DC-AAPOR/WSS 2019 Summer Review-Preview Conference

Program

July 12, 2019

Bureau of Labor Statistics Conference Center

2 Massachusetts Avenue, NE, Washington DC

Thank you to our conference sponsors!

Platinum Sponsor

Gold Sponsor

Silver Sponsor

PROGRAM

8:30-9:00	Registration	
9:00-9:15	Welcome & Opening Remarks	
9:15-10:30	Session 1A: Weighting & Imputation Methods for Probability and Nonprobability Samples	Session 1B: Cognitive testing and questionnaire development
	A Modeling Approach to Compensate for Nonresponse and Selection Bias in Surveys –Tien-Huan Lin An Imputation Solution for Differentiating between Unreported Attitudes and Genuine Nonattitudes in Survey Data— Jeff Gill Estimation Methods for Combining Probability and Nonprobability Samples—Michael Yang Approaches for Measuring Bias and Variance Components when Combining Probability and Non-probability Samples—Nadarajasundaram Ganesh Lock Sampling, or: Yes, Panels are Different - Now What?— Jake Soffronoff	Record Keeping Practices, Data Quality and Perceived Burden: Results from a Cognitive Interview Study Evaluating the National Study of Long-Term Care Providers— Meredith Massey The Role of Respondent Experience in Answering Survey Questions on Opioid-Related Impairment— Stephanie Willson Multinational cognitive interviewing project evaluating UNICEF questionnaire— Jonathan Vickers The Effect of Socio-economic Status on the Think-Aloud Quality in Children— Paul Schroeder Unpacking the Role of the Interview Guide in the Research Conversation— Casey Langer Tesfaye
10:30-10:45	Coffee Break	
10:45-12:00	Session 2A: Measurement Errors in Survey Research	Session 2B: Response Rates and Potential Nonresponse Bias
	Rethinking the classic social trust question wording— Anna Brown The Effect of Mode of Data Collection on Mental Health Measurement— Adena Galinsky A new scale for measuring tolerance— Kelsey Jo Starr	What Does Extra Effort Yield in the Current Telephone Survey Climate?— Sarah Dipko Exploring the Characteristics of Partial Interview Respondents in the Consumer Expenditure Survey— Laura Erhard

	The implications of sample-based versus self-reported measures of urbanicity—Alexandra Castillo	Increasing Representativeness Through the Use of Predictive Modeling and Targeted Outreach—Amy Djangali Comparing Estimates of Newsroom Employees in U.S. Government and Private-Sector Surveys— Elizabeth Grieco
12:00-1:30	Lunch on Your Own	
1:30-2:45	Session 3A: Innovative Methods of Survey Research	Session 3B: Redesign Federal and National Surveys (1)
	A Multiple Method Approach to Testing a Complex Web-based Establishment Survey Instrument—Aryn Hernandez Completion Rate Analysis: Community College Survey— Kim Dorazio Identifying Interviewer Falsification using Speech Recognition: A Proof of Concept Study—Hanyu Sun Constructing better coverage intervals for estimators computed from a complex sample survey— Phillip Kott	The NHIS Redesign: Adapting an Ongoing Survey to Changing Times—Sarah Lessem Overview of the 2016-2025 National Health Interview Survey Sample Design—Chris Moriarity Exploring the Impact of Outlet Questions on Data Quality and Respondent Burden-- Yezzi Lee Improving Data Quality in the Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS)—Peter Einaudi
2:45-3:00	Dessert Reception	
3:00-4:15	Session 4A: Moving Mountains with Social Marketing: Survey Findings, Focus Groups, and Audience Segmentation Behind the 2020 Census Communication Campaign	Session 4B: Redesign Federal and National Surveys (2)
	2020 Census Barriers, Attitudes, and Motivators Study Survey Results: Knowledge Gaps, Privacy Concerns, Fear of Repercussions, and Motivating Messages— Kyley McGeeney Empirical Evidence to Understand the 2020 Census Citizenship Controversy--Gina Walejko	Reinventing the Messaging Strategy in the American Community Survey Mail Contact Materials— Jonathan Schreiner Using Eye-Tracking to Evaluate New American Community Survey Mail Materials Design Strategies— Alfred Tuttle

	Hope, Fear, and Political Efficacy: Exploring 2020 Census Participation Motivators and Barriers through Focus Groups with Non-English Speakers, Puerto Ricans, Small Race and Ethnic Groups, and other Audiences— Sara Evans Mindsets and Segmentation: Promoting 2020 Census self-response— Shawnna Mullenax Making Data Count: Research and Analytics Applications to the 2020 Census Integrated Communications Campaign— Yazmín Argentina García Trejo	Creating a Redesigned Questionnaire for the CE Survey Using Colectica— Brett McBride Developing a Standard Measurement of Housing Insecurity in Surveys— Jessica Graber
4:15-4:30	Closing Remarks	
5:00	Happy Hour– Union Pub 201 Massachusetts Ave, NE	

ABSTRACTS—CONCURRENT SESSIONS

Session 1A: Weighting & Imputation Methods for Probability and Nonprobability Samples

Presentation Title	A Modeling Approach to Compensate for Nonresponse and Selection Bias in Surveys
Authors	Tien-Huan Lin (Westat), Ismael Flores Cervantes (Westat)
Presenter Email	amylin@westat.com

In surveys, errors such as selection bias, nonresponse, or noncoverage are all potential causes of biased estimates. This paper focuses on selection bias, which could be self-inflicted due to erroneous sample selection or could occur as missing not at random (MNAR) nonresponse. As examples, tobacco use surveys may be subject to selection bias since young males who are more prone to tobacco use are also less likely to participate; and surveys of domestic violence with an unbalanced sample of older females could induce biased results since the prevalence is highly correlated with age and gender. The common approach of mitigating bias using weighting adjustments justified by models for response propensity may increase the variance of weighted estimates. This paper examines empirically the bias and variance via gradient boosting, a popular statistical learning method, which develops weighting adjustments taking into account the correlation between survey outcomes and response propensity. Simulations are used to study the impact on bias and variance in three settings: 1) missing at random; 2) MNAR with partial model specified; and 3) MNAR with selection bias and partial model specified.

Presentation Title	An Imputation Solution for Differentiating between Unreported Attitudes and Genuine Nonattitudes in Survey Data
Authors	Jeff Gill (American University), Natalie Jackson (PRRI)
Presenter Email	jgill@american.edu

Most survey analyses treat “don’t know” or nonattitude responses as missing values and drop them from analysis with case wise (list wise) deletion. To date, considerable research has been devoted to minimizing such responses, so that missing data are minimized. There are two problems with this approach: (1) We know that case wise deletion is the wrong way to deal with unrecorded data unless it is missing completely at random (not conditional on other data, observed or unobserved). Otherwise, statistical principles dictate that we should use some form of imputation. Imputation, though, implies that these respondents actually have attitudes on the questions but have declined to state them, leading to the second issue: (2) We do not know whether non-substantive responses are true nonattitudes or the respondent is choosing not to reveal an existing attitude. In this work we demonstrate first that nonattitudes and “don’t know” responses are not random, but rather come from a distinct group of survey respondents. This is shown by modeling relevant missingness as a dichotomous outcome variable explained by various characteristics, including demographic attributes, other attitudinal questions, and group level contexts. This model allows us to produce an imputational model to predict missingness due to ignorance versus intransigence. We use these "data" as part of the survey analysis, using the appropriate statistical

treatment of the coefficient variability, to produce estimates that are not plagued by case wise deletion or fictitious attitudes generated by imputation. Our results demonstrate that this approach is useful for a wide range of survey research, including pre-election polls and non-political surveys.

Presentation Title	Estimation Methods for Combining Probability and Nonprobability Samples
Authors	Michael Yang (NORC at the University of Chicago)
Presenter Email	yang-michael@norc.org
There has been growing demand for methods to incorporate nonprobability samples in survey estimation to improve cost efficiency and timeliness of data dissemination. Researchers have proposed a range of estimation approaches involving nonprobability samples. Our literature review identified five general approaches: (1) Calibration—calibrate total estimates to known control totals; (2) Statistical Matching—statistically match nonprobability and probability samples; (3) Superpopulation Modeling—use a superpopulation model to derive estimates; (4) Propensity Weighting—model the propensity of inclusion in a nonprobability sample, and (5) Small Area Modeling—a small area estimation approach developed in-house at NORC (Ganesh et al., 2017). This paper presents our evaluations of these methods based on two studies: 1) the Food Allergy Survey, measuring adult and child food allergy prevalence, that NORC conducted on behalf of Northwestern University; and 2) an internal NORC AmeriSpeak study, measuring a wide range of topics such as happiness, health status, health insurance coverage, economic wellbeing, political orientation, voting behavior, civic engagement, and more. Both studies are based on probability samples selected from NORC's AmeriSpeak Panel and nonprobability samples selected from SSI's opt-in panel. The AmeriSpeak® Panel is a multistage probability sample selected from NORC's National Frame that represents the U.S. household population. Preliminary analysis shows that the different estimation approaches produce different pseudo weights but comparable point estimates based on the nonprobability sample (Yang et al., 2018). We extend the earlier analysis to compare composite estimates (point estimates, bias, variances) derived from the different estimation approaches to combining probability and nonprobability samples. Conclusions from this research may be used to guide survey estimation practice and motivate future investigations in this important area.	

Presentation Title	Approaches for Measuring Bias and Variance Components when Combining Probability and Non-probability Samples
Authors	Nadarajasundaram Ganesh (NORC at the University of Chicago), Edward Mulrow (NORC at the University of Chicago), Vicki Pineau (NORC at the University of Chicago), Michael Yang (NORC at the University of Chicago)
Presenter Email	nada-ganesh@norc.org
Probability sampling has been the standard basis for design-based inference from a sample to a target population. In the era of big data and increasing data collection costs, however, there has been growing demand for methods to combine data from probability and	

nonprobability samples in order to improve the cost efficiency of survey estimation without loss of statistical accuracy. In a prior presentation, we discussed the use of small area estimation models to generate unbiased estimates and to estimate the bias associated with a non-probability sample assuming the smaller probability sample yields unbiased estimates. In this presentation, we discuss methods to estimate the variance associated with such unbiased small area estimates. Furthermore, we consider a class of biased small area estimators that could potentially result in estimates with smaller total survey error compared to the previously described unbiased small area estimators. We investigate the properties of our estimators using a survey of adults 18+ years and a simulation study.

Presentation Title	Lock Sampling, or: Yes, Panels are Different - Now What?
Authors	Jake Soffronoff (USPS OIG)
Presenter Email	Jsoffron@gmail.com

Most everyone in the survey research world is aware that the use of nonprobability online panels is different than using probability-based methods. While many have attempted to correct for those differences using after-the-fact methods such as data weighting or raking, fewer have approached the problem as an issue to be actively dealt with through the sample design and data collection processes. This presentation details the "Lock Sampling" methodology that Jake Soffronoff has developed for fielding survey projects via nonprobability online panels. “Lock Sampling” involves several components: - The use of quota sampling designs that are specifically built to counter both the demographic and behavioral biases present in online panels. - The inclusion of multiple data quality checks to ensure that the data being collected is high quality, and is not being provided by respondents that are simply trying to qualify (and be paid) for completing surveys as “hard to reach.” - The active management of data collection to accommodate for – and ensure the collection of – responses from slow-responding, underrepresented groups. Like the locks in a canal, here the process of data collection includes multiple phases during field where fast-responding groups are held back until slow-responding, underrepresented groups have responded proportionately. Once those slow-responding, underrepresented groups have filled adequately, the sample flow is reopened with the aim of achieving overall quota sample balance. This process is repeated until the full sample is achieved. - Weighting of the cleaned data collected on both demographic and behavioral quotas. While employing the lock sampling method requires considerable attention and effort, the data achieved through the approach have repeatedly triangulated well against data collected through probability methods (published data comparing "lock sampled" results and data collected through probability telephone methods will be included in the presentation).

Session 1B: Cognitive testing and questionnaire development

Presentation Title	Record Keeping Practices, Data Quality and Perceived Burden: Results from a Cognitive Interview Study Evaluating the National Study of Long-Term Care Providers
Authors	Meredith Massey (National Center for Health Statistics), Lauren Harris-Koteijn (National Center for Health Statistics), Manisha Sengupta (National Center for Health Statistics)
Presenter Email	wnx6@cdc.gov

The National Study of Long-Term Care Providers (NSLTCP) is a biennial study that monitors trends in the supply, provision, and use of the major sectors of paid, regulated long-term care services. NSLTCP uses survey data on the residential care community and adult day services sectors, and administrative data on the home health, nursing home, and hospice sectors. For the adult day services center (ADSC) and residential care community (RCC) components, LTCSB has been conducting mixed-mode surveys to collect information about characteristics of the services providers and users. The services user questions are at the aggregate-level (individual services users not sampled) and are designed to collect information on the percentage of services users with a characteristic of interest, at the ADSC or RCC level. When responding to establishment surveys such as the NSLTCP, respondents are often asked to provide information that is maintained in the form of administrative records. However, there is often no standard for what information is maintained, how it is maintained or who has access to it. This presentation will explore the impact of record keeping practices on both data quality and response burden using the results from a cognitive interview study evaluating NSLTCP. The overall purpose of this project was to 1) determine whether there are data quality (measurement and response generation) challenges with the aggregate-level services users surveys questions and 2) investigate provider perceptions of the burden in looking up the records, and the impact of the perceived burden on data quality. Results from this study will be shared as well as how this can inform future study design and questionnaire development through a better understanding of record keeping practices among providers.

Presentation Title	The Role of Respondent Experience in Answering Survey Questions on Opioid-Related Impairment
Authors	Stephanie Willson (National Center for Health Statistics)
Presenter Email	swillson@cdc.gov

This paper explores the idea that survey questions produce better construct validity when respondent experiences are incorporated into the concept being measured. Findings are based on a cognitive interview study on opioid use conducted by the Coordinating Center for Questionnaire Design and Evaluation Research. The project consisted of 140 cognitive interviews completed across seven different geographical regions of the United States, including Washington, DC/Maryland, Kentucky, Alabama, Washington State, Massachusetts, Kansas, and Texas. This paper will focus specifically on side effects as a question designed to capture different aspects of impairment and how respondents misreported on that question. Respondents did not always understand the concept of side effects as intended. The nature

of their experiences with opioid use shaped their interpretation of the question, and these experiences did not always match the medical perspective embedded in the question. For example, side effects are medically understood as secondary to the intended therapeutic effect of a drug. Moreover, while some side effects can be beneficial, most are considered unpleasant or harmful. Respondents often did not understand their experiences with opioids in this manner. For instance, respondents who used opioids to manage chronic pain often understood the question on side effects through the lens of pain relief. As a result, the "side effect" they reported had more to do with how they felt in the absence of pain than it did with any adverse effects of the drug. As a result, the answer they provided was based more on the therapeutic effect of the drug and not its side effects. Respondents who were long-term, dependent users thought about adverse "side effects" as something they experienced when NOT taking the opioid. The idea that an adverse side effect could occur from taking opioids was not something they considered. These and other patterns will be discussed.

Presentation Title	Multinational cognitive interviewing project evaluating UNICEF questionnaire
Authors	Jonathan Vickers (National Center for Health Statistics), Kristen Miller (National Center for Health Statistics)
Presenter Email	nnx9@cdc.gov

In this case study, National Center for Health Statistics (NCHS) researchers document their recent experiences evaluating survey questions for the forthcoming UNICEF Module for Inclusive Education in Cambodia and Kazakhstan. The module is being designed to compare the barriers to inclusive education for children with disabilities across nations. UNICEF has enlisted the help of NCHS Collaborating Center for Question Design and Evaluation Research (CCQDER) to test and evaluate the questions with a diversity of respondents in different nations. The CCQDER uses cognitive interviews to assess the validity of survey questions. Cognitive interviews produce in-depth qualitative summaries, or question summaries, on respondents' interpretations of each question on the survey instrument. Researchers then analyze the question summaries to determine if the survey questions capture the intended item being measured. Data from the first round of testing in Cambodia demonstrated that respondents, particularly those with limited educations, struggled with questions from two domains: (1) attitudes towards education for all children and (2) barriers to education for out of school children. UNICEF and CCQDER revised the questions and held a second round of testing in Kazakhstan. The data from Kazakhstan suggested that the revisions had improved questions on attitudes towards education. However, Kazakhstani's interpretations of what it meant to be "in school" wrongly screened respondents into the section intended for respondents with children not receiving formal educations. As a result, CCDQER did not collect useful cognitive interviewing data on the revised "barriers to education" questions. Lessons learned from the analyses, strategies, and logistics of these two rounds of cognitive interviewing will guide the next rounds of testing.

Presentation Title	The Effect of Socio-economic Status on the Think-Aloud Quality in Children
Authors	Paul Schroeder (EurekaFacts, LLC), Mila Sugovic (EurekaFacts, LLC), Anh-Thu Ton (EurekaFacts, LLC), Cecilia Teal (EurekaFacts, LLC), Michael Plotkin (EurekaFacts, LLC)
Presenter Email	schroederp@eurekafacts.com

Research suggests that young children typically find the think-aloud method to be difficult (Someren et al, 1994); additionally, it is unclear whether the quality of verbal reporting produced by children also differs across demographic variables, such as SES. This study examines the application of a think-aloud technique in survey-item testing across students in grade 4 (age 9-10). We compared the quality of the think-aloud produced by students of different SES groups (Low and High) by evaluating verbal report length, relevance of the verbal report, and the number of reported problems identified. Results suggest that the quality of information gathered from the think-aloud interviewing method is not affected by SES among grade 4 students.

Presentation Title	Unpacking the Role of the Interview Guide in the Research Conversation
Authors	Casey Langer Tesfaye (Research Support Services)
Presenter Email	casey@researchsupportservices.com

Qualitative interviewers conducting In-depth interviews (IDI's) have a discursive choice regarding the role of the Interview guide within the research conversation. Some interviews, although guided by an interview guide or written set of questions, take place without any explicit references to the interview guide or the questions themselves, and other interviews make explicit mention of the interview guide or even use the presence of the scripted questions as a frequent tool to facilitate the type of responses that would be most helpful for a respondent to provide. This begs the question, "Is it good practice to bring the interview guide into the forefront by making it present in the discussion, or should that be avoided?" To address this question in a data-driven way, we conducted a turn by turn secondary data analysis of a set of three transcripts from IDI's that were intended to be up to 60 minutes long. The transcripts were chosen because the interviewers mention the interview guide to varying degrees, from occasional mentions of the guide to using the guides as an integral part of the conversation. We examined the varying roles that the interview guide and scripted questions play within these research conversations. We evaluated mentions of the guides by function; for example, rapport building, co-creating responses that are purportedly more useful for analysis, clarification of unclear responses, and maintaining politeness when follow-up questions may be undesirable. By closely examining the role that the guide plays within the research conversation, we are able to provide practical, data-driven advice for interviewer training regarding the mention of interview guides during IDI's.

Session 2A: Measurement Errors in Survey Research

Presentation Title	Rethinking the classic social trust question wording
Authors	Anna Brown (Pew Research Center), Juliana Horowitz (Pew Research Center)
Presenter Email	abrown@pewresearch.org
The question, "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?" has been used to measure social trust by many U.S. and international survey research organizations for at least the last 40 years. The question consistently finds that a majority of U.S. adults say "you can't be too careful in dealing with people" in recent decades. However, recent OECD guidelines call into the question the validity of this wording, pointing out that it measures two separate concepts – trust and caution. Using Pew Research Center's probability-based online American Trends Panel, we tested the traditional question wording against four alternative measures of trust: response options of "most people can be trusted" vs. "most people cannot be trusted;" the same response options but with "…to do the right thing" added at the end; an 11-point scale of how trustworthy most people are; and a four-point scale of how trusting the respondent says they themselves are. Using each of these alternative measures, we find Americans to be far more trusting than the traditional formulation suggests. While this is the case across demographic groups, the discrepancy is particularly notable among women, who appear to be far less trusting than men when asked the question with the "can't be too careful" wording but are as trusting as men in each of the alternative measures. This suggests that, while women may be more cautious than men, they are not necessarily less trusting. The results of this question wording experiment have important implications for researchers who are interested in understanding the level of social trust among Americans (and publics around the world) and how it varies across demographic groups.	

Presentation Title	The Effect of Mode of Data Collection on Mental Health Measurement
Authors	Adena Galinsky (National Center for Health Statistics), Ben Zablotsky (National Center for Health Statistics), James Dahlhamer (National Center for Health Statistics), Aaron Maitland (National Center for Health Statistics), Catherine Simile (National Center for Health Statistics), Hee-Choon Shin (National Center for Health Statistics)
Presenter Email	wpm0@cdc.gov
Stigma attached to mental health makes it potentially sensitive to ask about in surveys. Modes of data collection that require an interviewer to administer questions may elicit socially desirable answers that hide the true nature of a respondent's mental state. Indeed even within interviewer-administered modes of data collection there may be differences in how respondents answer questions about their own mental health depending on the social presence of the interviewer, which varies by mode of data collection. For example, interviewers have a higher social presence in face-to-face interviews compared to telephone	

interviews. This paper examines responses to six survey questions on psychological functioning including whether respondents experienced feelings of sadness, nervousness, restlessness, hopelessness, worthlessness, or found everything to be an effort. These six survey questions can be used together to create the K6 index of psychological distress (Kessler, 2003) and have been included annually on the National Health Interview Survey (NHIS) for several years. The NHIS conducts interviews face-to-face using computer-assisted personal-interviewing (CAPI) with telephone follow-up of nonrespondents. We first analyzed NHIS data to determine if serious psychological distress differs by the mode in which the respondents completed the NHIS. This analysis found different levels of serious psychological distress between modes that largely disappear after controlling demographic differences due to self-selection into the two modes. Next, we analyzed data from two experiments with one comparing telephone data collection (CATI) to face-to-face data collection and the other experiment comparing audio computer assisted self-interviewing (ACASI) to face-to-face data collection. We found significantly higher levels of serious psychological distress in both CATI and ACASI compared to CAPI, although sometimes this is again due to self-selection into the different modes. We discuss the implications for measuring psychological distress and other mental health topics in interviewer-administered surveys.

Presentation Title	A new scale for measuring tolerance
Authors	Kelsey Jo Starr (Pew Research Center), Jonathan Evans (Pew Research Center), Neha Sahgal (Pew Research Center), Ariana Salazar (Pew Research Center), Omkar Joshi (University of Maryland)
Presenter Email	kstarr@pewresearch.org

With the recent rise in populism and nationalism around the world, interest in measuring public opinions on these topics has also risen. While specific questions measuring specific attitudes have been widely used, few analyses have aggregated a variety of measures to understand broad levels of tolerance across countries. Using data from a recent survey on religion and public life in 15 Western European countries, Pew Research Center conducted factor analyses on a series of questions that looked at views on nationalism, immigration and religious minorities. Twenty-two of these questions were then combined into a weighted scale to determine overall levels of tolerance. Higher scores on the scale indicate a respondent agrees with a greater number of intolerant viewpoints. Logistic regression was then used to better understand if specific characteristics are more associated with higher scores, such as religious identity, age, political leanings or gender. This scale proved better than using individual survey variables to indicate tolerance and to understand the factors correlated with tolerance.

Presentation Title	The implications of sample-based versus self-reported measures of urbanicity
Authors	Alexandra Castillo (Pew Research Center), Kat Devlin (Pew Research Center), Janell Fetterolf (Pew Research Center), and Courtney Johnson (Pew Research Center)
Presenter Email	acastillo@pewresearch.org

Scholars have used urbanicity to explain meaningful differences in public opinion data, from predicting ideological affiliations to attitudes on divisive issues. While urbanicity provides a layer of nuance to our understanding of public opinion, it warrants further methodological exploration in domestic and international surveys alike. When measuring urbanicity, researchers often choose from a variety of geographical measures. Another solution is to ask respondents to describe the area where they live, which may reveal similar substantive patterns. A 2018 Pew Research Center survey included a degree of urbanization indicator and a self-reported measure of urbanity in four face-to-face countries–Greece, Hungary, Italy and Poland. The degree of urbanization indicator was populated from the Eurostat geographic variable DEGURBA, while respondents chose an appropriate descriptor for the area where they live for the self-reported measure. The presence of both variables in European, face-to-face surveys allows for Pew Research Center to analyze the viability of the self-reported measure on an international stage. This project explores the following issues: How does self-reported urbanity correspond with the degree of urbanization? What explains discontinuity between self-reported and sample-based measures of urbanity? Does self-reported urbanity yield expected demographic differences? Across a selection of indicators, does self-reported urbanity produce similar results as the degree of urbanization? This project finds that the self-reported measure of urbanity strongly correlates with the Eurostat degree of urbanization measure in some countries, but there are significant mismatches, particularly between self-reported suburban and urban respondents. Despite this, both measures yield similar findings across demographic variables, such as respondents with higher education being more likely to reside in urban areas. Similarly, the self-reported measure of urbanicity largely replicates the findings from the degree of urbanization measure across a selection of substantive indicators, suggesting that it is a viable substitute for survey-based location measures in other contexts.

Session 2B: Response Rates and Potential Nonresponse Bias

Presentation Title	What Does Extra Effort Yield in the Current Telephone Survey Climate?
Authors	Sarah Dipko (Westat)
Presenter Email	sarahdipko@westat.com
In the current telephone survey climate, making contact with and interviewing sampled cases by telephone presents numerous challenges. The primary challenge is to get sample members to answer the telephone, as high non-contact rates plague both random digit dial and list sample designs. Once that barrier is crossed, the next challenge is to gain cooperation with sample members. In addition, when calling a list sample one faces the prospect of tracing – this may or may not be warranted depending on the degree of mobility of sample members. For telephone centers that use calling algorithms, one approach to constraining effort is to set a maximum number of calls before finalizing cases as nonresponse. Some systems permit these maximums to be exceeded in order to increase the contact rate, depending on sample performance and available study resources. In addition, it is common to use refusal conversion to attempt to gain cooperation with those who initially refuse to participate. This paper addresses the effect of using these techniques, as well as tracing, for a recent multi-mode survey of SNAP participants. The research questions addressed include the following. • What does extra effort yield in terms of shifting the demographics of survey participants – does it yield more of the same types of respondents interviewed without extra effort, or shift the distribution? • What percentage of overall study completes is yielded by employing extra effort? • Are the same effects observed for both English-speaking and Spanish-speaking sample cases, or are the dynamics different?	

Presentation Title	Exploring the Characteristics of Partial Interview Respondents in the Consumer Expenditure Survey
Authors	Laura Erhard (Bureau of Labor Statistics)
Presenter Email	erhard.laura@bls.gov
One of the primary purposes of the Consumer Expenditure Surveys (CE) is to provide data to calculate the relative importance of goods and services in the market basket of the Consumer Price Index (CPI). In order to have a complete picture of spending for an average household, we require respondents in the CE Interview Survey to be asked all expenditure questions. Any households that do not complete all expenditure questions are considered non-interviews and data collected before the break-off point is not used in CE's processing published estimates, or in the data provided to CPI. To date there has been limited exploration of these partial interviews, in part due to difficulty in distinguishing these interviews from non-contacts and complete refusals. With a recent project to make accessible audit trails from the interview survey that capture data on interview and question timing, among other valuable paradata, the task of identifying and exploring partial interviews is more easily achieved. This research sets out to identify partial interviews using audit trail data to find the furthest question in the CE Interview survey that the respondent reached. After identifying partial interviews, the author will explore the characteristics of	

respondents that break off and the quality of data collected prior to break off, among other traits of a CE partial interview. This will be the first step in identifying potential uses for the captured data from partial interviews that could include use in imputation processes, nonresponse bias studies, or exploring the use of completed sections alongside imputed sections in survey estimates (similar to a split-questionnaire design). This presentation will be of interest to survey methodologists, data producers, and survey statisticians.

Presentation Title	Increasing Representativeness Through the Use of Predictive Modeling and Targeted Outreach
Authors	Amy Djangali (IMPAQ International), Jacob Joseph-David (IMPAQ International), Lily Trofimovich (IMPAQ International)
Presenter Email	adjangali@impaqint.com

IMPAQ has conducted The Kidney Disease Quality of Life Survey (KDQOL) since 2016 using a sequential mixed mode design (mail with telephone follow-up) to collect quality of life data from end-stage renal disease (ESRD) Medicare beneficiaries. Since the first year of data collection in 2016, we have continually analyzed our data in order to identify patterns of nonresponse within this medically fragile population. In 2017, we implemented methodological changes in order to increase representation from low responding subgroups including Hispanics, African-Americans, and 18-54 year olds. The protocol changes included: • Spanish Language targeting: This includes the creation of a predictive model to identify respondents who were most likely to speak Spanish and provided English and Spanish materials in the same envelope. • Mode preference: By analyzing 2016 data, we identified groups that were less likely to respond to the mail survey and those who would require additional telephone follow-up efforts (i.e. African Americans, Hispanics, and 18-54 year olds). We began telephone follow-up with these groups two weeks earlier than previous survey administrations. As a result of these changes, we saw a 4% response rate increase in 2017 with larger response rate gains for African Americans and Hispanics (5% and 8% respectively). We implemented the same protocols again for 2018 and continued to see response rate increase over the initial survey administration (42% vs. 38%). We also saw additional response rate gains compared to 2017 for Hispanics (3% increase) and 18-54 year olds (12% increase). Our presentation will pull together IMPAQ's findings from our previous AAPOR presentations in order to review on how we identified these patterns of nonresponse as well as provide information on how we implemented these methodological changes in order to increase representation within this Medicare population.

Presentation Title	Comparing Estimates of Newsroom Employees in U.S. Government and Private-Sector Surveys
Authors	Elizabeth Grieco (Pew Research Center), Mason Walker (Pew Research Center)
Presenter Email	egrieco@pewresearch.org

There are four surveys that produce estimates of newsroom employees in the United States, two private-sector and two government surveys. This poster will assess the comparability of estimates of the size and characteristics of news workers derived from these four surveys. Media researchers have traditionally relied on survey data produced by two professional organizations for estimates of newsroom employees: the American Society of News Editors' (ASNE) Newsroom Census and Employment Diversity Survey and the Radio Television Digital News Association's (RTDNA) Staffing and Women and Minorities Surveys. These surveys are designed to provide detailed information on specific media sectors – ASNE for newspaper and RTNDA for television and radio. Two government surveys, while less-used by media researchers, include industry-wide as well as sector-specific estimates. The Bureau of Labor Statistics Occupational Employment Statistics (OES) survey produces estimates of the overall number of newsroom employees as well as by media sector (newspaper, radio, television, cable, and digital), but includes no demographics. The U.S. Census Bureau's American Community Survey (ACS) produces demographic and social characteristics estimates of newsroom employees, but its estimates of the number of news workers by sector are limited. All of these sources provide valuable information about newsroom employees, but no research has evaluated the comparability of the estimates derived from them. This poster compares estimates of the size, sex, and race/ethnicity of residence of newsroom employees among these four data sources. As the analysis shows, estimates from government data sources generally track those from private-sector sources, with differences generally explained by differences in data collection methodology. This suggests that, as declining response rates push the cost of conducting surveys beyond the reach of many private organizations, government surveys may be an alternative source for some – but not all – of the information on newsroom employees traditionally provided by private-sector surveys.

Session 3A: Innovative Methods of Survey Research

Presentation Title	A Multiple Method Approach to Testing a Complex Web-based Establishment Survey Instrument
Authors	Aryn Hernandez (U.S. Census Bureau), Temika Holland (U.S. Census Bureau), Rebecca Keegan (U.S. Census Bureau), Amy Anderson Riemer (U.S. Census Bureau)
Presenter Email	aryn.hernandez@census.gov

Every five years the U.S. Census Bureau conducts an Economic Census (EC). This mandatory, self-administered establishment-based survey collects a wide range of financial and production data, such as payroll, employment, revenue by product type, and expenses. Between 2012 and 2017, the survey instrument for the EC was radically transformed from a downloaded software application to an entirely web-based platform. This new instrument was designed to meet the needs of widest variety of companies from small sole proprietorships with only one location to businesses with tens of thousands locations nationwide. The respondent-facing portion of the instrument is composed of two sites, the Respondent Portal and the survey instrument known as Centurion, connected by a program that provides a seamless link. The Portal delivers options to respondents that were previously only available through telephone contact with survey clerks. It also offers respondents the ability to share the survey with others within their business, a useful feature for business surveys, where data are often gathered from multiple sources. In addition, the Centurion instrument offers multiple reporting methods, easy-to-use review features, and 'how-to' videos. To facilitate the leap to this complex, web-based instrument, researchers conducted multiple rounds of usability testing and respondent debriefings. The usability testing consisted of a mix of traditional methods, such as having respondents perform tasks while thinking aloud, and more modern methods, such as screen recordings. Additionally, the 'how-to' videos underwent separate cognitive testing. Respondent debriefings, conducted via phone and internet, provided insight into remaining issues experienced during actual reporting. Finally, researchers are utilizing paradata to identify features that performed successfully and those that may benefit from further testing. This presentation will highlight innovative interface design and system functionality, as well as selected methodology and findings from the various rounds of testing.

Presentation Title	Completion Rate Analysis: Community College Survey
Authors	Kim Dorazio (M. Davis and Company, Inc.)
Presenter Email	kim@mdavisco.com

This study uses data from a longitudinal study that was conducted with Houston Community College students. The students, though, do not come from the traditional 18-22 year old demographic and half of the sample is food insecure. Our efforts to minimize non-response included a mixed mode approach that included texting and phone follow up. Results of these interventions are presented.

Presentation Title	Identifying Interviewer Falsification using Speech Recognition: A Proof of Concept Study
Authors	Hanyu Sun (Westat), Gonzalo Rivero (Westat), Ting Yan (Westat)
Presenter Email	hanyusun@westat.com

Survey management staff have long used Computer Audio-Recorded Interviewing (CARI) to monitor field interviewers' work, whether for assessing interviewer performance, validating interviews, or evaluating survey questions. CARI allows studies to monitor interviewer performance and provide feedback to interviewers in real time during the field data collection period. The success of such evaluation, however, often depends on labor-intensive coding in a timely manner. The coders first listen to the conversations between the interviewer and the respondent, and then provide their assessment of interviewer performance based on some coding items, such as if the interviewer falsified the case, and if the interviewer follow the standardized interviewing techniques. Due to resource constraints, often a small number of items in the questionnaire or a small portion of the interview will be coded. In recent years, there is a blooming interest in the survey field to explore the use of machine learning on different aspects of the data collection process (e.g. Eck et al. 2018, Thompson, Book, and Tamby, 2018). However, there is little research on how to use of machine learning approach to monitor interviewer performance and to detect falsification. Here we report a proof of concept study that explores the use of speech recognition to detect interviewer falsification. At Westat, we developed an assessment tool that automates the coding and the evaluation process. The tool first transcribes CARI audio recordings into text, and then measures the distance between the transcript and the questionnaire to create a score on how likely the interviewer falsified the interview. The tool also detects the number of conversational partners in the interview to help detect falsification. In the presentation, we will show how the tool works using recorded lab interview with varied features. We will also evaluate how the tool performs at detecting falsification using the lab data.

Presentation Title	Constructing better coverage intervals for estimators computed from a complex sample survey
Authors	Phillip Kott
Presenter Email	pkott@rti.org

Coverage intervals for a parameter estimated from complex survey data are usually constructed by assuming that the parameter estimate has an asymptotically normal distribution, and the measure of the estimator's variance is roughly chi-squared. The size of the sample and the nature of the parameter being estimated render this conventional "Wald" methodology dubious when constructing coverage intervals, especially for proportions. A revised method of coverage-interval construction has been developed in the literature that "speeds up the asymptotics" by incorporating an estimated skewness measure. We will discuss how skewness-adjusted coverage intervals can be computed in some common situations and why it may be inappropriate to call them "confidence intervals."

Session 3B: Redesign Federal and National Surveys (1)

Presentation Title	The NHIS Redesign: Adapting an Ongoing Survey to Changing Times
Authors	Sarah Lessem (National Center for Health Statistics), Ben Zablotsky (National Center for Health Statistics), Aaron Maitland (National Center for Health Statistics), Stephen Blumberg (National Center for Health Statistics)
Presenter Email	slessem@cdc.gov

Redesigning a large-scale survey while it is still in the field is a massive undertaking. The National Health Interview Survey (NHIS) redesign took over 4 years. The primary objectives of the redesign were to lower the burden on households in a difficult survey environment with falling response rates and increase the relevance of the health content in the survey to adapt to a dynamically changing health environment. The NHIS redesign was much more than a revision of the questionnaire to accomplish these objectives. The redesign was embraced as an opportunity to revise all processes and systems that make up one of the federal statistical systems flagship surveys. We will give an overview of the steps we took to redesign the NHIS including lessons learned and how these steps apply to redesigns of other ongoing surveys. The presentation begins with a discussion of the process of recognizing the need for a redesign and prioritizing content while involving multiple stakeholders. Next, we discuss the question development process including identifying new content, writing questions, testing questions, and translating into other languages. Instrument programming involved layers of specification to ensure an adequate baseline for programming. Multiple iterations of system testing that involved review of data input and output ensured that the instrument was working according to specifications. Interviewer training was completely revamped to highlight new features of the redesigned NHIS including the new flow of the interview along with new question content. In addition, interviewer training focused on identifying new and better ways for interviewers to sell the redesigned survey to sample households. We conclude with the development of plans for educating data users on the redesigned NHIS and its associated data releases in the future. We also discuss plans for continually adapting the survey over time.

Presentation Title	Overview of the 2016-2025 National Health Interview Survey Sample Design
Authors	Chris Moriarity (National Center for Health Statistics), Van L. Parsons, (National Center for Health Statistics), Kim Jonas (U.S. Census Bureau)
Presenter Email	cdm7@cdc.gov

A new sample design was implemented for the National Health Interview Survey (NHIS) at the beginning of 2016. The new NHIS sample design contains several new features. One new feature is increased flexibility to implement changes in the sample size and/or sample allocation. Another new feature is a different source of most of the sample addresses, relative to the 1985-2015 survey period. We discuss the new sample design features, and lessons learned during the first several years of the sample design period.

Presentation Title	Exploring the Impact of Outlet Questions on Data Quality and Respondent Burden
Authors	Yezzi Lee (U.S. Bureau of Labor Statistics), Safia Abdirizak (U.S. Bureau of Labor Statistics)
Presenter Email	lee.yezzi@bls.gov

The Consumer Expenditure Surveys (CE) and the Telephone Point-of-Purchase Survey (TPOPS) are two of the main data sources for the Consumer Price Index (CPI). Like many Random Digit Dial surveys, TPOPS is facing higher costs and declining response rates. To address these challenges, the CE program will begin collecting point-of-purchase and mode of purchase questions in 2019. A test of these additions were incorporated into the CE Interview Survey's fourth and final wave sample starting in April 2017. There are potential benefits to combining the surveys; for example, the CPI market basket will have the same target population and the point-of-purchase questions can serve as recall cues for respondents. In addition, having such details of reported items can provide additional information when reviewing anomalous data during the editing process. However, as with any addition to a survey, there are concerns about the effect on respondent burden, and data quality. This presentation will use the 2017 test data to provide a high-level overview of the outlet study. In addition, we will present preliminary results of the impact on expenditure data quality and respondent burden with the addition of outlet questions to the CE Interview Survey. This research makes use of the structured audit trails database to investigate how adding the outlet questions may have affected data quality. This information is of interest to survey managers, survey methodologists, and data analysts.

Presentation Title	Improving Data Quality in the Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS)
Authors	Peter Einaudi (RTI International), Mike Yamaner (National Science Foundation), Jonathan Gordon (RTI International), Stephanie Eckman (RTI International), Herschel Sanders (RTI International)
Presenter Email	peinaudi@rti.org

In the last 10 years, many have advocated the use of administrative data to replace or supplement survey data collection, and the use of electronic data interchange (or EDI, the transfer of forms directly to the survey organization) is often considered a more efficient method of data collection—reducing respondent burden and providing cost savings to the survey organization. GSS is an annual census of U.S. academic institutions granting research-based postsecondary degrees in science, engineering, and health fields. The survey sponsors, the National Science Foundation and the National Institutes of Health, began a major redesign of the survey in 2015 with two goals: 1) collecting more granular data; 2) reducing respondent burden. To accomplish these goals, we redesigned the data collection instrument to rely less on manual data entry and more on uploading of administrative records. This change required respondents to export counts from their administrative systems, format these data according to a template, and upload them. The redesign was piloted in 2016 and

fully implemented in the 2017 cycle. The transition from survey-based to record-based data collection through EDI often is more challenging than expected. Our paper will describe the process undertaken to successfully implement this transition, and the lessons we learned along the way. Results from the pilot and the 2017 full-scale data collection will illustrate the impact of the redesign on data quality. Based on pilot results, respondents reported reduced burden, measured in time needed to gather and upload the data. We also saw reduced missing-item rates. This presentation will be of interest to other researchers wishing to increase the use of administrative data by establishment survey respondents.

Session 4A: Moving Mountains with Social Marketing: Survey Findings, Focus Groups, and Audience Segmentation Behind the 2020 Census Communication Campaign

Presentation Title	2020 Census Barriers, Attitudes, and Motivators Study Survey Results: Knowledge Gaps, Privacy Concerns, Fear of Repercussions, and Motivating Messages
Authors	Kyley McGeeney (PSB), Brian Kriz (PSB), Shawnna Mullenax (PSB), and Monica Vines (U.S. Census Bureau)
Presenter Email	kmcgeeney@ps-b.com

As was done prior to the 2010 Census, the U.S. Census Bureau conducted a survey to understand what might motivate or prevent people from completing the 2020 Census. The results of this survey, conducted as part of the 2020 Census Barriers, Attitudes, and Motivators Study, will inform the creative strategy for the 2020 Census Integrated Partnership and Communications Program. We conducted the survey between February 20 and April 17, 2018 to a sample of 50,000 households across all 50 states and DC. Adults 18 or older were eligible to respond via the internet or mail in either English or Spanish. The 61-question survey covered a range of topics designed to provide insights into the barriers, attitudes, and motivators to participation across several demographic subgroups. Approximately 17,500 people responded, resulting in a weighted response rate (AAPOR Response Rate 3) of 39.4 percent. Findings suggest fewer than seven in ten householders reported they were “extremely likely” or “very likely” to fill out a census form. Gaps in knowledge about the census exist – many people were unfamiliar with the census, and there were misconceptions about its purpose and process. These knowledge gaps, including how census data are used, were uneven across demographic groups. The findings revealed five barriers that might prevent participation: concerns about data privacy and confidentiality, fear of repercussions, distrust in all levels of government, a lack of efficacy (being personally counted does not matter), and a belief that completing the census has no personal benefit. Finally, potential motivators include funding for public services, which ranked highest across groups. We conclude by discussing how specific findings will inform multiple campaign components especially creative development. We will also briefly highlight how the survey and its findings flowed into subsequent foundational research efforts such as the creation of mindsets and audience segments.

Presentation Title	Empirical Evidence to Understand the 2020 Census Citizenship Controversy
Authors	Gina Walejko (U.S. Census Bureau), Nancy Bates (U.S. Census Bureau), Yazmín A. García Trejo (U.S. Census Bureau), Anna Sandoval Girón (U.S. Census Bureau), Brian Kriz (PSB), Kyley McGeeney (PSB)
Presenter Email	gina.k.walejko@census.gov

In support of the 2020 Census communication campaign, the Census Bureau conducted research including a nationally representative survey and focus groups. Both were designed

to measure intent to participate as well as public knowledge, attitudes, and opinions about the census to inform the advertising campaign's creative platform and messaging.
Five weeks into the survey field period and with twelve focus groups completed, the Secretary of Commerce directed that the Census Bureau add a citizenship question to the 2020 Census in response to a request from the Department of Justice. This request resulted in an unprecedented amount of media coverage. The timing also resulted in a natural survey experiment and afforded the opportunity to add a discussion topic of citizenship to the remaining focus groups. Indirect evidence from previous research suggests the addition of this citizenship question will decrease self-response rates to the decennial census (Brown et al. 2018). Focus groups conducted with hard-to-survey groups (e.g. non-English speaking populations) indicate the citizenship question may impede participation among some immigrant-origin audiences. The barrier was highest among individuals who believe the purpose of the question is to find illegal immigrants and that their information will be shared across agencies. We augment focus group results with responses from the survey pre- and post-citizenship question announcement. We examine pre- and post- answer distributions to questions regarding whether the census counts both citizen and non-citizens, whether answers are shared with government agencies, and whether answers can be used to harm respondents. We then report multivariate models predicting intent to respond to the census controlling for date of response, nativity, race, ethnicity and other mediating and control variables.

Presentation Title	Hope, Fear, and Political Efficacy: Exploring 2020 Census Participation Motivators and Barriers through Focus Groups with Non-English Speakers, Puerto Ricans, Small Race and Ethnic Groups, and other Audiences
Authors	Sara Evans (PBS), Jenna Levy (PSB), Anna B. Sandoval Girón (U.S. Census Bureau), Jennifer Miller Gonzalez (PSB)
Presenter Email	sevans@ps-b.com

The 2020 Census Integrated Partnership and Communications program will encourage participation in the decennial census. To inform the design of a communications campaign to motivate self-response to the 2020 Census, 42 focus groups were conducted across 14 states and territories. The goal of this research was to better understand self-response motivators and barriers experienced by minority demographic groups and other audiences prior research suggests are at risk of low self-response to the census and for whom tailored communications will be developed. Focus groups were conducted in four languages: English, Spanish, Vietnamese, and Chinese (Mandarin and Cantonese). Using a systematic thematic analysis process, data were grouped into themes from existing research questions and new themes that emerged from the focus group data. These data suggest significant barriers ultimately revolve around lack of trust in government. First, findings point to strong overall skepticism related to "benefits messaging." Since 2000 the Census Bureau has emphasized messages about the way census data is used to determine and apportion resources. In some cases, however, such messaging runs counter to deep-seated personal experiences suggesting that some communities have not realized the positive changes for which they have hoped and worked. Second, intense anxieties related to providing data to the

government were pervasive and suggest information about government data protection policies will not alone mitigate concerns around privacy and confidentiality. Crucially, however, the data show that encouragement from trusted voices in the community could successfully increase participation among people with the greatest trust-based fears and government disaffection. Additionally, hope for a better future in the midst of challenges and anxieties was a common thread with which all audiences grappled in their own ways. We will discuss both implications for communications strategies as well as broader considerations for increasing participation and representation of harder-to-count and harder-to-reach audiences.

Presentation Title	Mindsets and Segmentation: Promoting 2020 Census self-response
Authors	Laura Kail (PSB), Gina Walejko (U.S. Census Bureau), Brian Kris (PSB), Robert Kulzick (PSB), Laura Kail (PSB), Shawnna Mullenax (PSB), Hubert Shang (PSB)
Presenter Email	smullenax@ps-b.com
In preparation for the 2020 Census, the Census Bureau developed mindsets – clusters of individuals' decennial census knowledge, attitudes, and barriers to participation – from the 2020 Census Barriers, Attitudes, and Motivators Study Survey. The Census Bureau also used k-means clustering to group census tracts into segments with similar response profiles (i.e., mode and propensity) and sociodemographic characteristics. These tract-level segments also differentiate on mindsets and media consumption habits. Mindsets and tract-level segments will inform overall program strategy, planning for advertising and partnership activities, creative development, and campaign optimization. This presentation describes the six mindsets, eight tract-level segments, and the methodology for producing each. The presentation concludes by describing how the communications team will use mindsets and tract-level segments to promote 2020 Census self-response across sociodemographic groups with diverse response patterns, knowledge bases, attitudes towards the census, and media consumption habits.	

Presentation Title	Making Data Count: Research and Analytics Applications to the 2020 Census Integrated Communications Campaign
Authors	Yazmín Argentina García Trejo (U.S. Census Bureau), Sarah Evans (PSB), Anna Sandoval Girón (U.S. Census Bureau)
Presenter Email	yazmin.argen.garcia.trejo@census.gov
In one of the largest social marketing efforts in United States history, the U.S. Census Bureau will devote millions of dollars to the development of creative messages and targeted media purchases with the goal of maximizing self-response during the 2020 Census. Crucial to maximizing the effectiveness and efficiency of the campaign was a foundation of research that included a nationally representative survey, 42 focus groups across the U.S. with hard-to-count populations and other audiences that could not be reached via a survey and for whom tailored communications will be developed, and modeling using Census Bureau and	

third-party data. The focus of this paper is the application of that research to the development of (a) messaging that reflects the beliefs, attitudes, motivators, and barriers of a diverse public and (b) a plan for advertising spending and broader dissemination. Three broad topic areas will be discussed. First, we will discuss best practices for synthesizing data across disparate yet complementary research methodologies. Second, we will highlight considerations for striking an optimal balance between tailored and broad-reaching insights and communications when synthesizing research across diverse audiences. Third, we will share tactics for navigating the needs of diverse research end users. More specifically, we will explore challenges related to designing and executing rigorous public opinion and qualitative research in a research-centric government environment and a communications context. At this intersection of research and communications, research is judged on both its methodological merit and its ability to provide appropriate and actionable insight for practitioners (e.g., advertising/marketing agency partners). To illustrate these discussion areas, example draft messages and creative work from the 2020 Census campaign will be shared to provide attendees a behind-the-scenes look at the evolution of a campaign of this magnitude and to highlight pivotal, research-informed decisions made at various stages of campaign development.

Session 4B: Redesign Federal and National Surveys (2)

Presentation Title	Reinventing the Messaging Strategy in the American Community Survey Mail Contact Materials
Authors	Jonathan Schreiner (U.S. Census Bureau)
Presenter Email	jonathan.schreiner@census.gov

Each year, the U.S. Census Bureau asks 3.5 million households to respond to the American Community Survey (ACS). This presentation will outline how an ongoing, multi-year, research project on strategic messaging has influenced the design of new mail contact materials sent to recruit participant households. In 2017, the Census Bureau published, "Strategic Framework for Messaging in the American Community Survey Mail Materials," which combined research on potential respondent mindsets with a cross-disciplinary review of survey communications literature to outline best practices recommendations for messaging in ACS mail contact materials (letters, brochures, instruction cards, envelopes). This analysis found that messaging in mail contact materials across multiple mailings should feel like a single conversation between the Census Bureau and potential respondents. The report also found that messaging should be simple, focused, and vary across mailings to strategically communicate three things: 1) trust that the survey is legitimate and secure, 2) the benefits of survey participation, and 3) messages that reduce the perception of burden of responding. In 2018, we used the findings from that report to conduct a systematic analysis of the messaging contained in the current ACS contact materials. By tracking the placement and frequency of each message across mail materials, this analysis highlighted both the strengths in the ACS mail contact strategy as well as potential ways to improve the strategic communication of trust, benefits and burden reduction messaging. This presentation will provide data users and survey methodologists an overview of the Census Bureau efforts to incorporate findings from these reports into the design of new mail contact materials. These new materials combine visual design updates, the use of plain-language, and a staged, strategic messaging plan to hold a less repetitive and more impactful conversation with potential respondents across multiple mail pieces.

Presentation Title	Using Eye-Tracking to Evaluate New American Community Survey Mail Materials Design Strategies
Authors	Alfred Tuttle (U.S. Census Bureau), Jon Schreiner (U.S. Census Bureau), Elizabeth Nichols (U.S. Census Bureau), Erica Olmsted-Hawala (U.S. Census Bureau), Rodney Terry (U.S. Census Bureau)
Presenter Email	alfred.d.tuttle@census.gov

The U.S. Census Bureau has undertaken an effort to develop new mail materials for the American Community Survey, using novel design strategies and out-of-the box thinking with regard to government correspondence. Our development of new materials emphasized strategic use of specific types of messages intended to motivate respondents to participate in the survey (e.g., beneficial uses of survey results, burden reduction features, etc.) and a holistic approach that considers the entire sequence of mail contacts in the selective deployment of messages. In addition, we applied visual design strategies to make the

materials appealing and easy to read. Our visual strategies include: color; simple images and icons; minimal, concise text; plain language; and text structures not typically found in government letters. Three treatment groups were developed, each based on a distinctive design strategy. To optimize the mail materials prior to the experiment, we plan to evaluate a representative selection of the materials in an eye-tracking study. Eye-tracking allows researchers to determine what subjects look at when presented with a visual stimulus. In the survey world, eye-tracking has mostly been used in the development and testing of online instruments and other web pages, but much less often with printed survey materials. Using eye-tracking, we attempted to learn how respondents read our letters (e.g., careful reading, skimming, missing parts, etc.), and whether our visual devices draw attention to important information. In our presentation, we will share findings from our preliminary qualitative analysis of eye-tracking output (videos and still images). We found that most people read the letters systematically and carefully, perhaps as a result of the laboratory setting. However, we also found evidence that our visual design elements drew respondents' attention to important information during both systematic reading and pre-attentive processing.

Presentation Title	Creating a Redesigned Questionnaire for the CE Survey Using Colectica
Authors	Brett McBride (Bureau of Labor Statistics), Parvati Krishnamurty, (Bureau of Labor Statistics)
Presenter Email	mcbride.brett@bls.gov

Changes to questionnaire items are routine in surveys. In addition, as surveys attempt to be more flexible to reduce respondent burden, skip patterns become more complex. The documentation of changes to survey items and the interdependencies between the items, and how easily this documentation can be accessed, significantly affect the efficiency with which survey metadata can be queried. The quality and accessibility of this documentation also significantly affect the efficiency of updates to data processing systems needed to accurately process these survey changes. One of the key elements of the comprehensive redesign of the Consumer Expenditure (CE) Survey involves creating an updated and streamlined version of the current CAPI questionnaire. This instrument redesign also provides an opportunity for the CE program to move towards more efficient and comprehensive documentation of survey metadata. Colectica Questionnaires software is being used to create specifications for the new CAPI instrument which will later be programmed into Blaise. Colectica Questionnaires, is part of the Colectica suite of software based on the Data Documentation Initiative (DDI) international standard for describing surveys, and generates various outputs including source code for computer assisted information systems. We will discuss the important features of the CAPI instrument redesign, such as aggregation, question order, use of screeners, record use, as well as the CE program’s experience with using Colectica to generate specifications for the redesigned instrument.

Presentation Title	Developing a Standard Measurement of Housing Insecurity in Surveys
Authors	Jessica Graber (U.S. Census Bureau), Matt Virgile (U.S. Census Bureau), Dave Tuttle (U.S. Census Bureau), Nicole Watson (U.S. Department of Housing and Urban Development), George Carter (U.S. Department of Housing and Urban Development)
Presenter Email	jessica.ellen.graber@census.gov

The concept of "housing insecurity" is used by researchers and policymakers, but has not been measured in one consistent way throughout the research literature. The lack of a comprehensive consensus measure makes it challenging to track prevalence of housing insecurity from year to year and to examine its relation with health, education, employment, and criminal justice outcomes. There is no comprehensive transferable instrument for measuring housing insecurity that is currently used by national surveys. For this reason, the U.S. Department of Housing and Urban Development identified the development of a housing insecurity module as a research priority. The goal of developing the housing insecurity module is to construct a standardized series of questions to measure the continuum of housing insecurity. A transferable survey module will help researchers build a more robust and coherent body of knowledge around housing needs, trade-offs, and correlates, enhancing the quality and consistency of policy-relevant research, and amplifying the visibility of the continuum of housing needs. As part of this research, the U.S. Census Bureau conducted cognitive pretesting in 2018 on a new housing insecurity module, which will be administered to low-income households as a follow-up survey after completing the 2019 American Housing Survey (AHS). Questions on the new survey address three primary elements of housing insecurity: "Affordability," "Stable Occupancy," and "Decent and Safe." We discuss findings about question difficulty and sensitivity from our cognitive pretesting, and revised or reordered question wordings that will be administered in the 2019 AHS based on our findings. We also discuss next steps in the scale development process of the housing insecurity research module, which will be based on data from the 2019 survey production.