



DC|AAPOR



Washington
Statistical
Society

**DC-AAPOR/WSS 2024 Fall Super Review Conference
Program**

September 24, 2024

The National Academy of Sciences
2101 Constitution Avenue NW, Washington, DC

PROGRAM

8:30 - 9:00	Registration (Outside of the Lecture Hall)	
9:00 - 9:15	(Lecture Hall) Welcome & Opening Remarks	
	SESSION A ROOM: Lecture Hall	SESSION B ROOM: NAS 125
9:15 - 10:30	Session 1A: Advanced Techniques in Sampling	Session 1B: Perceptions, Information Security, and Privacy in Survey Data Management
1A moderator: Josh Langeland	Optimal allocation under anticipated nonresponse - Jonathan Mendelson	Model Based Synthetic Data with a Formal Privacy Guarantee - Terrance Savitsky
1B moderator: Sarah Spell	Adapting Cochran's rule for simple random sampling to an estimate from a complex sample - Phillip S. Kott	Protecting Establishment Data with Formally Private Methods: Application to the QCEW - Daniell Toth
	The Use of Big Data-Based Model Prediction for Stratification of Household Addresses - Noah Bassel	Generational Effects on Understanding of the "Need to Know" Principle - Autumn Perkey
10:45 - 12:00	Session 2A: Frame Development, Imputation, and Visualizations	Session 2B: AI and Advanced Text Analysis in Survey Research
2A moderator: Melissa Cidade	Constructing Sample Frames in Business Surveys When No Master Sample Frames Are Available - Jason Kosakow	Modernization of Qualitative Research: Will AI Ever Replace Us? - Darby Steiger
2B moderator: Casey Langer Tesfaye	Imputation of Missing Values in Auxiliary Variables by Low Rank Matrix Approximation - MoonJung Cho	Cracking the Code: AI vs. Human Coding for Open-Ended Survey Questions - Sarah Kahl
	Revising the Census Bureau's Low Response Score - Luke J Larsen	Understanding Mentions of BLS Products Through Topic Modeling of News Articles - Erin Boon
	Using Linked Micromaps for Evidence-Based Policy - Randall Powers	A Framework to Apply Natural Language Processing Techniques to Analyze Public Opinions on Peace and Governance in Africa - Brian Zhou
12:00 - 1:30	Lunch (on your own)	

1:30 - 3:00	Session 3A: Translation Issues and Sensitive Topics	Session 3B: Optimizing Recruitment, Contact Strategies, and Panel Management in Survey Research
3A moderator: Shelley Feuer Session 3B moderator: Stephanie Coffey	Disparate Backgrounds and Competing Priorities: How to Set up a Survey Translation Project and Bridge the Gaps between Team Members - Patricia Goerman Address Collection on an Internet Instrument in English and Spanish - Marcus Berger Lessons in Effective Community Engagement and Trust Building from a Community Based Participatory Research (CBPR) Project - Casey Langer Tesfaye Protocols for Interviewing Teens About Sensitive Topics - Mandi Martinez	When Does Survey Burden Become Too High (Or Too Low) in Probability Panels? - Michael Jackson Evaluating the Use of QR Codes in Panel Recruitment - Kyle Berta Assessing the Value of Inbound Telephone Recruitment in Probability-based Panels - Jared Gumbs Collection and Use of Email Addresses in Establishment Surveys: Results from Two Contact Strategy Experiments - Josh Langeland Exploratory Analysis of Contact Strategies for Annual Economic Surveys - Katharyn Loweth
3:00 - 3:15	Afternoon break/social (Outside of the Lecture Hall)	
3:30 - 4:45	Session 4A: Multi-Mode Survey Design and Incentive Strategies	Session 4B: Supporting Response and Response Outcomes
4A moderator: Marcus Berger 4B moderator: Sarah Lessem	Knowing When to Zoom in: Mode Management When Video Interviewing Takes a National In-Person Study By Storm - Casey Fernandes From me to you': Adapting an interviewer-administered survey to self-administered survey mode - Jonathan Katz Equating Methods in Surveys: Providing Trends when Response Options Change - Rachel Levenstein How Low Can We Go: Optimal Post Incentives in Era of Rising Cost and Declining Response Rates - Molly Fisch-Friedman	Roadmap for Respondents: Developing a Survey Preview for the Annual Integrated Economic Survey - Krysten Mesner When was that? Methods to Improve Use and Recall of a Defined Reference Period - Douglas Williams Climate Crisis and Radicalization: Exploring the Link between Environmental Stress and Political Extremism - Madeline Romm Are Evangelicals a monolith? Exploring attitudes to immigration among Hispanic, Black and white Evangelicals - Angela Eichhorst
4:45 - 5:00	Closing Remarks (Lecture Hall)	
5:00 - 5:30	Travel time to Happy Hour	
5:30 - 7:30	Happy Hour - Hotel Hive Rooftop (2224 F St NW)	

ABSTRACTS – CONCURRENT SESSIONS

Session 1A: Advanced Techniques in Sampling

Presentation Title	Optimal allocation under anticipated nonresponse
Authors	Jonathan Mendelson (U.S. Bureau of Labor Statistics), Michael Elliott (University of Michigan)
Presenter Email	mendelson.jonathan@bls.gov

Survey response rates have declined dramatically in recent years, increasing the costs of data collection. Despite this, there is little existing research on how to most efficiently allocate samples in a manner that incorporates anticipated response rates. Existing mathematical theory for single-stage stratified sample allocation generally assumes complete response. A common practice is to allocate sample under complete response, then to inflate the sample sizes by the inverse of the anticipated response rates. However, we show that this method can fail to improve upon an unadjusted allocation, due to ignoring the associated increase in the cost-per-interview.

We provide mathematical theory on how to allocate sample for single-stage designs in a manner that incorporates the effects of nonresponse on cost efficiency. We derive the optimal allocation for the poststratified estimator under nonresponse, which minimizes either the unconditional variance or the expected costs, holding the other constant, while accounting for uncertainty in the number of respondents. We assume a cost model that incorporates effects of nonresponse. We provide theoretical comparisons between our allocation and common alternatives, which illustrate how response rates, population characteristics, and cost structure can affect the methods' relative efficiency. In an application to public-use data from a self-administered survey of U.S. military personnel, our proposed allocation increases the effective sample size by 25%, compared with common practice.

Presentation Title	Adapting Cochran's rule for simple random sampling to an estimate from a complex sample
Authors	Phillip S. Kott (RTI International)
Presenter Email	philkott1@gmail.com

Given the mean of a simple random sample without replacement, Cochran claimed that calculating a two-sided 95% confidence interval for the population mean using a standard normal-distribution table is often reasonable when the sample size is 25 times the square of the skewness coefficient of the population. We adapt a variant of this crude rule to a two-sided confidence interval for a parameter estimated from a complex probability sample. We conjecture that the standard two-sided 95% confidence interval for an estimated parameter is usually reasonable when the absolute value of the skewness coefficient of the nearly unbiased parameter estimate is less than 0.2. Conversely, it is not reasonable to use the standard two-sided 95% confidence interval for an estimated parameter when the absolute value of the skewness coefficient of its estimate is greater than 0.2. This warning is particularly germane for estimated proportions and differences between proportions.

Presentation Title	The Use of Big Data-Based Model Prediction for Stratification of Household Addresses
Authors	Noah Bassel (NORC)
Presenter Email	bassel-noah@norc.org

The National Survey of Early Care and Education (NSECE) is the most comprehensive study of the availability and use of early care and education (ECE) in the U.S. Because the target population of the NSECE's household survey is a relatively small proportion of all households, the cost of screening households to determine eligibility has always been an important constraint for the NSECE. Like many household surveys the NSECE also faces the twin challenges of declining response rates and rising data collection costs. In response the 2024 NSECE incorporates big data classification and disproportionate stratification into its frame construction and sampling design. Household commercial data are used as inputs for a machine learning model that predicts the probability that a given household on the frame falls within the target population. Household addresses are then stratified accordingly and households with a high probability of eligibility are oversampled. In this study we will evaluate the tradeoff between cost savings and weight variation, along with other implications for sample design.

Session 1B: Perceptions, Information Security, and Privacy in Survey Data Management

Presentation Title	Model Based Synthetic Data with a Formal Privacy Guarantee
Authors	Terrance Savitsky (Bureau of Labor Statistics), Matthew Williams (RTI International), Monika Hu (Vassar College)
Presenter Email	savitsky.terrance@bls.gov

We propose a Bayesian pseudo posterior mechanism to generate record-level synthetic databases equipped with an (ϵ, π) -probabilistic differential privacy (pDP) guarantee, where π denotes the probability that any observed database exceeds ϵ . The pseudo posterior mechanism employs a data record-indexed, risk-based weight vector with weight values $\in [0, 1]$ that surgically downweight the likelihood contributions for high-risk records for model estimation and the generation of record-level synthetic data for public release. The pseudo posterior synthesizer constructs a weight for each datum record by using the Lipschitz bound for that record under a log-pseudo likelihood utility function that generalizes the exponential mechanism (EM) used to construct a formally private data generating mechanism. By selecting weights to remove likelihood contributions with non-finite log-likelihood values, we guarantee a finite local privacy guarantee for our pseudo posterior mechanism at every sample size. Our results may be applied to any synthesizing model envisioned by the data disseminator in a computationally tractable way that only involves estimation of a pseudo posterior distribution for parameters, θ , unlike recent approaches that use naturally-bounded utility functions implemented through the EM. We specify conditions that guarantee the asymptotic contraction of π to 0 over the space of databases, such that the form of the guarantee provided by our method is asymptotic. We illustrate our pseudo posterior mechanism on the sensitive family income variable from the Consumer Expenditure Surveys database published by the U.S. Bureau of Labor Statistics. We show that utility is better preserved in the synthetic data for our pseudo posterior mechanism as compared to the EM, both estimated using the same non-private synthesizer, due to our use of targeted downweighting.

Presentation Title	Protecting Establishment Data with Formally Private Methods: Application to the QCEW
Authors	Daniell Toth (Bureau of Labor Statistics)

Presenter Email	Toth.Daniell@bls.gov
This talk presents a practical application of a formally private method to Quarterly Census of Employment and Wages data. We will show discuss a specific application with define parameters to this important economic dataset. We will demonstrate the guaranteed level of protection for each establishment and the effect this limitation procedure has on the utility of the dataset; specifically, on the accuracy of key economic statistics.	

Presentation Title	Generational Effects on Understanding of the “Need to Know” Principle
Authors	Autumn Perkey (University of Maryland), Laura Short (University of Pittsburgh), Ella Reid (University of North Georgia), Gary Brown (The Bush School of International Affairs)
Presenter Email	perkey@umd.edu
A valid “Need to Know” is one of the three requirements for access to classified information. Given recent violations on upholding this principle through examples such as Private Manning, Airman Teixeira, and Senior Airman Winner As younger generations enter the national security workforce understanding generational effects on understanding national security information policy is critical. We examine the relationship between generational effects and understanding of the “need to know principle” using subject matter expertise interviews and a novel survey examining. Our survey examined how key demographic characteristics, experience and role in the security community, values, and other measures of personality characteristics informed an individual’s understanding on “need to know” expectations. We sought to understand how generational perspectives on “need to know” differed, and how our findings could be used to informs DoD information security policy and future research. Overall, our research points to a more mixed result of generational effects being driven more by individuals’ characteristics, organizational culture, and understanding of “need to know” than generations alone.	

Session 2A: Frame Development, Imputation, and Visualizations

Presentation Title	Constructing Sample Frames in Business Surveys When No Master Sample Frames Are Available
Authors	Jason Kosakow (Federal Reserve Bank of Richmond), Acacia Wyckoff (Federal Reserve Bank of Richmond)
Presenter Email	jason.kosakow@rich.frb.org
To conduct high-quality surveys, it is important for survey professionals to have a sampling frame that covers the entire target population. However, access to frames that have high levels of coverage for establishment surveys are only available for some government agencies. Due to this, surveys among establishment usually rely on using third-party data to build sampling frames. Third party data has been known to have high rates of inaccurate contact information and not cover the entire population, thus, increasing the challenges of conducting high quality establishment survey research. The Federal Reserve Bank of Richmond attempted to create a sample frame of businesses with high-levels of coverage by using the USPS CDS file, similar to what is done for many household surveys. Using third party data and the Google Places API, the Richmond Fed was able to create a frame that covers a majority of addresses listed as businesses in the USPS CDS file.	

Presentation Title	Imputation of Missing Values in Auxiliary Variables by Low Rank Matrix Approximation
Authors	MoonJung Cho (Bureau of Labor Statistics)
Presenter Email	mjeccho@gmail.com
Analysis of sample survey data often requires adjustments for missing values in the variables of interest. Standard adjustments based on item imputation or on propensity weighting factors rely on the availability of auxiliary variables for both responding and non-responding units. Their application can be challenging when the auxiliary variables are themselves subject to incomplete-data problems. This paper shows how low rank matrix approximation can be applied to impute missing auxiliary variables. Performance depends on the number of auxiliary variables and the extent of missingness. We evaluate imputation by low rank matrix approximation in terms of bias and mean squared error.	

Presentation Title	Revising the Census Bureau's Low Response Score
Authors	Luke J. Larsen (Census Bureau), Joanne Fane Lineback (Census Bureau), Eric Jensen (Census Bureau), Yazmín García Trejo (Census Bureau)
Presenter Email	luke.j.larsen@census.gov
In an effort to predict which areas across the United States are less likely in aggregate to self-respond to the decennial census, the Census Bureau developed the Low Response Score (LRS) in the early 2010's to prepare for the 2020 Census. Published at the tract and block-group levels, the LRS is an OLS-based metric derived with predictors from the decennial census and the American Community Survey (ACS). While the LRS has been employed by field operations managers, data scientists, and the general public for many uses over the past decade, it has become less useful over time after the 2020 Census had come and gone. With the latest decennial census came many changes, some expected (for example, the implementation of an internet self-response mode) and others unexpected (most infamously, impacts on data collection due to the coronavirus pandemic). Clearly, the LRS model would have to be revised for a post-2020 environment, and so a Census workgroup was established to research, develop, and deploy this new metric onto a forthcoming release of the Census Planning Database (PDB). In this paper, the authors employ data from the 2023 PDB to study the impact of applying several “common sense” adjustments to the LRS model (e.g., using total self-return-rate as response variable, adding internet access predictors, and implementing regional interactions). They conduct regression diagnostics to assess the results, and a comparative analysis of “old” versus “new” scores demonstrates the overall impact of the effort to improve prediction of low self-responding areas. Researchers who use the LRS as proxy for response propensity of their own survey programs will find that the updated metric will better reflect salient characteristics of local communities and more accurately predict which tracts and block groups are prone to lower self-response rates.	

Presentation Title	Using Linked Micromaps for Evidence-Based Policy
Authors	Randall Powers (Bureau of Labor Statistics), Wendy L. Martinez (Census Bureau)
Presenter Email	Powers.Randall@bls.gov

Linked micromaps were developed to display geographically indexed statistics in an intuitive way by linking them to a sequence of small maps. The approach integrates several visualization design principles, such as small multiples, discrete color indexing, and ordering. Linked micromaps allow for other types of data displays that are connected to geography, including scatterplots, boxplots, time series plots, confidence intervals, and more. Initial applications of micromaps used data from the National Cancer Institute and the Environmental Protection Agency. In this presentation, we will show how linked micromaps can be used to better understand and explore relationships and distributions of statistics linked to US states and DC. We will compare linked micromaps with other popular data displays, such as bubble charts, choropleth maps, and bar charts. We will illustrate how linked micromaps can be used for evidence-based decision-making using data from the Bureau of Labor Statistics, the Census Bureau, and the Economic Research Service.

Session 2B: AI and Advanced Analysis in Survey Research

Presentation Title	Modernization of Qualitative Research: Will AI Ever Replace Us?
Authors	Darby Steiger (SSRS)
Presenter Email	dsteiger@ssrs.com
Qualitative research methods have been evolving rapidly in recent years, especially in response to the COVID-19 pandemic, when focus groups and in-depth interviews shifted almost entirely to virtual methods (Keen et al, 2022; Roberts, Pavlakis & Richards, 2021). Since then, generative AI is now offering unprecedented access to tools that have potential to help us become more efficient in our qualitative research processes. For example, generative AI can perform tasks such as suggesting research hypotheses, drafting questions for discussion guides and recruitment screeners, suggesting probes for participants, transcription, summarizing qualitative data, and offering recommendations. However, these tools also raise concerns that do not always align with our needs and obligations as social science researchers (Christou, 2023). For example, recent literature cautions that adoption of AI may degrade the human element of research, may introduce societal biases, may produce different results than human analysis, and may introduce ethical concerns. However, these same studies also point to the benefits of embracing AI in qualitative research, noting that its efficiency, its ability to support the structure of researchers' established coding structures, and its potential to address equity issues (Hamilton et al, 2023; Anis and French, 2023; Marshall and Naff, 2023). Over the past year, SSRS has been exploring ways AI can be used to enhance our qualitative toolbox, including testing the incorporation of AI into translation, probing, and analysis. We will explore selected case studies that compare: (1) traditional cognitive interviews with AI-assisted asynchronous cognitive interviews, and (2) traditional qualitative analysis methods with AI-assisted analysis. We will discuss the benefits and downsides we have seen with each approach, along with considerations and implications for the future.	

Presentation Title	Cracking the Code: AI vs. Human Coding for Open-Ended Survey Questions
Authors	Sarah Kahl (Ipsos Public Affairs), Kelly Bell (Ipsos), Ruth Young (Ipsos), Sherin Mattappallil (Ipsos), Jordan Forrest (Ipsos)
Presenter Email	sarah.kahl@ipsos.com

AI advancements are ushering in a significant transformation within the field of public opinion research, with coding and data processing standing out as particularly ripe areas for innovation. This presentation delivers a detailed analysis comparing the coding precision of a Large Language Model (LLM) and human researchers across a wide spectrum of open-ended survey responses.

We utilize data sourced from the Ipsos KnowledgePanel, a probability-based online panel, to delve into several open-ended questions from a survey centered around Veterans. The responses were coded by Ipsos TACTIC (Test, Analyze, Conceptualize, Track, Improve, and Communicate), which utilizes an LLM (GPT-4), and human researchers, facilitating direct and comprehensive comparisons. The survey included both short and long open-ended questions, which further enables a comparative study across different question types. Our research will scrutinize the relative accuracy of AI coding versus human coding, evaluate the accuracy of the TACTIC LLM, and explore whether certain types of open-ended questions are more conducive to AI coding.

The potential for AI coding to drastically reduce the time and cost associated with analyzing open-ended survey questions could have profound implications for the industry.

Presentation Title	Understanding Mentions of BLS Products Through Topic Modeling of News Articles
Authors	Erin Boon (Bureau of Labor Statistics)
Presenter Email	Boon.Erin@bls.gov
The Bureau of Labor Statistics (BLS) measures labor market activity, working conditions, price changes, and productivity in the U.S. economy to support public and private decision making. To meet this mission, BLS not only publishes statistics and research on its own website but also seeks to understand when and where its products are mentioned in online news sources. Making sense of this huge volume of news articles is impossible without a means of summarizing and grouping them. Using article data collected by a third-party service, we experimented with several methods to model the topics contained in news articles that mention BLS products. We compared and optimized candidate models with a goal of meeting the needs of internal stakeholders who use the output to help evaluate the impact of their outreach efforts. Ultimately, we selected a model that provided the best balance of evaluation metrics and utility to these users. This presentation will include a summary of the models we explored and the process we developed to compare them.	

Presentation Title	A Framework to Apply Natural Language Processing Techniques to Analyze Public Opinions on Peace and Governance in Africa
Authors	Brian Zhou (Harvard University), Alvan Arulandu (Harvard)
Presenter Email	brianzhou@college.harvard.edu
The intersection of governance, development, and peace is a difficult prioritization problem for African governments and requires the participation of multilateral stakeholders including the United Nations, African Union, and other bodies. Historic studies in governance have analyzed datasets including resources such as the United Nation General Debates and surveyed citizens on public opinion (such as the Afrobarometer), but lack information about public opinion and discourse on governance; for example, social media was pivotal in showing public dissatisfaction on governments during the Arab Spring.	
This paper offers a framework to apply Natural Language Processing to fuse various barometers of public opinion together. It aggregates text articles sourced from international news media, African media sources (IRIN, BBC Africa, Radio France Africa, etc.), and emerging/social media (Twitter, Instagram, etc.), to develop a quantitative statistical model that is a real-time tracker on governance and issues as they emerge in a country.	

Combined, these sources of public opinion allow this paper to analyze specific community opinions on political governance, economic governance, institutional governance, and general governance. The paper applies and explores natural language processing techniques to analyze a large variety of Corpora (compilations of text) to assist in qualitative studies that analyze the strengths of each platform in providing insight on public opinion.

By synthesizing sources of public opinion on various governance aspects through advanced NLP techniques, this study provides nuanced insights into public discourse on African governance and suggests new approaches that can combine strengths and weaknesses of current governance polling, news media, and social media. This framework can significantly inform qualitative studies, enhancing the understanding of governance issues as they unfold in real-time.

Session 3A: Translation Issues and Sensitive Topics

Presentation Title	Disparate Backgrounds and Competing Priorities: How to Set up a Survey Translation Project and Bridge the Gaps between Team Members
Authors	Patricia Goerman (Census Bureau), Ariana Muñoz Maurás (Census Bureau)
Presenter Email	patricia.l.goerman@census.gov

The survey translation process can involve many practitioners with different backgrounds, such as survey sponsors, translators, survey methodologists, subject matter experts and linguists. These different team members often have competing priorities and constraints to consider as they work towards the creation of different language versions of a survey instrument. Tensions between these different actors can arise and can be difficult to resolve when each is approaching the process from a different perspective. A great deal of 3MC (multinational, multicultural, or multiregional) survey literature delineates the work required to achieve functional equivalence across different language versions of a survey instrument (Cross Cultural Survey Guidelines, 2016). A recent book by Pan, et.al. 2020 addresses many of these issues, providing an overview of the complexity and steps that should ideally take place in a survey translation project such as design, translation, review and pretesting and recommending a sociolinguistic framework as a basis to bridge the gaps between team members with different backgrounds. What is missing from the book is where a particular project team should start in terms of finding or developing the necessary skill set and implementing the recommendations. This presentation will focus on gaps in setting up and carrying out survey translations from recent real-world experiences in a large U.S. survey organization and will propose practical guidance for surveys and agencies new to the process.

Presentation Title	Address Collection on an Internet Instrument in English and Spanish
Authors	Marcus Berger (Census Bureau), Elizabeth Nichols (Census Bureau), Alda Rivas (Census Bureau), Angie O'Brien (Census Bureau)
Presenter Email	marcus.p.berger@census.gov

The U.S. Census Bureau has been conducting a small-scale iterative testing program to plan and prepare for the 2030 Census. This program consists of iterative rounds of usability testing in English and Spanish, followed by testing with nationally representative field test samples. As part of the testing of the online instrument for this survey, respondents are asked to enter their street address. They are then presented with an address standardizer screen, which asks them to confirm their address using standard formatting and abbreviations. This

presentation analyzes address submissions from three rounds of usability interviews, as well as one round of field testing.

Address collection is critical to the census, but also many other online forms. During the usability testing, we noticed that sometimes, participants enter data into the wrong field, especially during Spanish-speaking sessions. We will summarize our findings from three rounds of usability interviews on this screen, focusing on errors in the address collection screen. For any apparent errors, we will analyze the errors and any relationship they have to other characteristics such as the type of device the participant used, the language presented (English or Spanish), or other possible correlates, like age of the respondent or whether the respondent is the owner or renter of the home.

We also analyze responses to the address collection screen from the field test, conducted in spring of 2023, analyzing any clear errors and comparing responses of English and Spanish speaking respondents. Finally, based on the results of this comparison, we make recommendations to avoid response errors on the address collection screen and provide avenues for future testing.

Presentation Title	Lessons in Effective Community Engagement and Trust Building from a Community Based Participatory Research (CBPR) Project
Authors	Casey Langer Tesfaye (Westat), Joy Thomas (Westat), Jennifer Brooks (Westat), Anna Cupito (Westat)
Presenter Email	caseytesfaye@westat.com

In the last two decades, the Centers for Disease Control and Prevention has emphasized community engagement, which involves working collaboratively with groups of people who share similar interests, contexts or circumstances, or live in the same geographic area, to address issues affecting their well-being. Notably, the CDC has supported community-engaged research methods such as Community-Based Participatory Research (CBPR) projects, which involve significant input from the community throughout the project lifecycle from design through dissemination

(1). The aim of CBPR projects is to produce an outcome or product that has long-lasting positive effects within the community. Ultimately, effective engagement in CBPR leads to community empowerment—the ability of communities to develop, sustain, and grow an organized base of people who act together to set agendas, influence public discourse, and make decisions about community priorities and how to address them

(2). CBPR is particularly well-suited when focusing on sensitive topics or working with vulnerable populations or communities with significant mistrust.

For the current project with CDC, we were interested in reaching out to communities and developing partnerships in three locations with relatively low uptake of the COVID-19 vaccine. The goal was to co-create effective health messaging for the community. The project involved the recruitment of community leaders and then vaccine-hesitant community members; an initial focus group of the vaccine-hesitant community members; and 16-20 sessions to build trust, work through design thinking exercises, and co-create health messaging. In this paper, we discuss some of the challenges and lessons learned for building and maintaining trust, finding and engaging participants, finding common ground, empowering communities cocreation and shared governance, and producing lasting benefits for and with communities facing distinct public health concerns, barriers, and sources of mistrust.

Presentation Title	Protocols for Interviewing Teens About Sensitive Topics
Authors	Mandi Martinez (Census Bureau), Betsarí Otero Class (Census Bureau), Ariana Muñoz Maurás (Census Bureau)
Presenter Email	mandi.martinez@census.gov
The purpose of cognitive interviewing studies is to understand how respondents interpret questions and evaluate how well respondents' interpretations match the intended purpose of these questions. Conducting cognitive interviews with teens between the ages of twelve through seventeen about a variety of sensitive topics presents some unique challenges. Researchers must coordinate with parents during the recruiting and screening phase, establish rapport in order to effectively probe on sensitive topics, and ensure that interviewers are trained on how to respond if teens experience distress while discussing sensitive topics. Some of the challenges posed by interviewing teens on sensitive topics can be addressed with a well thought out data collection plan. What does the process of conducting cognitive interviews look like for this vulnerable population? How do we implement and train researchers on the distress protocols, or procedures to follow if teen respondents get upset during the interview? How do we navigate the recruiting and consent process when parents are speaking on behalf of their children? Should researchers provide resources for support related to topics that may be discussed during the interview? In this presentation we describe common challenges and lessons learned regarding conducting cognitive interviews with adolescents and how these challenges are impacted by sensitive topics. We illustrate this through the discussion of three rounds of cognitive testing on the NHIS-Teen followback survey for the National Health Interview Survey. The survey spans several health topics, including teens' physical and mental health and well-being, social and emotional support, and experiences with bullying and discrimination.	

Session 3B: Optimizing Recruitment, Contact Strategies, and Panel Management in Survey Research

Presentation Title	When Does Survey Burden Become Too High (Or Too Low) in Probability Panels?
Authors	Michael Jackson (SSRS)
Presenter Email	mjackson@ssrs.com
In recent years, multi-purpose probability panels—recruited via probabilistic methods and then made available for sampling into individual studies—have become increasingly popular. Because probabilistic recruitment is much more expensive than opt-in recruitment, panelist retention is a critical priority for such panels. Accordingly, probability panel providers typically make some effort to limit the frequency with which panelists are sampled for surveys, on the theory that excessive burden may drive attrition and/or other undesirable panelist behavior (such as providing low-quality data). However, the threshold at which invitations become “too frequent” in a general-population panel is unclear. Furthermore, there is a potential tradeoff between burden control and representative study-level sampling, as panelists from lower-response-propensity subgroups need to be oversampled (and thus sampled more frequently on average) to counteract lower expected completion rates. Finally, there also may be a point at which invitations become “too infrequent”, leading to attrition via insufficient panelist engagement. In this paper, we will attempt to shed light on these tradeoffs by assessing relationships between panelist burden and several undesirable outcomes, including attrition, survey- and item-level nonresponse, speeding,	

and satisficing. To do so, we will leverage a rich database of paradata aggregated from hundreds of studies run on the SSRS Opinion Panel, a nationally representative panel recruited primarily via address-based sampling.

We will primarily define burden in terms of the number of study invitations, which has the longest time series; but will also consider alternative definitions such as the average completion time. By quantifying the relationship between burden and various undesirable outcomes, we will attempt to derive data-driven guidelines for the maximum and/or minimum frequency with which probability panelists should be invited into surveys. Such guidelines could be used to inform within-panel sampling algorithms and/or targeted recruitment strategies to optimize panelist burden and improve panel health.

Presentation Title	Evaluating the Use of QR Codes in Panel Recruitment
Authors	Kyle Berta (SSRS)
Presenter Email	kberta@ssrs.com
Researchers are increasingly relying on probability-based panels for their data collection needs and moving away from traditional forms of data collection such as RDD (Random Digit Dial) phone. In this paper we explore the use of QR codes for panel recruitment. The use of QR codes has become increasingly commonplace in the mainstream and, in turn, survey research. As such, SSRS commonly uses QR codes on invitation letters for panel recruitment surveys. In this paper we explore demographic differences between respondents who participate in recruitment surveys via QR codes vs those who don't use the QR codes. We derive insights on what this may mean for researchers using or building panels. We utilize data from the recruitment survey for the probability based SSRS Opinion Panel for this research.	

Presentation Title	Assessing the Value of Inbound Telephone Recruitment in Probability-based Panels
Authors	Jared Gumbs (SSRS), Cameron McPhee (SSRS), James McKinstry (SSRS)
Presenter Email	jgumbs@ssrs.com
Probability-based panels continue to be a popular source of high quality, affordable, data collection. While probability-based panels provide an attractive alternative to large scale dual-frame RDD campaigns, we don't think it's time to throw out all the headsets quite yet. The SSRS Opinion Panel is a multi-mode, probability-based, nationally representative panel of adults 18 or older recruited via Address Based Sample (ABS). At the recruitment stage we give all respondents the option to either complete the registration survey online or by calling a toll-free phone number that is answered by live interviewers. In this paper we will examine demographic and attitudinal differences between the panelists that chose to register via inbound phone call and those that register online, to see if we are in fact attracting diverse panelists by offering the inbound phone option. Furthermore, we will investigate all the phone calls made to our inbound recruitment phone number to see what we can learn about those who were enticed enough to call the phone number but ultimately didn't participate versus those that called in and actually participated. To accomplish this, we will do a reverse look up of our contact system Automatic Number Identification records to get demographic information on those who called our inbound line but didn't participate.	

Presentation Title	Collection and Use of Email Addresses in Establishment Surveys: Results from Two Contact Strategy Experiments
Authors	Josh Langeland (Bureau of Labor Statistics), Victoria Narine
Presenter Email	langeland.joshua@bls.gov
Email is a convenient tool for survey communications due to its low cost and the ability to include a link to online reporting systems. Withholding a paper form makes respondents more likely to report their data online,	

thus decreasing costs and improving data quality. However, emailed invitations typically achieve a lower response rate than postal mailings. Firewalls and spam filters are obstacles for successfully delivering invitations and messages may become buried in busy inboxes. Further, since email addresses are typically only available for specialized populations (e.g., students at a university), much of the research on email communications in the survey literature is not generalizable to wider populations.

To better understand how emailed communications impact response behavior within the unique response model for business surveys (see Willimack and Nichols), a pair of experiments on the use of email communications in establishment surveys were conducted by the U.S. Bureau of Labor Statistics (BLS). The first experiment varied the timing and number of emailed nonresponse follow-ups and examined the impact on response behaviors. The study found that earlier and more frequent communications had a positive impact on survey response. The second experiment examined the effectiveness of placing phone calls to businesses to collect email addresses. This study found that placing a single phone call to establishments greatly increased the number of units that provided an email address. This presentation discusses the results of the experiments and how they were used to implement changes in data collection methodologies at BLS.

Presentation Title	Exploratory Analysis of Contact Strategies for Annual Economic Surveys
Authors	Katharyn Loweth (Georgetown University), Hannah Wen (University of California, Davis)
Presenter Email	ksl84@georgetown.edu
Communication strategies are central components of survey production. The emails, postal mailings, and phone calls of a communication strategy inform and remind potential survey respondents about the importance of survey completion. While these communications overall spur survey engagement, there is varying empirical evidence on the impact of these different communication methods and their particularities, such as sequencing and timing, on respondent engagement. As part of the Coding it Forward 2024 Fellowship program, the Economic Directorate of the U.S. Census Bureau sponsored an investigation into the contact strategies of several annual legacy surveys utilizing data from survey years 2015-2022. The goal of the analysis was to identify trends in communication strategies over time and to better understand the relationship between contact efforts and survey engagement. The recommendations from this study could then inform the future communication efforts of business surveys such as the Annual Integrated Economic Survey. This presentation provides an overview of the communication strategy analysis performed as part of the Fellowship program. First, we discuss the data cleaning and transformation activities completed to utilize information from the communication strategies documentation as part of quantitative analysis. We then discuss the historical trends in communication strategies over time and the contact methods' different engagement windows. We end with recommendations for best practices in data handling and planning future communication strategies to better discern the causal impact of different contact methods on respondent engagement.	

Session 4A: Multi-Mode Survey Design and Incentive Strategies

Presentation Title	Knowing When to Zoom in: Mode Management When Video Interviewing Takes a National In-Person Study By Storm
Authors	Casey Fernandes (Westat), Ryan Hubbard (Westat)

Presenter Email	caseyfernandes@westat.com
The Medical Expenditure Panel Survey adopted a video interviewing option beginning in 2021 that is growing greatly in popularity. In 2023, thousands of interviews were completed by video rather than the traditional 5 rounds of in-person data collection. We examine the evolving experimental and operational decisions for allowing a video interview to replace an in-person interview, the success of these efforts, and the growing need to model mode choice based on contact history, case characteristics, past health-care utilization, and interviewer/field considerations. The goals of increasing response rates, lowering costs, and reducing respondent burden are factors that shape future case management.	

Presentation Title	From me to you': Adapting an interviewer-administered survey to self-administered survey mode
Authors	Jonathan Katz (Census Bureau), Alexandra Piccirillo (Census Bureau), Caitlyn Keeve (Census Bureau), Jessica Holzberg (Census Bureau)
Presenter Email	jonathan.m.katz@census.gov
Many interviewer-administered surveys are considering adding a self-administered web survey mode to their data collection strategy. Transitioning interviewer-administered surveys to self-administered can reduce costs and respondent burden. However, there are concerns about how to effectively adapt questionnaires for multi-mode data collection in a way that minimizes mode effects. In this presentation, we will discuss the process of developing and cognitively testing a self-administered questionnaire for the National Telecommunications and Information Administration Internet Use Survey (IUS), a supplement of the interviewer-administered Current Population Survey. We first tested the interviewer-administered IUS to identify any necessary changes to the current questionnaire. Following this testing, we drafted an adapted self-administered web questionnaire of the IUS. Although we attempted to maintain the comparability of wording when possible, some larger changes were necessary (e.g., converting an open-ended question to select all that apply and modifying conversational and instructional text). Some changes were also made to take advantage of features of the mode and ideally improve data quality and reduce respondent burden (e.g., combining multi-step questions into a single question). We then conducted exploratory cognitive and usability testing of the self-administered instrument to gain an understanding of how these questions performed and of possible mode effects that may occur. In this presentation, we will review the process of creating the self-administered instrument and findings of both interviewer and self-administered testing. Finally, we will discuss select questions where we anticipated differences in comprehension and difficulty between the two modes and whether the two questionnaires were successful at capturing the same measurement.	

Presentation Title	Equating Methods in Surveys: Providing Trends when Response Options Change
Authors	Rachel Levenstein (Zencity), Cachal Neuburger (Tel Aviv University)
Presenter Email	rachel.levenstein@zencity.io
Questionnaires in longitudinal surveys often stay the same over time to ensure comparability from cycle to cycle. When improvements to scales can or need to be made, we can choose to stop reporting trends between changes, or we can find methods to equate the scores for both versions. These issues also present themselves in standardized tests used in education. Frequently, additional test items are added with varying levels of difficulty. But when assessments change their content, trends must still be maintained to measure student growth. Kolen & Brennan (2014) and many others outlined methods of	

equating test scores when the content of the assessments change. In this paper, we compare these methods using survey data, potentially allowing us to make comparisons even when response scales have changed.

The Zencity Blockwise Survey measures trust in police and perceptions of safety in cities and counties in the United States and the UK. Blockwise has been used by law enforcement agencies to monitor and ultimately improve public trust in their cities since 2017. Police leadership uses trends in the survey results to understand trust and safety, particularly at more granular geographic levels (e.g., police precinct) and demographic categories. These data can be used for communications and for evaluating policing practices and initiatives. Prior to 2023, the Blockwise survey consisted of three questions on an unlabeled 0-10 scale. In 2023, to provide more reliable estimates of trust and safety, additional questions were added and the scales for the three original items were put on a four point, fully labeled scale. To ensure comparability of the previous survey to the new methodology, we converted scores on the 0-10 scale to the four point scale using equating methods. This paper will compare these methods, showing change over time by geographic area and by demographic characteristics.

Presentation Title	How Low Can We Go: Optimal Post Incentives in Era of Rising Cost and Declining Response Rates
Authors	Molly Fisch-Friedman (SSRS), Robyn Rapoport (SSRS), Elizabeth Sciupac (SSRS), Rob Manley (SSRS), Cameron McPhee (SSRS)
Presenter Email	mfisch-friedman@ssrs.com

Prior research has shown the power of monetary pre-incentives and post-incentives in encouraging cooperation and boosting response rates. Typically, the higher the post-incentive, the better the response rate. However, most studies are subject to budget limitations and therefore seek the most bang (i.e., response rate) for their buck. So, what is the optimal post-incentive amount, given the inclusion of a pre-incentive, that maximizes response rate within budget?

In 2023, SSRS conducted a large, national address-based sample (ABS) study with budget constraints around the pre- and post-incentives. As such, building on past experimental research, we tailored a design that included different pre- and post-incentive structures for key groups.

One group – about 20% of the sample in the highest income strata – received a survey invitation with a cash pre-incentive of \$1.25 but were not offered a post-incentive. This intentional targeting was based on results from previous experimentation, which indicated the yield among this high-income group would be relatively high without offering a post-incentive, as higher-income households are typically more responsive.

The remainder of the sample, roughly 80%, were randomly assigned to be offered either a \$5 or \$10 post-incentive, along with a \$1 cash pre-incentive, allowing us to evaluate the difference between each group in terms of both yield and cost-per-interview (CPI). Overall, we found the \$10 post-incentive performed markedly better than the \$5 offer on both measures, allowing us to adjust our protocol appropriately in Wave 2 of the project and maximize the limited budget.

This presentation will evaluate the effectiveness of the full bespoke design, including how different incentive treatments may or may not affect the demographics of completed interviews. In addition, we will explore avenues for future research, including the potential to reconsider incentivization theories that challenge the benefits and feasibility of the “pre-incentive, no post-incentive” treatment, even for constrained budgets.

Session 4B: Supporting Response and Response Outcomes

Presentation Title	Roadmap for Respondents: Developing a Survey Preview for the Annual Integrated Economic Survey
Authors	Krysten Mesner (Census Bureau), Melissa Cidade (Department of Justice Statistics)
Presenter Email	krysten.mesner@census.gov
The Annual Integrated Economic Survey (AIES) launched in March 2024, integrating and replacing seven long-standing annual business surveys into one collection. The AIES is designed to combine Census Bureau collections to increase data quality, reduce respondent burden, and allow the Census Bureau to operate more efficiently. Establishment surveys often rely on a survey preview to support response. However, there is no designated survey form for the AIES as questions are specific to a company's industry (or industries). As such, the Census Bureau developed an interactive content selection tool for previewing questions specific to a company's assigned North American Industry Classification System (NAICS) code(s) to support the AIES. This presentation provides an overview of the evolution of the AIES survey preview materials over the course of the research phases preceding the production launch in March. We then discuss the iterative testing of the content selection tool in preparation for the March launch, including cognitive and usability testing to refine the tool. We end with recommendations for continued improvements and best practices for survey previews for large, complex establishment surveys.	

Presentation Title	When was that? Methods to Improve Use and Recall of a Defined Reference Period
Authors	Douglas Williams (Bureau of Labor Statistics), Aleia Clark Fobia (Census Bureau)
Presenter Email	williams.douglas@bls.gov
Surveys commonly use reference periods to define a domain for respondents to report the frequency of behaviors, or the incidence of events. Reference period research has generally focused on the effects of period length on recall, or how the implied precision (e.g., typical week vs last week) affects frequency reports. However, there is a lack of research on how consistently reference period statements are interpreted, or how well respondents adhere to defined periods over the course of a survey. Consistent understanding of these periods is important for surveys such as the Current Population Survey (CPS), which measures employment-related activities over a defined 7-day period covering the 12th of the survey month. Previous research on the CPS determined that the use of "last week" led to varied interpretations by respondents, resulting in specific month and day bounds provided as respondent instructions at the start of the survey. For the remainder of the survey "last week" is used. As the survey progresses, the cognitive demands of processing new information can affect respondents' ability to accurately recall the reference period as defined. In this paper we review the results of an experiment conducted with a panel of respondents, where repeating bounded dates throughout the survey was varied (versus using "last week"). At the conclusion of the self-administered survey, respondents were asked to report the time period they used when answering questions. Repeating the date demonstrated improved recall accuracy. We present results across respondent characteristics and examine the types of recall rules respondents employed when deviating from the bounded period (e.g., last seven days from interview; using a scheduled workweek, etc.). Findings also contribute to research into how to render interviewer-administered instruments into a self-administered mode.	

Presentation Title	Climate Crisis and Radicalization: Exploring the Link between Environmental Stress and Political Extremism
---------------------------	--

Authors	Madeline Romm (START), Marcus A. Boyd (ARLIS), Samuel Henkin (NSI), Sean Doody (START), Allison Bredder (University of Maryland), Melissa Cidade (George Mason University)
Presenter Email	mromm@umd.edu
The presentation delves into the complex nature of climate change, describing it as a "hyperthreat" that requires global cooperation and innovative approaches beyond traditional centralized responses. It underscores the generational divide in perceiving and addressing climate change, with younger individuals experiencing higher levels of climate anxiety due to the immediate and severe impacts they are likely to face. This anxiety fuels both peaceful and radical activism as conventional strategies seem insufficient. The study examines climate anxiety across different demographics, revealing how proximity to climate change impacts influences attitudes towards radical and incremental solutions. The Verasight panel results show that younger respondents are more likely to support radical changes to address climate change compared to older individuals with decreasing effects of climate anxiety as age increases. Immigration status, regionality, and political ideology also play a role, with naturalized citizens, current immigrants, first-generation Americans, residents of the Midwest and West, and those with radical political ideologies more likely to support drastic measures. The study highlights the greater inclination for radical climate action among younger individuals and those more directly impacted by climate change. It concludes by emphasizing the importance of understanding and addressing climate anxiety to prevent radicalization and promote effective climate policies, urging further research to engage younger populations and explore the psychological effects of climate change exposure.	

Presentation Title	Are Evangelicals a monolith? Exploring attitudes to immigration among Hispanic, Black and white Evangelicals
Authors	Angela Eichhorst (NORC), David Sterrett (NORC)
Presenter Email	eichhorst-angela@norc.org
Evangelicals make up about a quarter of the United States public, including about 40% of the Republican Party, and are a powerful force in American politics (Suomala 2017). They were one of the most impactful voting blocs in delivering the 2016 election to Trump (Wong 2018), who emphasized policy points popular among Evangelicals. Immigration reform is a priority for many Evangelicals, as well as many others in the US electorate, and Evangelicals are more likely than non-Evangelicals to say immigrants hurt rather than help the country. At a time when the Evangelical electorate is becoming increasingly diverse, we explore immigration attitudes among different racial subgroups. We used AP-NORC's Immigration Attitudes and Conspiratorial Thinkers poll to analyze how Hispanic, Black, and white Evangelicals differ in their attitudes towards the provision of public assistance to migrants in the US, the deportation of migrants, and the prioritization of immigration policy at the federal level. The poll was conducted using probability and non-probability sample sources featuring online and telephone interviews with 4,173 adults, including 248 Hispanic, 160 Black and 737 white Evangelicals. Our research also pulls immigration attitude data from the 2022 AP VoteCast survey, a 120,896 interview survey conducted the week leading up to the midterm elections on phone and web using probability and non-probability sources. The VoteCast dataset supports the robustness of our findings with data from 308 Hispanic, 180 Black and 1,433 white Evangelicals.	