tax data (Soloveichik 2023b). By themselves, those two case studies account for \$2 trillion of 'free' data creation in 2019. The paper extrapolates total data creation using BEA's published input-output tables. Based on that extrapolation and expert judgment, the paper calculates that total private creation of 'free' data was \$10 trillion in 2017.

Session H-5: Advancing the Federal Statistical Ecosystem TODAY! A 'How-to' Session for Supporting the FSS

Organizer: Jennifer Nielsen, *National Center for Education Statistics*

Chair: Jennifer Nielsen, *National Center for Education Statistics*

Location: *Vessey 2*

FCSM: A Brief Overview of FCSM and Ways You Can Get Involved

Jennifer Nielsen, *National Center for Education Statistics*

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

Ellen Galantucci, *Federal Maritime Commission*

CNSTAT: Helping Improve Federal Statistics, Methods, and Practices for Over 50 Years

Melissa Chiu, *CNSTAT*

The Committee on National Statistics (CNSTAT), established in 1972 at the National Academies of Sciences, Engineering, and Medicine, serves a unique role in federal and national statistics, data, and public policy. CNSTAT's mission is to advise the federal government and the nation grounded in the current best scientific knowledge and practice that will lead to improved statistical methods and information upon which to base public policy. CNSTAT works with federal agencies and experts around the globe on consensus studies and workshops that advance ideas, best practices, and solutions to complex problems. This presentation will feature some of CNSTAT's recent contributions, relevance, and impact on improving statistical methods, data, and evidence and point to CNSTAT's ongoing work on a vision for a new national data infrastructure for the 21st century. The presentation will also discuss CNSTAT's flagship publication, *Principles and Practices for a Federal Statistical Agency* (the "purple book"), which highlights to the executive and legislative branches the important public good provided by strong federal statistical agencies while also providing a code of ethics for agencies that underlies their culture and missions. The presentation will include plans for the upcoming 8th edition and how the federal statistical community can provide input.

COPAFS: Advancing and Promoting Understanding of Federal Statistics and Federal Statistical Agencies

Paul Schroeder, *COPAFS*

The Council of Professional Associations on Federal Statistics (COPAFS) is devoted to educational activities and to 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. COPAFS' goal is to advance excellence in Federal Statistics. During this session COPAFS will discuss their objectives and the activities they conduct to support these objectives.

ASA: Promoting the Practice and Profession of Statistics®

Steve Pierson, *American Statistical Association*

The American Statistical Association was founded in 1839 out of interest in improving the 1840 decennial census. While a much more diverse professional society now—with members throughout the public and private sectors and around the world—we continue to strive to support the federal statistical system and its employees in myriad ways. I will overview these activities and how to participate.

WSS: Live or Work in the Metropolitan Washington, DC Area? Use Statistical or Data Science Methods or Data? If the Answers to These Questions are "Yes," Join the Washington Statistical Society (WSS)

Erin Tanenbaum, *NORC at the University of Chicago*

The Washington Statistical Society (WSS) provides a unique opportunity for employees, contractors, and grant recipients working for or with the Federal Government. Even before its official founding in 1926, prominent Federal statisticians have used the WSS as an opportunity to offer continuing education, networking opportunities, and accolades to further the profession. The WSS continues this tradition by organizing events and volunteer opportunities in and around the Washington, DC area. In this talk, we will outline recent and upcoming WSS happenings, discuss ways in which Federal statisticians are contributing to the WSS, and explore ways statisticians can get involved.

Thursday, October 26, 2023

8:30 AM

Session H-6: What is Consumer Inflation? A Family of Indexes Tells the Story

Organizer: Anya Stockburger, *Bureau of Labor Statistics*

Chair: Thesia Garner, *Bureau of Labor Statistics*

Discussant: Ana Aizcorbe, *Bureau of Economic Analysis*

Location: Room 0105

A History of BLS Research Consumer Price Indexes and the Equivalized Income Quintile Series

Joshua Klick, *Bureau of Labor Statistics*

Anya Stockburger, *Bureau of Labor Statistics*

Since the introduction of the urban population target in the 1970s, the BLS has produced more than the headline measure of consumer inflation. In this presentation, we will review a range of research series such as the R-CPI-RS, R-CPI-E, and R-HICP-CPI-U. The R-CPI-U by equivalized income quintiles is the latest series to be added to the research page and will be updated semiannually. Results through June 2023 are compared to historical periods to evaluate inflation across quintiles.

Disentangling Rent Index Differences: Data, Methods, and Scope

Brian Adams, *Bureau of Labor Statistics*

Lara Lowenstein, *Federal Reserve Bank of Cleveland*

Hugh Montag, *Bureau of Labor Statistics*

Randall Verbrugge, *Federal Reserve Bank of Cleveland*

Prominent rent growth indices often give strikingly different measurements of rent inflation. We create new indices from Bureau of Labor Statistics (BLS) rent microdata using a repeat-rent index methodology and show that this discrepancy is almost entirely explained by differences in rent growth for new tenants relative to the average rent growth for all tenants. Rent inflation for new tenants leads the official BLS rent inflation

by four quarters. As rent is the largest component of the consumer price index, this has implications for our understanding of aggregate inflation dynamics and guiding monetary policy.

Household Cost Indexes: Prototype Methods and Results

Robert Martin, *Bureau of Labor Statistics*

Joshua Klick, *Bureau of Labor Statistics*

William Johnson, *Bureau of Labor Statistics*

Paul Liegey, *Bureau of Labor Statistics*

We propose prototype methods to estimate a family of price indexes known as Household Cost Indexes (HCI) using U.S. data. HCIs aim to measure the average inflation experiences of households as they purchase goods and services for consumption. These differ from the Bureau of Labor Statistics' headline Consumer Price Index (CPI) products in two main respects. First, the upper-level aggregation of the HCIs weights households equally. Second, the HCIs use the payments approach to value owner-occupied housing explicitly using household outlays. In contrast, the U.S. CPIs use rental equivalence. The HCI for all urban consumers has an average 12-month change of 1.52% over December 2011 to December 2021, compared to 1.86% for the CPI-U. The bulk of the difference is due to the payments approach.

Thursday, October 26, 2023

10:30 AM

Session I-1: Data Driven Insights: The Utility and Policy Building of Integrated Data from Federal Statistical Agencies

Organizer: Lisa B Mirel, *National Center for Science and Engineering Statistics*

Chair: Elizabeth Mannshardt, *National Center for Science and Engineering Statistics*

Discussant: Lisa B Mirel, *National Center for Science and Engineering Statistics*

Location: Chesapeake A

Linking the Policy and Utility of Linked Data in Science and Engineering

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

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

The National Center for Science and Engineering Statistics (NCSES) is a leading provider of statistical data on the U.S. science and engineering enterprise. As a principal federal statistical agency, NCSES serves as a clearinghouse for the collection, acquisition, analysis, reporting, and dissemination of statistical data related to the United States and other nations. NCSES conducts more than a dozen surveys in the areas of education of scientists and engineers, R&D funding and expenditures, science and engineering research facilities, and the science and engineering workforce. Type of respondent ranges from individual, academic department and institution, state agency, federal agency, to company, business, and establishment. The Foundations for Evidence-Based Policymaking encourages the secure use of existing federal data to inform critical policy discussions and decisions. In recent years, NCSES has explored linking data survey and alternative data sources to bolster the utility of its surveys. Working with linked data includes implementing policy frameworks that consider potential use cases, development challenges, stakeholder accessibility, and individual privacy. This talk will describe recent NCSES data linkage efforts and discuss its policy-driven research and linkage framework. It will conclude with insights on the way forward in this critical field of data integration.

Does the Decade Matter? Examining the Impact of Using Geocodes from Different Decades in the Analysis of Merged Survey and Contextual Data

Jessie Parker, *National Center for Health Statistics*

Luis Roberto del Pozzo, *National Center for Health Statistics*

Cordell Golden, *National Center for Health Statistics*

The National Center for Health Statistics (NCHS) has a well-established data linkage program designed to enhance the analytic utility of the survey data that NCHS collects by linking survey data with administrative records and geocoding the address data for survey participants. NCHS routinely updates geocodes for participant addresses when the US Census Bureau releases updates following the decennial census. The geocodes allow researchers to merge the survey data with contextual geographic information from other sources, such as the Agency for Healthcare Research and Quality's (AHRQ) Social Determinants of Health (SDOH) Database. This presentation will examine temporal changes at the census tract level from 2010 to 2020 for participants in the 2016 National Health Interview Survey (NHIS) by merging in select information from the AHRQ SDOH database from the following domains: social context, economic context, education, physical infrastructure, and healthcare context. This geographic level information provides critical context to the health information reported by survey participants and adds a longitudinal perspective to the cross-sectional survey data. The presentation will examine how using geocodes from different decennial censuses impact analyses. Examining temporal change in SDOH demonstrates the value of the geocoded survey data for the public health research community.

Using Linked Data to Train and Validate Machine Learning Prediction Models

Orlando Davy, *National Center for Health Statistics*

Cordell Golden, *National Center for Health Statistics*

Linked survey and administrative data can be used to facilitate richer analyses by augmenting the information collected from the surveys with vital or administrative data. However, data linkage requires survey participants to provide consent for linkage and sufficient personally identifiable information (PII). There has been a growing reluctance of survey participants to provide the PII needed for linkage. When data linkage is not possible, machine learning (ML) prediction models can be used to predict outcomes, such as morbidity and mortality, for survey participants. To perform as intended, the models require quality and accurate training data and a validation source. The National Center for Health Statistics (NCHS) Data Linkage Program has produced an extensive repository of high-quality linked data files that can be used to address a wide-range of health-related research topics and a variety data science applications. This presentation will explore the use of NCHS linked data for training and validating ML prediction models. Data from NCHS' National Health Interview Survey linked with the National Death Index will be used to investigate the performance of three machine learning models used for the prediction of all-cause mortality.

Thursday, October 26, 2023

10:30 AM

Session I-2: An Updated Measure of Poverty: (Re)Drawing the Line; Recommendations and Their Implementation

Organizer: David Johnson, *National Academies of Sciences, Engineering, and Medicine*

Chair: David Johnson, *National Academies of Sciences, Engineering, and Medicine*

Discussants: Chris Mackie, *National Academies of Sciences, Engineering, and Medicine*, and Michelle Ver Ploeg, *Economic Research Service*

Location: Chesapeake B

Recommendations and Conclusions from the CNSTAT Report

Jim Ziliak, *University of Kentucky*

The recent CNSTAT report, *An Updated Measure of Poverty: (Re)Drawing the Line*, highlights the critical importance of poverty measurement for the evaluation of the nation’s well-being and the evaluation of the effectiveness of government programs. The report recommends that the SPM be elevated to the nation’s headline poverty statistic and renamed accordingly (e.g., to the Principal Poverty Measure [PPM]). In proposing the PPM, the report reconsiders the SPM approach to both threshold setting and resource estimation. In its specification of the PPM, the set of threshold categories has been expanded to explicitly recognize additional expenditures required of most American families. Specifically, the report argues for including medical care and, with further research, childcare in the basic needs bundle, and for updating how basic housing costs are estimated. The fact that medical care and childcare are broadly recognized as basic needs—needs that typically account for a bigger budget share than SPM-included categories such as clothing—means that the threshold component of the PPM comes closer to representing the full resource level needed to keep a family out of poverty. This creates a transparency advantage for the PPM over the SPM, which hides some of these costs by subtracting them from resources. This presentation will highlight the importance recommendations and conclusions in the CNSTAT report.

Updating the SPM: Plans, Schedule and Using Households

Liana Fox, *U.S. Census Bureau*

Currently, updates to the SPM are governed by an Interagency Technical Working Group. With the release of the new CNSTAT report, there are many new revisions that need to be reviewed, and possibly implemented. As the CNSTAT report pointed out in Recommendation 2.7, the SPM should consider using the household as the unit of analysis, as many households share housing and other costs. This research examines SPM poverty rates using alternative units of analysis as compared to the current SPM units. The presentation will also provide comments on some of the other recommendations in the report and highlight the process and timelines discussed in the ITWG.

Health Inclusive Poverty Measure Estimates in the United States: 2022

John Creamer, *U.S. Census Bureau*

In 2023, the National Academies’ Committee on National Statistics released a consensus report offering recommendations to update the Supplemental Poverty Measure (SPM). One of the recommendations is to change the way medical expenses are handled in the poverty measure, moving from deducting these expenses from a household’s resources to incorporating a value of health insurance needs and resources using the Health Inclusive Poverty Measure (HIPM) framework (Korenman and Remler 2016). This research will explore the differences between the SPM and HIPM treatment of medical expenses by comparing poverty rates between the two measures from 2014 to 2022. The research will also estimate the impact of public assistance programs on HIPM poverty rates, a novel feature of the HIPM framework. Altogether, the estimates are a proof-of-concept of how a new poverty measure produced by the Census Bureau would incorporate values of health insurance in poverty measurement.

Accounting for Rents and Net Implicit Rental Income in an SPM-like Measure through the Use of Rental Equivalence for Owner-occupied Housing

Thesia I. Garner, *Bureau of Labor Statistics*

Jake Schild, *Bureau of Labor Statistics*

Currently differences in shelter expenditures for owners and renters are accounted for in the Supplemental Poverty Measure (SPM) by the use of separate thresholds for homeowners with a mortgage, homeowners without a mortgage, and renters. The Committee on National Statistics (CNSTAT) SPM panel recommended discontinuing this approach and instead noted that these differences be accounted for in resources (recommendation 5.3). The purpose of this research is to examine the potential impact of adopting this recommendation; other CNSTAT panel’s recommendations will not be considered. A rental equivalence based SPM-like measure will be produced for the year 2021 using three years of data from the U.S. Consumer Expenditure Survey. The threshold will include rental equivalence for owners, as opposed to out-of-pocket expenditures, in the threshold; resources will include net implicit rental income from homeownership. The estimation sample will be restricted to single consumer unit households only. Expenditures will be

equivalized using the same 3-parameter equivalence scale in use to produce the current SPM measure. A reference unit threshold will be produced for consumer units/households composed of two adults and two children. Poverty rates based on this measure will be produced for the total population and for select demographic characteristics.

Thursday, October 26, 2023

10:30 AM

Session I-3: A Roadmap for Disclosure Avoidance in the Survey of Income and Program Participation

Organizer: Brad Chaney, *National Academies of Sciences, Engineering, and Medicine*

Chair: Brad Chaney, *National Academies of Sciences, Engineering, and Medicine*

Discussant: Caleb R. Floyd, *U.S. Census Bureau*

Location: Chesapeake C

Assessing the Level of Disclosure Risk in SIPP

V. Joseph Hotz, *Duke University*

The SIPP panel examined the current level of disclosure risk faced by SIPP in two ways. One part of the analysis involved looking at the SIPP data and how risk is affected by the data granularity and by a structure that includes both household-level and longitudinal data. A second part involved reviewing a re-identification study performed by the Census Bureau, looking at its strengths and weaknesses. The panel offers recommendations on how disclosure risk assessments might be improved. The findings have important implications on what can be released and how data are made available.

Promoting Usability and Accessibility in SIPP While Respecting Respondent Confidentiality

Jennifer Lynne Van Hook, *The Pennsylvania State University*

There often is a tradeoff between promoting data usability and protecting confidentiality. The panel considered three different dimensions of usability —(1) accuracy of the statistics generated, (2) the feasibility in generating statistics that answer significant questions that the SIPP was designed to address, and (3) ease and equity of access. It then examined how to create an appropriate balance between usability and privacy protection for SIPP, while considering how SIPP is currently used by researchers (in terms of both the content and methodology), what resources are available to researchers, and how different types of modes of access affect usability.

Virtual Data Enclaves and Secure Data Access

Heeju Sohn, *Emory University*

Federal agencies and others often provide confidential data through secure virtual data enclaves. The panel evaluated current access to the restricted SIPP through Federal Statistical Research Data Centers (FSRDC) and examined whether alternative forms of secure online data access may be appropriate for SIPP. The panel considered multiple models of secure online access and discussed their implications for protecting confidentiality, burdens on researchers to access the data, and implications for equity in disclosure risk for minority groups. The panel also examined potential strategies for the Census Bureau to consider for creating and maintaining an alternative secure online data access system.

The Potential of a Remote Analysis Platform as a Tool for Protecting Confidentiality

Thomas Richard Krenzke, *Westat*

Natalie Shlomo, *University of Manchester*

One potential tool for promoting data access to SIPP is by creating a remote analysis platform that includes a table generator for creating customized tabulations and other types of analyses. The remote analysis platform would allow users to define and generate their own statistics via a query-based system using drop-down lists without actually seeing the data. These platforms are intended to meet increasing demands for more open and accessible data and to utilize the power of the internet to disseminate tabular outputs and other statistical analyses through flexible web-based applications. They also could have the potential to expand the usage of SIPP by widening access to the public. We examine how disclosure limitation rules-of-thumb can be built into the platform to avoid identity and attribute disclosures based on a set of criteria and thresholds, and how the criteria and thresholds may be assessed automatically. We also examine whether disclosure limitation methods can be applied on-the-fly to protect outputs obtained from the remote analysis platform and whether the disclosure limitation methodology can be transformed into a non-interactive differential privacy mechanism. The panel examined the potential capabilities of such a platform, what types of users would make use of such a platform, how it would be used and what would be needed to support current users. We conclude with a timeline for developing a complete platform for SIPP.

Geography and Small Area Estimation in SIPP

Scott Holan, *University of Missouri*

Privacy protection in the context of providing small area estimates constitutes an important problem in official statistics. Although there is a vast literature on basic approaches to small area estimation, it is often advantageous to consider the specific data product and needs of the constituent data users in route to production and dissemination of small area estimates. At the same time, it is necessary to simultaneously impose some form of disclosure avoidance methodology. This talk describes general SAE methodology and focuses on avenues forward to producing privacy preserving SAE for the U.S. Census Bureau's Survey of Income and Program Participation (SIPP). Emphasis on geographic aspects of the SIPP will be highlighted in this presentation.

Thursday, October 26, 2023

10:30 AM

Session I-4: Building on GDP: The Future of Economic Statistics

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

Chair: Vipin Arora, *Bureau of Economic Analysis*

Location: Vessey 1

Global Supply Chains: New Resources and the Path Forward

Tom Howells, *Bureau of Economic Analysis*

Policymakers and data users have long sought new ways of measuring and understanding global supply chains and how goods and services flow through the global economy. The COVID-19 pandemic and its aftermath have only increased the desire for these statistics. Building on a multiyear collaboration with the National Science Foundation and deploying new resources, BEA will expand on its existing Trade in Value Added (TiVA) statistics and build out new and existing products to provide a more comprehensive picture of this important aspect of the world economy.

BEA's Measures of Economic Well-Being

Alyssa Holdren, *Bureau of Economic Analysis*

BEA regularly seeks ways to further measure economic well-being using both traditional and new statistics. One product BEA has produced for several years are prototype Measures of Economic Well-Being, which combine statistics from its core accounts with data from statistical agency partners in ways that provide new perspectives on economic well-being and the distribution and long-term growth of the economy. Advances in automation and visualization will allow BEA to build out these prototype measures, add new features, and develop open-source code for data users to further build upon the foundation laid by BEA.

Technical Advances through Environmental-Economic Accounting to Expand Statistical Measures

Charles Rhodes, *Office of Management and Budget*

As the U.S. begins to develop its 15-year program to build a System of Environmental-Economic Accounts, as laid out in January 2023's National Strategy to Develop Statistics for Environmental-Economic Decisions, technical progress needs to occur on multiple fronts. Official U.S. Environmental-Economic Accounting Classifications will be necessary to formalize accounts as we develop from pilot to prototype to production-grade accounts, and the Office of the Chief Statistician of the United States and BEA are in Classifications development discussions. As with other U.S. economic account production, data sharing across agencies will be a concern, and demand for data sharing could be high for these new accounts. Valuation of Environmental-Economic physical measures will continue to present issues that have persisted in environmental economics for decades, and not all of the issues are purely technical. Beyond new interagency coordination to expand the U.S. economic statistics system to include these accounts, formal and informal collaboration internationally to exchange information, techniques, and possibly protocols presents further challenges and opportunities.

Expanding the Frontier of Economic Statistics Using Alternative Data: A Case Study of Regional Employment

Abe Dunn, *Bureau of Economic Analysis*

Big data offers potentially enormous benefits for improving economic measurement, but it also presents challenges (e.g., lack of representativeness and instability), implying that their value is not always clear. We propose a framework for quantifying the usefulness of these data sources for specific applications, relative to existing official sources. We specifically weigh the potential benefits of additional granularity and timeliness, while we carefully consider the error of any new or improved estimates, relative to comparable errors produced in existing official statistics. We apply the methodology to employment estimates using data from a payroll processor, considering both the improvement of existing state-level estimates, but also the production of new, more timely, county-level estimates. We find that incorporating payroll data can improve existing state-level estimates and produce county-level estimates with reasonable levels of error, given the level of granularity.

Thursday, October 26, 2023

10:30 AM

Session I-5: Enhancing Survey Programs by Using Multiple Data Sources

Organizer: Brian Harris-Kojetin, *National Academies of Sciences, Engineering, and Medicine*

Chair: Melissa Chiu, *National Academies of Sciences, Engineering, and Medicine*

Discussant: Joseph Salvo, *University of Virginia Biocomplexity Institute*

Location: Vessey 2

Overview of Toward a 21st Century National Data Infrastructure: Enhancing Survey Program by Using Multiple Data Sources

Elizabeth Stuart, *Johns Hopkins University*

Much of the statistical information produced by federal statistical agencies has come from sample surveys. However, surveys have faced a number of challenges in recent years, including decreasing response rates, increasing costs, and demands for more timely and more granular data. There has been a proliferation of data from other sources, including administrative records, satellite and sensor data, and private-sector data that could be harnessed to enhance survey data. The National Academies' Committee on National Statistics recently issued a report on Toward a 21st Century National Data Infrastructure: Enhancing Survey Programs by Using Multiple Data Sources, which identifies four main ways that multiple data sources could improve

national statistics. These include 1) providing information to evaluate, improve quality of data sources; 2) giving additional information about survey respondents; 3) producing statistics for small populations; and 4) creating data products directly from administrative data. The report includes a small set of “use cases” from the areas of income, health, crime, and agricultural statistics that represent different ways in which multiple data sources are, or could be, exploited and that illustrate the types of challenges to be faced.

Using Multiple Data Sources to Enhance Data Equity

Kimberlyn Leary, *Harvard Medical School*

The use of multiple data sources can benefit data equity—promoting the collection and use of data in which all populations, and especially those that have been historically underrepresented or misrepresented in the data record, are visible and accurately portrayed. This presentation will discuss different aspects of data equity and examine how multiple data sources can be used to investigate or improve coverage, enable finer data disaggregation, produce model-based estimates for small subpopulations, and assess and reduce measurement error.

Measuring and Mitigating Racial Disparities in Tax Audits

Tom Hertz, *Internal Revenue Service*

Government agencies around the world use data-driven algorithms to allocate enforcement resources. Even when such algorithms are formally neutral with respect to protected characteristics like race, there is widespread concern that they can disproportionately burden vulnerable groups. We study differences in Internal Revenue Service (IRS) audit rates between Black and non-Black taxpayers. Because neither we nor the IRS observe taxpayer race, we propose and employ a novel partial identification strategy to estimate these differences. Despite race-blind audit selection, we find that Black taxpayers are audited at 2.9 to 4.7 times the rate of non-Black taxpayers. The main source of the disparity is differing audit rates by race among taxpayers claiming the Earned Income Tax Credit (EITC). Using counterfactual audit selection models for EITC claimants, we find that maximizing the detection of underreported taxes would not lead to Black taxpayers being audited at higher rates. In contrast, in these models, certain policies tend to increase the audit rate of Black taxpayers: (1) designing audit selection algorithms to minimize the “no-change rate”; (2) targeting erroneously claimed refundable credits rather than total under-reporting; and (3) limiting the share of more complex EITC returns that can be selected for audit. Our results highlight how seemingly technocratic choices about algorithmic design can embed important policy values and trade-offs.

Thursday, October 26, 2023

10:30 AM

Session I-6: Innovations in Gathering Health Data: Pilots, Methods, and Lessons

Organizer: Rich Levy, *U.S. Census Bureau*

Chair: Rich Levy, *U.S. Census Bureau*

Location: Room 0105

The 2021 Physician Pain Management Questionnaire Pilot Study: Lessons Learned and Future Implications

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

Doreen Gidali, *National Center for Health Statistics*

The 2021 Physician Pain Management Questionnaire (PPMQ) was a pilot study conducted by the National Center for Health Statistics (NCHS) and sponsored by the National Institutes of Health’s National Center for Complementary and Integrative Health. It was conducted as a mail survey sent to 1,000 physicians between June 2021 and February 2022. The sample consisted of nonfederal-employed physicians providing care to

patients in all 50 U.S. states and the District of Columbia. The PPMQ pilot study objectives were to: evaluate survey questions on physician knowledge, awareness, and use of prescription opioid guidelines, and assess the feasibility of conducting a nationally representative physician survey to gain better understanding of physician opioid prescribing practices for pain management. This presentation will discuss findings from the cognitive evaluation of survey questions on pain management and guidelines for prescribing opioids; outline modifications to the PPMQ based on the evaluation to better capture these concepts; highlight results from statistical analyses supporting the questionnaire's efficacy in capturing data on intended concepts; and discuss lessons learned from conducting the PPMQ pilot study as well as its subsequent impact on future health care surveys.

Design of a Pilot Test of Businesses to Study the Effects of Omitting Select Survey Processes on Survey Response and Data Quality

David Kashihara, *Agency for Healthcare Research and Quality*

Matthew Thompson, *U.S. Census Bureau*

The Covid-19 health pandemic caused surveys to experience delays and work stoppages as operations were shut down and either postponed or cancelled. In 2021, with both future shutdowns and escalating survey costs considered, the Medical Expenditure Panel Survey – Insurance Component (MEPS-IC), a large federal business survey covering health insurance offerings, fielded a pilot test to study the effects of conducting the survey without some of the more costly and labor-intensive processes: Research, Prescreener, and Telephone Follow-Up (TFU). During Research, analysts use various resources to update contact information for businesses. The Prescreener is a telephone screening call to determine if establishments offer insurance. If insurance is offered, the establishment will receive the full survey. Otherwise, questions will be asked about the business and the survey will end. The TFU operation attempts to reach businesses that did not respond to the full survey via web or mail. The pilot test of about 5,000 small businesses was drawn independently from the annual MEPS-IC. Data were collected concurrently with MEPS-IC cases but processed separately so they would not contribute to the official survey estimates. This presentation will discuss the sampling design, data collection and processing, and estimation for the pilot test. The presented results of the pilot test will inform the MEPS-IC and other surveys of the potential effects of omitting common survey processes (not necessarily voluntarily) despite possible gains of reduced survey costs.

Can High-quality Health Tracking Data be Collected Using a Probability Panel Sample? Results From a National Bridge Study

Jazmyne Sutton, *SSRS*

Ashley Kirzinger, *Kaiser Family Foundation*

Cameron McPhee, *SSRS*

As the rising cost of survey research continues to drive many studies away from the use of traditional probability samples toward less resource-intensive options, there is still valid hesitation on the part of many large-scale tracking studies to move to alternative sample frames, notably probability-based online panels. Concerns about the representativeness of the online probability panels and the impact of respondent conditioning on data quality and the reliability of measured trends has contributed to this wariness. However, in recent years, several notable organizations have begun to “test the probability-panel waters” in an effort to sustain these important sources of data. In 2022, KFF worked with SSRS to conduct two concurrent versions of their Health Tracking Survey to assess the similarities and notable differences between surveys of U.S. adults conducted using a probability-based online panel and an Address-based Sample, with a push-to-web methodology, comparing both projects to long-term trend questions on health care experiences and attitudes towards health care policy. This presentation details the methodology and results of this bridge study, highlighting both demographic, behavioral, and attitudinal differences between the two sets of respondents. We also discuss whether standard weighting techniques mitigate any of the differences and explore whether the addition of a supplementary random sample of prepaid cell phones can offer a way for the most rigorous U.S. health studies to transition to a less costly methodology.

Measuring the Reliability of Method Comparison Study Methodology Using Data from the National Health and Nutrition Examination Survey

Jennifer Rammon, *National Center for Health Statistics*

Kevin Chuang, *National Center for Health Statistics*

Te-Ching Chen, *National Center for Health Statistics*

Hee-Choon Shin, *National Center for Health Statistics*

The Division of Health and Nutrition Examination Surveys at the National Center for Health Statistics conducts internal method comparison studies whenever laboratories undergo instrumental or methodological measurement changes or when contract laboratories change. Difference plots and Deming regressions guide decision making and if substantial differences are observed, then adjustment equations are released with the data for analysts planning to combine survey cycles or conduct trends analyses. This project aims to assess the reliability of the methodology by analyzing blinded pseudo-crossover-studies created from publicly available National Health and Nutrition Examination Survey data. Specifically, we address three questions: are adjustment equations overproduced, is it sufficient to rely on single measurements by each method versus duplicate measurements, and how are overall trend estimates influenced by the inclusion of adjustment equations? Preliminary results suggest that adjustment equations are not overproduced, that duplicate measurements are less biased than single measurements, and that effects on trends analyses are minimal.

Lessons Learned from the National Electronic Health Records Survey

Kelly L. Myrick, PhD, CPH, *National Center for Health Statistics*

Mohsin Mahar, MPH, *Reli Group Inc.*

The National Electronic Health Records Survey (NEHRS) has provided valuable information on the use and adoption of electronic health records and health information exchange among American office-based physicians for over a decade. Over the years, NEHRS has had to adapt to changes in office-based care and respondent preferences. This presentation will discuss the lessons we have learned about changes in response by mode, who responded, and how they responded over time. In the most recent years of NEHRS, response from the sampled physician ranged from 68.8% (standard error [SE] 3.4 percentage points) in 2018 to 82.1% (SE 1.5 percentage points) in 2021. NEHRS is conducted by the National Center for Health Statistics and sponsored by the Office of the National Coordinator for Health Information Technology. NEHRS began as an annual mail survey known as the National Ambulatory Medical Care Survey Electronic Medical Records Supplement (EMR survey 2008–2011). In 2012, the survey became NEHRS, which was fielded annually from 2012–2021 with the exception of 2016 and 2020.

Thursday, October 26, 2023

1:45 PM

Session J-1: Enhancing Food Policy Research Through Administrative Data Linkages

Organizer: Matthew P. Rabbitt, *Economic Research Service*

Chair: Mark Prell, *Economic Research Service*

Discussant: Constance Newmann, *Food and Nutrition Service*

Location: Chesapeake A

SNAP Participation Dynamics in A Long-Term Administrative Data Panel

Laura Tiehen, *Economic Research Service*

Kegan O'Connor, *Economic Research Service*

Christian A. Gregory, *Economic Research Service*

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

Mark Prell, *Economic Research Service*
Michelle Ver Ploeg, *Economic Research Service*

This study examines the dynamics of SNAP participation, using monthly program administrative data from selected states in different regions of the U.S. We analyze SNAP participants' spell duration patterns, including patterns of entry, exit and re-entry into SNAP over a sample period that extends, in some States, from 2004 to 2021. We use Kaplan-Meier estimates of the hazard and survival functions for spells of SNAP participation. This descriptive analysis allows us to examine the distribution of participation spell durations in a long panel that contains a relatively large number of uncensored participation spells (where we can observe both entry into and exit out of SNAP). The analysis provides further and updated evidence that SNAP serves as temporary assistance for some participants and longer-term support for others. The length of the panel also allows us to examine how participation patterns have changed over a period that spans the Great Recession and slow economic recovery, as well as the COVID-19 pandemic. Additionally, we examine how changing patterns of entry and exit have contributed to SNAP caseload changes.

Developing a Longitudinal Database to Facilitate Research on SNAP Participation

Renuka Bhaskar, *U.S. Census Bureau*
Maria Perez-Patron, *U.S. Census Bureau*
Brad Foster, *U.S. Census Bureau*
Brian Knop, *U.S. Census Bureau*

The Supplemental Nutrition Assistance Program (SNAP) is the nation's largest federal effort to reduce hunger, providing benefits to 41 million people in 22 million households in fiscal year 2021. Cross-sectional estimates of SNAP participation do not provide an understanding of patterns of program participation in terms of long-term versus temporary periods of access. The SNAP Longitudinal Data Project (SNAP-LDP) was set forth by the 2018 Farm Bill with the objective to create a longitudinal database to facilitate research on SNAP participation, including duration of participation over time and across states. The SNAP-LDP involves the joint effort of the USDA-Food and Nutrition Service, the U.S. Census Bureau, and state SNAP agencies that will share their data with the U.S. Census Bureau. This presentation will describe the methodology involved in the development of the SNAP-LDP database, going from data acquisition, assignment of unique identifiers for data protection and data linkage, and database development. We will present some preliminary results of the type of estimates that researchers will be able to obtain using this database including entry and exit rates, spell duration, gaps in coverage, and changes in caseloads over time.

Estimating the Prevalence of Food Security for HUD Subsidized Housing Units through Data Linkage

Veronica E. Helms, *Department of Housing and Urban Development*
Emily Molfino, *U.S. Census Bureau*

Prior cross-agency research using administrative data linked with the National Health Interview Survey (NHIS) showed that persons receiving federal housing assistance from the U.S. Department of Housing and Urban Development (HUD) faced alarmingly high rates of food insecurity; approximately 37.2% of HUD-assisted adults reported household food insecurity and rates varied significantly by HUD program type. Results suggest that individuals in residential choice assistance programs had higher rates of food insecurity when compared to individuals in project-based programs, who are assigned specific housing units. As part of the Biden-Harris Administration National Strategy on Hunger, Nutrition, and Health, HUD's Office of Policy Development and Research (PD&R), in collaboration with the U.S. Census Bureau and the U.S. Department of Agriculture (USDA), Economic Research Service (ERS) have committed: "to better understand and address persistent food insecurity among HUD-assisted individuals." To achieve this goal, the research team linked the 2015-2020 Current Population Survey: Food Security Supplement with HUD administrative data on individuals living in public and assisted housing. Preliminary findings underscore the uniqueness of HUD-assisted households experiencing food insecurity and include key findings related to HUD program type, household structure, main source of income, housing costs, and food costs.

Addressing Nonresponse Bias in Food Security Measures Using Weighting Adjustments

Jonathan Eggleston, *U.S. Census Bureau*
Matthew P. Rabbitt, *Economic Research Service*
David C. Ribar, *Georgia State University*
Alisha Coleman-Jensen, *Economic Research Service*

Food security—defined as consistent, dependable access to enough food for an active, healthy living—has been measured by the United States Department of Agriculture using the Current Population Survey Food Security Supplement (CPS-FSS) since 1995. Response rates to the CPS-FSS have declined in recent years, dropping below the 80 percent threshold for federal surveys set by the Office of Management and Budget. As of 2020, 75.8 percent of monthly CPS households completed the Food Security Supplement interview. As the response rate for the CPS-FSS declines there is potential for nonresponse bias in food security measures which can impact the monitoring of food security in the United States. In this study, we will examine nonresponse bias in food security measures from the CPS-FSS using a survey weighting adjustment procedure proposed by Eggleston et al. (2022). This procedure relies on federal and third-party administrative data matched to CPS-FSS respondents and nonrespondents to address survey nonresponse through the construction of new weighting variables. After constructing the new weighting variables, we compare food security prevalence rates using both the new and traditional supplement weights.

Thursday, October 26, 2023

1:45 PM

Session J-2: Putting Statistics on the Map, Using Spatial Analysis and Geography to Reveal the Story

Organizer: Ed Strocko, *Bureau of Transportation Statistics*
Chair: Ed Strocko, *Bureau of Transportation Statistics*

Location: Chesapeake B

Tax Filing Literacy Gaps in the Small Business and Enterprising Gig Ecosystem: Exploring Spatial Socio-economic and Demographic Factors

Barbara J. Robles, *Retired, Federal Reserve Board*
Caroline Bruckner, *American University*

We analyze new survey data of gig ecosystem respondents managing 1099 income and taxes. Working with our research partners, we administered an online survey to approximately 89,000 small business, gig and freelance work-for-pay respondents in September 2022. We capture key demographic information of the tax literacy bandwidth of respondents. We explore linked spatial contributions that impact enterprising gig workers filing taxes with 1099 income. We estimate logit models that investigate spatial characteristics displayed by respondents along with their tax knowledge by linking our survey (n=562) to the IRS Individual zip-code public data. Our analysis probes:

- (i) What are the socio-economic demographics of those generating enterprising and gig income?
- (ii) What are 'spatial' tax literacy gaps of enterprising income generation?
- (iii) What industry sectors are impacted by the growing gig ecosystem and the tax knowledge gap?

The results provide insight into linked spatial zip-code information contributing to tax compliance and tax literacy gaps.

Reimagining Redistricting Data Products

Emily Molfino, *U.S. Census Bureau*
Sallie Keller, *U.S. Census Bureau*

Redistricting happens every 10 years after the U.S. Census Bureau collects new population data via the

decennial census. That population data informs how states redraw boundaries to reflect shifts in population between the district being represented. Since the 1980 Census, the U.S. Census Bureau has been legally required to provide the redistricting data file (Public Law 94-171) to help inform these efforts. Today, the redistricting data are used in a growing range of non-redistricting research and policy planning. There is interest in examining the redistricting data to see how they can be reimagined in the context of current formal privacy and confidentiality protections while still meeting data user needs. The presentation will discuss research focused on the complex balance between ensuring the redistricting data meet redistricting needs, protecting privacy and confidentiality, and supporting the other uses of the redistricting data.

Discovering Hidden Patterns in County-Level Diagnosed Diabetes Incidence in The United States Using Neural Networks: A Spatio-Temporal Analysis From 2011 to 2020

Hui Xie, *Centers for Disease Control and Prevention*

Deborah B. Rolka, *Centers for Disease Control and Prevention*

Despite a slowdown in United States diabetes incidence rates over the past decade, trends in the incidence of newly diagnosed diabetes are highly variable at lower geographic levels. This study aims to unveil complex patterns and relationships in diabetes incidence at the county level using a neural network approach. To estimate county-level diagnosed diabetes incidence rates, we applied small area estimation (SAE) techniques to ten years (2011-2020) of data from the Behavioral Risk Factor Surveillance System (BRFSS). We then designed a recurrent neural network (RNN) architecture that could capture the complex patterns and relationships in the estimates. The RNN model was trained and evaluated using a combination of cross-validation and performance metrics such as accuracy, precision, and recall. Our analysis revealed several key findings. First, we identified distinct temporal patterns in the incidence rates, with 12.41% of counties having a continuously increasing trend, while others showed more complex patterns. Second, we observed notable spatial variations, with 22.89% of counties (mostly located in the Southeast) having consistently higher incidence rates than others. Our study demonstrates the potential of using a neural network approach to identify hidden patterns and relationships at the county level. The findings provide valuable insights into the temporal and spatial dynamics of diabetes incidence and could inform targeted interventions to help reduce diabetes incidence rates in the most affected areas.

Accurately Identifying and Enumerating Multiunit Housing with Remote Sensing Data for Address Frame Enhancement

Lee Fiorio, *NORC at the University of Chicago*

Ned English, *NORC at the University of Chicago*

Recent developments have minimized the need for field-based housing unit listing when constructing frames for in-person household surveys. For most places in the United States, the USPS Delivery Sequence File (CDS) provides a near-complete list of addresses. In places where CDS coverage is insufficient, advancements by NORC have it possible to improve coverage by integrating the CDS with spatial data including satellite imagery, building footprint data, and county assessor tax parcel data – that is, list these areas remotely without having to visit them in the field. However, an important challenge when remote listing is identifying and enumerating multiunit housing. This paper will offer an assessment of NORC's approach to multiunit housing used while remote listing a portion of NORC's 2020 Decennial National Master Sample. In a sample of areas, we conduct field validation to evaluate the accuracy with which we identified multiunit housing during our remote listing procedure. Results will define the scope of the multiunit housing problem and highlight the contexts in which remote listing performs well.

Geospatial Heterogeneity in Inflation: A Market Concentration Story

Michael Navarrete, *University of Maryland*

Seula Kim, *Princeton University*

Inflation is an important economic indicator that is usually measured and reported at the national level. However, there is systematic geographical variation across metropolitan statistical areas (MSAs). Using

Nielsen retail scanner dataset, we find that the poorest MSAs experience higher inflation rates than the richest MSAs from 2006 to 2016 for PCE food and beverages category. The official PCE inflation measure closely matches the inflation experienced by the richest decile, but understates the inflation experienced in most MSAs. If we cumulate the differences in inflation rates from 2006 to 2016, the poorest decile experiences about 10 percentage points higher inflation than the richest decile. We can look at disaggregated PCE categories such as milk and eggs using a BLS concordance that maps Nielsen product modules to PCE entry level items (ELIs). ELIs then map to PCE categories. We also find an association between market concentration and inflation. The richest MSAs have lower market concentration and experience less inflation. We look at exogenous supply shock, the bird flu of 2015, in the eggs market to tease out a causal relationship between market concentration and inflation.

Thursday, October 26, 2023

1:45 PM

Session J-3: Improving Price Index Measurement: Methods for Incorporating New or Alternative Data Sources

Organizer: Kyle Hood, *Bureau of Economic Analysis*

Chair: Jeffrey Gonzalez, *Bureau of Labor Statistics*

Location: Chesapeake C

A Bootstrap Variance Method to Evaluate Blended Import and Export Price Indexes with Census Trade Data

Daniel Yang, *Bureau of Labor Statistics*

Helen McCulley, *Bureau of Labor Statistics*

Xiao Fan, *Bureau of Labor Statistics*

Rozi Ulics, *Bureau of Labor Statistics*

The International Price Program (IPP) at the Bureau of Labor Statistics (BLS) produces Import and Export Price Indexes (MXPI) which consists of two components: import and export merchandise. MXPI is the primary national index that monitors the U.S. establishments' international trade prices for goods and services. The IPP is planning to replace a sizeable proportion of the MXPI survey with Census Trade Data (CTD) or administrative trade transaction data. CTD collects detailed shipment transaction records, e.g., international Harmonized System (HS) product classification code, trade value (\$), foreign country of import or export, shipment quantity, etc. A well-designed index-specific Match-Adjusted R-squared (MARS) method was established to determine which complex characteristics define a product variety and produce consistent prices in the transaction data (McCulley 2023), and which are used to produce blended (MXPI survey and CTD) price index estimates. In this study, we adopted for each product variety defined by MARS' statistics, we calculate each index strata and propose a bootstrap variance estimation method to take into account the combined sources of error from a survey and error propagation from a census.

A Match Adjusted R-Squared Method for Defining Products within Census Administrative Trade Data

Helen McCulley, *Bureau of Labor Statistics*

The Bureau of Labor Statistics (BLS) Import and Export Price Indexes (MXPI) measure price changes of U.S. imports and exports by surveying U.S. establishments to collect prices for goods and services they trade internationally. The BLS is planning to utilize administrative data in place of directly collected data in the MXPI for a large portion of measured trade. The administrative trade transaction data contains detailed shipment records that cover nearly all international merchandise trade by the United States and is currently used as the sampling frame for the MXPI. The data is collected by the U.S. Customs and Border Protection agency and maintained by the U.S. Census Bureau. Each shipment transaction record contains the international Harmonized System (HS) product classification code, trade dollar value, foreign country of

import or export, shipment quantity, trade party identifiers, and several other potentially price-determining data fields that can be used to define homogeneous products. Increasing the number of data fields used to define a product helps reduce index unit value bias. Too many defining characteristics can decrease the number of months a product is available. To balance these competing index calculation needs, a match-adjusted R-squared method (Chessa, 2021) was implemented to define product varieties.

Preliminary Market Basket Weights for Chained CPI-U

Kate Eckerle, *Bureau of Labor Statistics*

A primary objective of a consumer price index is to accurately capture changes in consumer behavior. In contrast to its traditional counterpart, the BLS' Consumer Price Index (CPI-U), which uses annualized weights, the Chained CPI-U (C-CPI-U), with its more dynamical basket, was explicitly designed to account for adjustments consumers make to distribute their overall spending as the relative prices of goods and services change, and as their needs change in general. Research suggests that alternatives to price index calculations that swap the stationary reference basket for forecasted monthly weights derived using filter-based ARIMA models and historic timeseries of monthly budget shares, improves the accuracy of the C-CPI-U prediction. Forecasting monthly weights may be superior, and consistent with the assumptions of the chained index. However, predictions may be limited to only accounting for long-term trends, and strong seasonal patterns in consumer behavior; dynamics that are in some sense independent from the state of prices. These methods may achieve a better baseline expectation for the course of expenditures shares over a given year. We investigate whether 1. frequency domain methods, particularly those which incorporate interference patterns among elementary items, further improves the baseline monthly weights forecast, 2. whether special treatment of outliers provides improvements, and 3. whether incorporating auxiliary non-CE data that is more up to date improves the accuracy of the preliminary C-CPI-U index, particularly during anomalous periods.

Variances of Combined Consumer Price Index Survey and Alternative data

Onimissi Sheidu, *Bureau of Labor Statistics*

The Consumer Price Index (CPI) is generally estimated using probability survey data. In recent years, estimates for some subsector indexes have become volatile due to low response rates. This issue poses a serious challenge in validating the unbiasedness of these estimates. This also leads to an increase in data collection costs needed to obtain large sample sizes for robust estimation. The abovementioned problems have made the use of alternative data more enticing to mitigate these issues. However, the problem of quantifying uncertainty of estimates derived from the alternative data, despite various adjustments, renders using alternative data on its own less feasible and challenging. Just joining the CPI data and the alternative data may not be enough to solve the problem of instability and uncertainty. This study presents the use of propensity score models that seamlessly blend two sets of gasoline data. We assess the precision and the reliability of the estimates from the combined CPI survey data and private sector nonprobability sample data source versus the benchmark (CPI) values. The result seems promising and shows that sufficient benefit can be achieved in using this method to enhance the CPI data.

Thursday, October 26, 2023

1:45 PM

Session J-4: The Importance of Private Sector Data to Federal Statistics

Organizer: Thomas L. Mesenbourg Jr., *CNSTAT*

Chair: Thomas L. Mesenbourg Jr., *CNSTAT*

Location: Vessey 1

This policy session discusses the importance of private sector data to federal statistics. The traditional survey paradigm is under threat and the Nation needs a national data infrastructure that blends data from multiple sources. The private sector is both an important data user and data holder, yet currently the federal statistical

system lacks the robust relationships and mechanisms needed to effectively use private sector data. This session describes recent initiatives and discusses the opportunities and implications of using private sector data to improve the quality, timeliness, granularity of federal statistics. Tom Mesenbourg, Study Director of the recently released consensus report, *Toward a 21st Century National Data Infrastructure: Mobilizing Information for the Common Good*, will focus on report findings related to private sector data. Bill Wiatrowski, BLS Deputy Commissioner (and currently Acting Commissioner) will highlight a recently released ICSP report, *The Use of Private Datasets by Federal Statistical Programs: Extent, Challenges, and Lessons Learned* as well as describing BLS uses. Ron Jarmin, Census Bureau Deputy Director will talk about current Census Bureau uses of private sector data and share his ideas about how statistical agencies need to transform the way they interact with the private sector. Fiona Greig, Vanguard, will be the discussant, providing a private sector perspective.

Panelists:

- Thomas L. Mesenbourg Jr., *CNSTAT*
- William J. Wiatrowski, *Bureau of Labor Statistics*
- Ron S. Jarmin, *U.S. Census Bureau*
- Fiona Greig, *Vanguard*

Thursday, October 26, 2023

1:45 PM

Session J-5: Leveraging Machine Learning to Improve the Accuracy and Reliability of Official Statistics

Organizer: Kyle Hood, *Bureau of Economic Analysis*

Chair: Kyle Hood, *Bureau of Economic Analysis*

Location: *Vessey 2*

Using Machine Learning to Assess Question Performance

Anil Battalahalli, *Westat*

Ting Yan, *Westat*

Hanyu Sun, *Westat*

Computer Assisted Recorded Interviewing (CARI) has long been used by field management to monitor interviewer performance and to assess questionnaire items (e.g. Hicks, Edwards, Tourangeau, McBride, Harris-Kojetin and Moss 2010). Conventionally, a human coder needs to first listen to the audio recording of the interactions between the interviewer and the respondent, and then evaluate and code features of the question-and-answer sequence using a pre-specified coding scheme. Such coding tends to be labor intensive and time consuming. Due to resource constraints, often a subsample of completed interviews can be evaluated in a timely manner. In this study, we will present a pipeline we developed at Westat that heavily draws on the use of modern machine learning methods to detect problematic survey items. Building on literature on behavior coding and question evaluation, we will show how to leverage this pipeline to process recordings and to detect problematic survey items through metrics generated at the end of the processing. We will evaluate the performance of the pipeline using both mock interviews produced in a laboratory setting, and field interviews from a nationally representative survey. We will also discuss the time and cost implications of using the pipeline, compared to the conventional human coding.

Matching Occupational Injury Data Using Augmented Siamese Neural Networks

Elan Segarra, *Bureau of Labor Statistics*

When record linkage efforts involve complex characteristics there is ample potential for general purpose machine learning (ML) techniques to succeed where traditional probabilistic approaches might fall short. However, there can still be pre-processing (e.g., geocoding) and hand-picked comparators that can further

improve linkage outcomes using standard ML models. In this project we present a fusion of these sides we are calling an Augmented Siamese Neural Network. This approach leverages the inherent flexibility of Siamese Neural Networks in a record linkage context while adding additional layers to allow for hand curated comparators that may be difficult for ML optimizers to implicitly identify without sufficiently large, labeled data sets. The framework is used to match establishments from the BLS Survey of Occupational Injuries and Illnesses to establishments in the OSHA Injury Tracking Application. The difficulties inherent in matching company names and addresses and the existence of multi-establishment firms make this a prime application for testing. Linkage outcome metrics of this augmented algorithm are compared both with results from probabilistic methods (i.e., Fellegi-Sunter) and standard machine learning methods to illustrate the added benefits.

Using Machine Learning for Quality Assessments of Call Center Interactions – A Case Study

Elizabeth Nichols, *U.S. Census Bureau*

Monica Puerto, *Accenture Federal*

Brian Sadacca, *Accenture Federal*

Shaun Genter, *U.S. Census Bureau*

Kevin Zajac, *U.S. Census Bureau*

The 2020 Census offered a telephone help line with approximately 7,000 Customer Service Representatives (CSRs) who took calls between March and October 2000. With callers' permission, over 5 million calls were audio recorded. Using nearly 300 scripted Frequently Asked Questions, CSRs answered callers' questions and, in some cases, collected the caller's census information using a scripted electronic census questionnaire. To determine whether CSRs were following the read-as-worded rule for these scripts, quality monitors reviewed a sample of recordings for each CSR. This quality review was a time- and resource-intensive process. It was also error prone because a sampled call might not have been representative of a given CSR's typical performance. As part of the research phase after the 2020 Census, we are investigating ways to automate this quality review using machine learning. Creating a labeled dataset and iterating through different machine learning model methodologies, we compare the words spoken by the CSR in the audio recordings to the scripts to determine how well they followed the read-as-worded rule. This research can potentially allow us to identify CSRs who needed additional training. It can also identify scripts that need reworking because they were difficult to read as worded for many CSRs.

Digitization and Capture as a Service (DCaaS): Facilitating the Transition to Electronic Records for Federal Agencies

Nevada Basdeo, *U.S. Census Bureau*

Kevin L Schweickhardt, *U.S. Census Bureau*

Brandon Michael Dubbs, *U.S. Census Bureau*

This paper outlines a Census Bureau pilot program, known as Digitization and Capture as a Service (DCaaS), aimed at facilitating the transition to fully electronic records in compliance with Memorandum M-19-21 jointly issued by the OMB and NARA. The memorandum requires federal agencies to move towards an electronic environment, including digitizing permanent records before transfer to NARA. However, the conversion of non-digital data to electronic records presents several challenges, including data quality degradation, lack of standardization, high variability, restricted equipment accessibility, high equipment expenses, and the requirement for skilled personnel. Moreover, statistical agencies must take extra measures to protect the data's confidentiality. To address these challenges, DCaaS provides scanning equipment, trained personnel, and quality control measures to accurately capture and preserve records in a secure environment. DCaaS, drawing on lessons learned from the Decennial Census Digitization and Linkage project, aims to improve access to historical records and promote the use of digital records for research and analysis. Methods used in DCaaS include image scanning in a secure government facility, machine (deep) learning, optical character recognition, and natural language processing. The paper also discusses the benefits and challenges of the pilot program and provides recommendations for future implementation.

Session J-6: Federal Statistical Infrastructure for Puerto Rico

Organizer: Carrie Dennis, *U.S. Census Bureau*

Chair: Tomás E. Encarnación, *U.S. Census Bureau*

Discussant: Lorena Molina-Irizarry, *U.S. Department of Commerce*

Location: Room 0105

New Economic Data Products and Efforts to Improve Data Access for Puerto Rico

Michael Lopez-Pelliccia, *U.S. Census Bureau*

This presentation will discuss current efforts at the U.S. Census Bureau to expand economic data products and improve data access for Puerto Rico. A FY24 budget initiative for ~\$11M was developed to support a new annual Puerto Rico Economic Survey as well as a monthly/quarterly economic indicator program for Puerto Rico. These products would provide significantly more insight than the current once-every-5-years view via the Economic Census, and will be designed as part of the Census Bureau's larger ecosystem of data collection and production – integrating with other programs, tools and datasets to provide unprecedented insights into Puerto Rico's economy and workforce. Along with the Deputy Secretary of the Department of Commerce, senior staff from the Census Bureau traveled to Puerto Rico in 2022 and 2023 to meet with key stakeholders from government agencies, academia, and the private sector, and receive input on their most critical data needs. This session will highlight these efforts as well as touch on improving data access via the Federal Statistical Research Data Centers as well as building human capital infrastructure by establishing a partnership agreement between the University of Puerto Rico and the Census Bureau.

The Value of Federal Data for Puerto Rico GDP

Amy Filipek, *Bureau of Economic Analysis*

In 2019 BEA initiated a program to estimate Puerto Rico GDP. This program depends heavily on data collected by other Federal agencies including the Census Bureau. Although more Federal data is collected for Puerto Rico than for the other U.S. territories, the data collections are currently less comprehensive than for the states and the District of Columbia. BEA uses a wide variety of information produced by the Puerto Rico government to supplement the Federal economic data collections. However, a number of information gaps remain, either because no relevant data are collected by any source or because the data are lagged by several years. New Federal data collections, like those proposed in the Census Bureau's FY24 budget initiative and R&D survey initiative, could fill critical gaps in Puerto Rico's economic data that would benefit both BEA and other stakeholders.

Supporting the Transformation of Puerto Rico's Address Infrastructure

Andrea Johnson, *U.S. Census Bureau*

Lynda Liptrap, *U.S. Census Bureau*

The U.S. Census Bureau has been working with public and private stakeholders to improve local addressing systems in Puerto Rico. Through these efforts we are building data equity for Puerto Rico and supporting an equitable address infrastructure for Puerto Rico address data users and creators. The Census Bureau's The Opportunity Project's (TOP) 2022 product development sprints facilitated cross-sector collaboration to solve novel issues – like the local addressing system in Puerto Rico – with data and technology. This presentation will highlight results from the TOP sprint and efforts to build a more robust address framework. The presentation will also feature two upcoming enhancements to the "Census Geocoder" and the "Puerto Rico Predominant Address Type Viewer," both tools to improve understanding and access to Census data for Puerto Rico.

Session K-1: Advances on Federal SOGISC Data: The Federal Evidence Agenda, OMB Guidance, and the NASEM Report

Organizer: Christina Dragon, *National Institutes of Health*

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

Discussant: Christina Dragon, *National Institutes of Health*

Location: Chesapeake A

Measuring Sex, Gender Identity, and Sexual Orientation Moving Forward

Kellan Baker, *Whitman-Walker Institute*

The growing visibility of sexual and gender diversity within the U.S. population, as well as the continued identification of sexual orientation and gender identity disparities in services access, quality, and outcomes, has prompted a re-examination of the ways in which data on sex, gender, and sexual orientation are collected. As part of these efforts, in 2021 the National Institutes of Health asked the National Academy of Sciences (NAS) to convene a committee of experts to review current data collection practices and develop recommendations for measuring sex, gender identity, and sexual orientation in general survey research, administrative settings, and health contexts. The purpose of these recommendations is to enhance the quality and usability of data on these multidimensional constructs for the U.S. population as a whole, improve strategies for identifying LGBTQI+ populations in various contexts, and outline future research directions in the collection and use of these data. This presentation will discuss the committee's approach to defining the constructs of sex, gender identity, and sexual orientation; the principles the committee developed to guide its deliberations; the committee's overall conclusions and recommendations regarding the measurement of these constructs in different settings; and the report's recommended directions for future research and policy efforts.

Best Practices for Collecting SOGI Data on Federal Surveys

Bob Sivinski, *Office of Management and Budget*

Federal surveys play a vital role in generating data the public, businesses, and government agencies need to make informed decisions. Measuring sexual and gender minority (SGM) populations in Federal surveys improves understanding of SGM populations and supports evidence-based policymaking. In accordance with the June 2022 Executive Order 14075 on Advancing Equality for Lesbian, Gay, Bisexual, Transgender, Queer, and Intersex Individuals, the Office of the Chief Statistician of the United States developed recommendations for Federal agencies on the current best practices for the collection of self-reported sexual orientation and gender identity (SOGI) data on Federal statistical surveys. These recommendations build on a long history of robust Federal effort to develop and refine SOGI measurement best practices. This presentation summarizes key aspects of the report, including the importance of continual learning, considerations for including SOGI items on surveys, example approaches for collecting and reporting this information, guidance on how to safeguard SOGI data, and a summary of challenges that need further research.

Leveraging Data to Advance Equity for LGBTQI+ People

Renee Ellis, *U.S. Census Bureau*

Executive Order 14075 on advancing equity for LGBTQI+ people, released in June 2022, specifically called out the role that data collection and evidence use play in advancing equity and full inclusion for this population. This EO emphasized the need for promoting inclusive and responsible Federal data collections and included a number of requirements in pursuit of that goal, such as the development of recommendations for agencies on the best practices for collecting sexual orientation, gender identity, and sex characteristics (SOGI) data on Federal statistical surveys. The EO also required the creation of an interagency working group focused on SOGI data and charged this group with developing the Federal Evidence Agenda on LGBTQI+ Equity, which Federal agencies will use to guide their required SOGI Data Action Plans. The Federal Evidence Agenda,

released in January 2023, marks the Federal Government's first-ever effort to provide a roadmap for agencies to build and use the evidence needed to improve the well-being of and advance equity for LGBTQI+ people. Across its three chapters, it discusses the data needs in this area, a whole-of-government learning agenda to advance LGBTQI+ equity, and guidelines for collecting SOGI data on administrative forms.

Thursday, October 26, 2023

3:30 PM

***Session K-2: Collecting Race and Ethnicity in Establishment Surveys:
Agency Methods and Results***

Organizer: Rebecca L. Morrison, *Bureau of Labor Statistics*

Chair: Rebecca L. Morrison, *Bureau of Labor Statistics*

Discussant: Jeri Mulrow, *Westat*

Location: Chesapeake B

**Challenges in Collecting and Reporting Race and Ethnicity Data in Establishments:
Examples from Two Federal Surveys**

Herman Alvarado, *Substance Abuse and Mental Health Services Administration*

Kathryn Piscopo, *Substance Abuse and Mental Health Services Administration*

Krysten Mesner, *U.S. Census Bureau*

Amy Anderson Riemer, *U.S. Census Bureau*

Establishment surveys face unique challenges when it comes to collection and proxy reporting of demographic data. Issues like inconsistent record keeping procedures, access to records, and proxy reporting burden are common among establishment surveys. These issues raise concerns on data quality and response burden, especially when demographic information is crucial to identify gaps in services provided to target groups in these establishments. The hybrid establishment response process (Willimack, 2007) can be useful in identifying specific areas of disruption that represent a challenge to survey respondents. This presentation will show findings from recent cognitive testing with respondents in two different establishment settings, as well as coding analysis of open-ended respondent feedback. These research methods can shed some light on how these issues affect data reporting of demographic information, particularly OMB's proposed revisions to the race and ethnicity items. This presentation will highlight collection issues in particular for federal surveys such as the National Substance Use and Mental Health Services Survey (N-SUMHSS) and the Survey of Sexual Victimization (SSV).

The Use of Moderated and Unmoderated Cognitive Testing to Evaluate Revised Race & Ethnicity Questions by Proxy Reporting in Establishment Surveys

Krysten Mesner, *U.S. Census Bureau*

Amy Anderson Riemer, *U.S. Census Bureau*

Jessica White, *U.S. Census Bureau*

Patrice Hall, *U.S. Census Bureau*

In early 2023, the Office of Management and Budget (OMB) released recommended revisions (88 FR 5375) to revise OMB's 1997 Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity (SPD 15). To improve the quality and usefulness of federal race and ethnicity data, initial proposals included in the revision process were as follows: collecting race and ethnicity together in a single question in both a long and short version, adding a response category for Middle Eastern and North African (MENA), and updating terminology, definitions, and question wording. As part of this federal initiative, the U.S. Census Bureau tested two versions of the race and ethnicity question on the Annual Business Survey (ABS) using moderated and unmoderated cognitive interviews. The ABS is an establishment survey that contains demographic questions about the business owner(s) as well as various characteristics about the business. Since ABS respondents are sometimes proxies for the business owner(s), we had the

unique opportunity to evaluate the challenges of proxy reporting for these questions. This allowed us to identify potential reporting errors due to data availability (i.e., records) and the perceived sensitivity of collecting race and ethnic identity data for establishments.

Evaluating the Feasibility of Collecting Detailed Race and Ethnicity Data About Students and Employees from Their Institutions

Michael Yamaner, *National Center for Science and Engineering Statistics*

Jennifer Beck, *National Center for Science and Engineering Statistics*

Laura Burns Fritch, *RTI International*

The National Center for Science and Engineering Statistics' (NCSES) portfolio includes surveys that collect demographic information on populations that study or work in STEM. The proposed changes to the standards for collecting race and ethnicity data pose unique challenges to establishment surveys. Employers and institutions often do not maintain detailed administrative records on the demographic characteristics of employees or students. This presentation will share the results of a record-keeping evaluation with academic institutions and research centers on the information they keep and maintain on students and post-docs. Record keeping evaluations are a method used to assess the data that businesses and institutions maintain that they use when responding to surveys. This evaluation includes whether academic institutions and research centers could provide the information that corresponds to the proposed revisions to the reporting requirements.

Public Schools' Student and Teacher Race/Ethnicity Data: Findings from the School Pulse Panel

Rebecca Bielamowicz, *National Center for Education Statistics*

Josue Delarosa, *National Center for Education Statistics*

Ryan Iaconelli, *National Center for Education Statistics*

In response to the proposed revisions to SPD 15, the National Center for Education Statistics (NCES) surveyed school principals about how race and ethnicity data are collected on their students and teachers to assess schools' ability to collect data in line with the proposed changes. Questions were fielded in February 2023 as a part of the School Pulse Panel (SPP) and asked principals about when student and teacher race/ethnicity data are collected, how it is reported, whether their information systems already include information about whether students and teachers are Middle Eastern or North African, and whether their information systems have detailed ethnic categories. This presentation will share these results from the SPP and consider the implications the proposed revisions may have on schools' ability to collect and report these data.

Thursday, October 26, 2023

3:30 PM

Session K-3: Using Data to Overcome the Supply Chain Challenges of Today and Tomorrow

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

Chair: Heather Evans, *International Trade Administration*

Discussant: Jason Bolton, *Bureau of Industry and Security*

Location: Chesapeake C

Re-imagining the Supply Chain at the Census Bureau

Carla Medalia, *U.S. Census Bureau*

Christian Moscardi, *U.S. Census Bureau*

Stakeholders across the federal government, private industry, and the public need insight into the nation's supply chain in order to ensure its resiliency. However, existing Census Bureau data products do not provide

enough comprehensive, timely, or granular information to meet those needs. To fill those data gaps, we are developing a portfolio of efforts seeking to improve our supply chain data infrastructure. This talk will highlight several of these efforts, including the Supply Chain Insights Platform (SCIP). The SCIP uses data visualizations and data science methodologies to combine new and existing data assets that measure different aspects of the supply chain, including manufacturing, international trade, the domestic movement of goods, and sales. Together, these efforts will transform the way the U.S. uses data to monitor its supply chain.

The Role of Transportation Statistics in Measuring the Supply Chain

Rolf Schmitt, *Bureau of Transportation Statistics*

Cha-Chi Fan, *Bureau of Transportation Statistics*

The Bureau of Transportation Statistics (BTS) is the preeminent source of statistics on commercial aviation, multimodal freight activity, and transportation economics, and provides context to decision makers and the public for understanding statistics on transportation. Transportation is one of key components throughout the entire supply chain, linking suppliers and markets, often through multiple modes and multiple facilities. Transportation affects costs and timeliness of shipments, and transportation capacity can be affected by availability of facilities, equipment, and labor and by unexpected surges in shipments. However, transportation might play a less critical role to supply chain disruptions when supply of goods is deficient. BTS measures the physical movement of goods from both shipper and carrier perspectives through multiple freight programs and measures the role of for-hire and in-house transportation in interactions of industries through the transportation satellite account. BTS also engages industry partners to share data to enable more effective supply chain logistic management through the FLOW program.

The Role of Data in Supporting a More Resilient Manufacturing Supply Chain

Nico Thomas, *National Institute of Standards and Technology*

Stephen Campbell, *National Institute of Standards and Technology*

The National Institute of Standards and Technology (NIST)'s Manufacturing Extension Partnership (MEP) is positioned at all levels of the supply chain. The program works with small to mid-sized manufacturers (SMMs), original equipment manufacturers (OEMs), and federal and state stakeholders to connect top-down data and needs requirements with bottom-up capabilities, helping to optimize the resiliency of the nation's manufacturing sector. This presentation will explore how data supports NIST MEP's interactions with the manufacturing supply chain, how data is critical in supporting the resiliency manufacturers and the supply chain – using the program's response to the pandemic as a case study, and how opportunities for increased access to data and information can help mitigate future manufacturing supply chain constraints.

Farm to Market: The Agricultural Supply Chain

Kranti Mulik, *USDA*

The United States Department of Agriculture (USDA)'s Transportation Services Division (TSD) serves as the definitive source for economic analysis of agricultural transportation from farm to market. TSD experts support domestic and international agribusinesses by providing market reports, economic analysis, transportation disruption reports, technical assistance, and outreach to various industry stakeholders. Tracking developments in truck, rail, barge, and ocean transportation, TSD provides information and analysis on the four major modes of moving food from farm to table, port to market.

Thursday, October 26, 2023

3:30 PM

Session K-4: Measuring the Health of the Federal Statistical System

Organizer: Steve Pierson, *American Statistical Association*

Chair: Steve Pierson, *American Statistical Association*

Location: Vessey 1

The statistics community maintains guidelines aimed at protecting federal statistics against threats to its objectivity and resources. These guidelines include the Statistical Policy Directives issued by the U.S. Office of Management and Budget, the National Academies Principles and Practices for a Federal Statistical Agency, and the United Nations Fundamental Principles of Official Statistics. But there is no comprehensive effort to measure the extent to which these guidelines are met in practice, and thus assess the health of the federal statistical system. In this session, panelists discuss strategies for measuring the health of the federal statistical system. Discussion will highlight recent publications on the topic organized through the American Statistical Association.

Panelists:

- Claire Bowen, *Urban Institute*
- Connie Citro, *CNSTAT*
- Nancy Potok, *NAPx Consulting*
- Brian Moyer, *National Center for Health Statistics*
- Jonathan Auerbach, *George Mason University*

Thursday, October 26, 2023

3:30 PM

Session K-5: Seize the Data: Program Oversight, Policy Recommendations, and Insights from Using Linked Survey and Administrative Data

Organizer: Sirin Yaemsiri, *Government Accountability Office*

Chair: Jared Smith, *Government Accountability Office*

Discussant: Jeff Tessin, *Government Accountability Office*

Location: Vessey 2

Blending Probability Sampling with API and Administrative Data: Estimating Recall Rates for Rideshare Vehicles in the 50 States

Abinash Mohanty, *Government Accountability Office*

Sirin Yaemsiri, *Government Accountability Office*

Joanie Lofgren, *Government Accountability Office*

The Infrastructure Investment and Jobs Act of 2021 directed GAO to study the extent of open recalls in passenger vehicles used for ridesharing in each state. GAO obtained vehicle records from two ridesharing companies that represent the vast majority of the US market. GAO used API calls to CARFAX's Vehicle Recall Search Service to measure the open recall status of a portion of vehicles driven in August 2022 for these companies. For the remaining vehicles, GAO used a probability sampling method to estimate the open recall status by collecting data from a manual NHTSA search tool. We used a stratified design and indirect estimation methods to obtain state-level estimates with available resources. While making national estimates of open recalls were relatively straightforward, we discuss our assumptions and their implications for making state estimates of open recalls. The review illustrates the broader potential for combining estimates from probability samples with data from API calls and administrative records.

Unequal Opportunity: Leveraging Blended Data to Assess Dress Code Disparities Among K-12 Students

Frances Tirado, *Government Accountability Office*
Sonya Vartivarian, *Government Accountability Office*

In recent years, researchers, advocates, and parents have expressed concerns about disparities among students in K-12 schools. In particular, school dress codes may disproportionately focus on girls' clothing and bodies, and exclusionary discipline—the practice of removing students from the classroom—for dress code violations may disproportionately harm Black and Hispanic students, among others. Related concerns have been raised about whether assigning law enforcement personnel to schools affects similar groups of students disproportionately. For this talk, we will discuss three analyses: a nationally representative sample designed by GAO to estimate the prevalence, characteristics, and disparities of dress codes in K-12 school districts and two observational studies conducted using NCES complex sample data merged with two data sources. Using these linked data, we performed generalized linear regressions to explore associations between school-level characteristics and policies, such as a strict dress code and the presence of school sworn law enforcement officers, and school-level student outcomes, such as percentage of students arrested or suspended, while controlling for other factors. We will discuss the policy implications of this work and GAO recommendations for agency action.

Disparities in Longevity across Communities in the United States

Lijia Guo, *Government Accountability Office*

Studies on geographic inequalities in life expectancy in the United States have generally focused on single-level analyses of aggregated data at state or county level. Less is known about how socioeconomic conditions and neighborhood-level disadvantage may intersect to contribute to longevity disparity. Reports on COVID-19 mortality variations in small areas within counties further highlighted the needs on community-level analysis. We will discuss how longevity varies by communities and the socioeconomic and other community characteristics associate with longevity disparity across the US. We used government databases from multiple agencies and developed models that identified socioeconomic factors for estimating life expectancy while accounting for other characteristics. Our analysis indicated that US residents in less socially vulnerable communities have experienced larger gains in life expectancy than those living in more socially vulnerable communities. In addition, we found an inverse relationship between community risks of natural disasters and longevity disparity--an indication of how climate change may affect longevity risks. Finally, we found that socioeconomic factors impact longevity most strongly at younger ages, which could help identify policies to reduce longevity disparities across communities.

Thursday, October 26, 2023

3:30 PM

Session K-6: Advances in Estimation and Forecasting: Examples from the Digital Economy, Population Migration, and SNAP

Organizer: Suzanne Strong, *Bureau of Justice Statistics*

Chair: Suzanne Strong, *Bureau of Justice Statistics*

Location: Room 0105

New Estimates of the SNAP Multiplier

Erik Scherpf, *Economic Research Service*

Pat Canning, *Economic Research Service*

Sarah Rehkamp, *Economic Research Service*

Abby Okrent, *Economic Research Service*

In addition to its important role as a safety net for low-income families, the Supplemental Nutrition Assistance Program (SNAP) serves as a unique fiscal lever with which the federal government can provide

stimulus to agriculture, food, and related industries. SNAP benefits not only provide direct added income to businesses where benefits are redeemed but also indirect added income to suppliers and their employees, generating an economic multiplier effect. In this paper we use the USDA Economic Research Service's (ERS) Social Accounting Matrix (SAM) multiplier model, combined with the latest available data, to produce new estimates of the SNAP multiplier. The SAM is an extension of an input-output model—a widely-used framework for evaluating the economic impact of government spending—and offers several improvements over the Food Assistance National Input-Output Multiplier (FANIOM) model that was previously used by USDA ERS to estimate the SNAP multiplier in 2010. Prior estimates from the ERS SAM model using 2016 data put the SNAP multiplier effect at about 1.5. However, the size of the multiplier depends in part on the degree of slackness in the economy and may have changed as economic conditions have changed in recent years.

Forecasting Migrant Encounters at the Southwest Border: Leveraging Quantitative and Qualitative Insights to Maximize Forecasting Accuracy

Justin Schon, *Department of Homeland Security*

Nadwa Mossaad, *Department of Homeland Security*

Douglas Baals, *Department of Homeland Security*

How many migrants will attempt to cross the United States Southwest land border with Mexico in the next 6 months? The Office of Immigration Statistics (OIS) in the Department of Homeland Security (DHS) seeks to answer this question with a mixed methods approach that provides new forecasts each month. First, OIS estimates separate machine learning models for 45 country-family type groups (top 14 countries plus all other, each broken down into Single Adults, Family Units, and Unaccompanied Children). Then, predictions from the statistical models are adjusted based on insights into legal and policy dynamics by subject matter experts from across DHS components and the wider federal government. This mixed methods approach maximizes forecasting accuracy during periods that follow normal historical patterns and during periods with anticipated migration surges or sharp reductions. The final predictions are rigorously validated against actual migrant encounter numbers to evaluate forecasting accuracy. Overall, OIS migrant encounter predictions have an average mean absolute percent error (MAPE) under 10% one month ahead – with the greatest accuracy for Single Adults and Mexican nationals – with growing error further into the future, an unprecedented accuracy given the high uncertainty around these migration flows.

Measuring Digital Intermediation Services in the Digital Economy: Experimental Estimates of Rideshare, Homeshare, and Delivery Services

Tina Highfill, *Bureau of Economic Analysis*

Brian Quistorff, *Bureau of Economic Analysis*

The Bureau of Economic Analysis (BEA) produces economic statistics through its system of satellite accounts that highlight specialized areas of the economy that are not directly apparent in BEA's official economic statistics published under the North American Industry Classification System (NAICS), such as outdoor recreation and arts and culture. BEA recently developed a Digital Economy Satellite Account to better understand this area of the economy as it involves production that spans multiple NAICS industries, ranging from computer manufacturing to internet-based retail trade (e-commerce) to software production. Currently, BEA's digital economy statistics do not fully capture production of digital intermediary services earned from operating a digital platform that facilitates the direct interaction between multiple buyers and multiple sellers for a fee (such as rideshare), resulting in an incomplete picture of the digital economy. In this paper, we discuss a conceptual framework for measuring digital intermediary services across industries, an area of interest to other international statistical agencies as well as BEA. We also provide experimental estimates of gross output for three areas of specific interest to our data users: rideshare, homeshare, and delivery services. Our conservative estimates find digital intermediation services represented at least \$30 billion in 2021 gross output.

Comparing Decennial Census Counts and Local Population and Household Estimates: A Case Study in Fairfax County, Virginia

Xuemei Han, *Fairfax County Government*

Fei Carnes, *Fairfax County Government*

Mabel Ansah, *Fairfax County Government*

Yan Jiang, *Fairfax County Government*

Michelle Gregory, *Fairfax County Government*

Xingyou Zhang, *Bureau of Labor Statistics*

Fairfax County is the most populous jurisdiction in Virginia, with nearly 1.2 million population, about 13% of the state's population. Since 2005, Fairfax County has developed a spatial-explicit systematic application, the Integrated Parcel Lifecycle System (IPLS), to produce official annual estimates and forecasts of population, households, and housing units at the individual parcel level. The IPLS produces highly accurate data, at a 99.3% accuracy rate for 5-year population forecasts and a 97.3% accuracy rate for 10-year forecasts. The system integrates the most current county administrative data across Fairfax County (e.g., real estate and tax files, land use, planning and development data, data from Fairfax Water, Fairfax County Public Schools, and County surveys data), and further integrate data from other external sources, including data from Federal Statistical Agencies (e.g., Census Bureau, CDC, USPS). We compare Decennial Census and local population and household estimates, identify the quantitative and spatial discrepancies, discuss the influence of privacy and disclosure control, and explore potentials for integrating local and federal-level statistical data systems to improve data quality and optimize mutual resource utilization.