DC-AAPOR & Washington Statistical Society
Summer Preview/Review Conference
Schedule and Program

July 18, 2016
USDA Conference Center
355 E Street SW, Washington DC

Time	Session	Title
8:00-8:30		Registration
8:30-8:40		Welcome & Opening Remarks
8:40-10:10	Session 1A	Cross Cultural Survey Methodology
	Session 1B	Advances in Questionnaire Design and Evaluation Methodology
10:10-10:35		Coffee Break & Posters
10:35-11:50	Session 2A	Geography in Surveying and Public Opinion
	Session 2B	New Methods in Gathering and Linking Data
11:50-1:20		Lunch on Own
1:20-2:50	Session 3A	Approaches to Coverage, Sampling, and Weighting
	Session 3B	Mode and Interviewer Effects
2:50-3:15		Coffee Break & Posters
3:15-4:45	Session 4A	Data Analysis and Dissemination
	Session 4B	Survey Design Issues
4:45-4:50		Closing Remarks
5:00		Happy Hour @ Rooftop Bar (400 E St SW)

Sponsored by:

Gold Sponsors – IMPAQ International & RTI International

Silver Sponsor – Fors Marsh Group

Session 1A: Cross Cultural Survey Methodology

Presentation Title	**Leadership Development Through Volunteer Activities**
Presenter Name	**Cathy Furlong**
Presenter Email	**cathy.furlong@cox.net**

Statistics Without Borders is an out-reach committee of the American Statistical Association. All work completed by officers, committee members, and project members is pro-bono. This situation has both advantages and disadvantages. However, this situation also provides a unique opportunity to develop and expand one's leadership skills. I will review my experiences and make suggestions to individuals looking into volunteer activities.

Presentation Title	**Assessing Translations: How different checking procedures compare under field conditions**
Presenter Name	**Danielle Cuddington**
Presenter Email	**DCuddington@PewResearch.org**

Achieving a good translation of a source questionnaire from one language into multiple other languages that maintains consistent measurement properties is often a challenge. Until recently, the most commonly used method for assessing questionnaire translations was back-translation. Research initiated by Harkness's work has shown the weaknesses of this approach. This led to the development of team and committee focused approaches, which are now widely used in multi-country projects. But, these types of approaches are rather complex, costly and have implications for the timeline of projects. Translation verification is another approach that has been discussed as an alternative to back-translation procedures and committee approaches. An independent translator reviews the original translation and discusses with the original translator how to modify the translation. In our paper, we outline the extent to which the verification approach resulted in changes to translation. Subsequently, we evaluate whether changes suggested through the verification approach to existing translations â€" that were back-checked in previous editions of the survey â€" produce different results in terms of item non-response and in terms of the distribution of the substantive answers. To do so, we will analyze the results of the 2014 and 2015 editions of the Global Attitudes survey, comparing changes over time between items for which the translation was changed following the verification, items for which the translation was flagged but not changed and items for which the translation was not flagged. Although the evaluation does not allow us to assess the effectiveness of either approach in an absolute sense, it should provide some insight into any relative improvements made by applying the verification approach.

Presentation Title	**Cross-Cultural Instrument Development and Testing: Lessons from India and Rwanda**
Presenter Name	**Mousumi Sarkar**
Presenter Email	**msarkar@impaqint.com**

With funding from the United States Department of Labor International Labor Affairs Bureau (USDOL/ILAB) Office of Child Labor, Forced Labor and Human Trafficking (OCFT) IMPAQ is conducting randomized controlled trials (RCTs) to evaluate the effectiveness of five child labor interventions in five countries: India, Malawi, Rwanda, Ecuador, and Costa Rica. The context of child labor varies in the five countries, from young children involved in mica mining in India, to tea plantation activities in Rwanda, working in tobacco farms in Malawi, and children leaving school and working primarily in the informal sector in a variety of urban environments in Ecuador and Costa Rica. The programs vary from educating and empowering the community to, facilitating returning to school, providing vocational training or intensive non-cognitive school curriculum modules to enable children to earn their secondary school certificate. Surveys are being implemented with 1) heads of household and women with children in India; 2) children aged 16-17 years in Rwanda in a vocational training program; 3) head of households in Malawi's tobacco growing area; 4) 15-21 years olds school dropouts in Ecuador enrolled in an intensive education program; and 5) children of all ages who have dropped out of school and are being enrolled back in Costa Rica. The questionnaires are designed to measure the impact of these programs. Questions include assessment of the prevalence of child labor. While the key research questions are shared across the five programs, the questionnaires are adapted to be relevant within the local context, the programs and populations. After the instruments are drafted, we will conduct cognitive testing with to refine the instruments. We will analyze the data from the testing paying special attention to change due to languages and cultures. We will analyze the differences in the ways respondents interpret and define concepts, and patterns in response and interpretation of response categories that might stem from cultural or linguistic factors. Once the instruments have been finalized based on the findings from the cognitive testing, we will pilot test each instrument with a sample that is similar to our target population. The pilot test will include some de-briefing questions for both the respondent and the interviewer. Analysis of these data will be conducted to assess if further changes are required in the instruments. We propose to present our analysis of these data from the cognitive and the pre-tests, and the resultant changes to the questionnaires.

Presentation Title	**Reducing non-response for income questions**
Presenter Name	**Steve Schwarzer**
Presenter Email	**SSchwarzer@pewresearch.org**

Creating comparable, cross-national measures of household income is inherently challenging, given differences in language and context as well as often prohibitively high non-response rates. In the case of public opinion surveys, which allow limited space for exploring the multiple dimensions of financial, human and social capital, the challenge is even more daunting. Yet, income is a key factor for understanding how individual attitudes vary within and across nations when it comes to issues such as personal aspirations, country satisfaction, and technology use. Given these limitations, what is the most useful survey method for reducing non-response rates while providing data that is comparable across a wide-range of countries?   The Pew Research Center has conducted surveys in more than 40 countries every year for the past three years, which provides a unique opportunity to test and validate question wordings and formats across a range of countries. Using empirical data from a survey experiment and several cross-national surveys, we explore one mechanism for reducing refusal rates on income questions. We were particularly interested in testing the effect of different follow-up questions after asking a closed-end or open-end income question. We tested (1) a closed-end question with a median-based follow up, (2) an open-end question with a series of follow-up questions, and (3) a median-based question with a closed-end follow up. The paper reports results demonstrating the effectiveness of the different approaches used. By combining a closed-end income question with a more general, median-based, follow-up question, we were able to significantly reduce refusal rates for our income questions in the majority of countries. We demonstrate that follow-up questions also have differential effects on various demographic groups.

Presentation Title	**The First Self-Administered Survey in North Korea: Self-Esteem of North Koreans Compared with Peers in 53 Other Countries**
Presenter Name	**Elliott Y. Chun**
Presenter Email	**ychun2@gmail.com**

The purpose of this interdisciplinary paper is to discuss methods and findings from a self-administered survey conducted for the first time in North Korea with its focus on assessing North Korean self-esteem and socioeconomic correlates. We tested the Rosenberg Self-Esteem scale in North Korea. Findings from this unprecedented North Korean survey are compared with self-esteem previously measured for peers in 53 other countries. Co-authors have been involved in this rare North Korean survey design, data collection and analysis. We will discuss research challenges, methodological approaches and findings relevant and useful to understanding social psychological fabric of North Korea. Co-authors are faculty and research associates affiliated with the ISR Foundation Center for Interdisciplinary Research and Center for Science Diplomacy as well as academic institutions.

Session 1B: Advances in Questionnaire Design and Evaluation Methodology

Presentation Title	**Respondent Processing of Rating Scales and Scale Direction Effects: An Eye-tracking Study**
Presenter Name	**Andrew Caporaso**
Presenter Email	**andrewcaporaso@westat.com**

Holding constant other scale features, the direction in which a scale is presented has been found to affect the resultant survey answers; respondents are more likely to select a scale point closer to the start of the scale regardless of its direction, producing primacy effects (Yan, 2015). What remains understudied is the mechanism underlying this scale direction effect. Two common response processing models are offered as possible explanations for these effects: satisficing, and anchoring and adjusting. The satisficing model treats the impact of scale direction as a special case of response order effect and argues that satisficers sequentially process the rating scale and select the first option that seems reasonable. The anchoring-and-adjustment heuristic assumes that respondents start with an initial anchor (the beginning of a scale) and make adjustments to the anchor until a plausible point is reached. Since both notions predict a primacy effect, it is hard to know which notion accounts for scale direction effect. In order to learn more about what's behind scale direction effects, we will collect eye tracking data from respondents as they respond to a web survey. As the eye movement data (e.g., fixation counts and fixation duration) show directly the amount of attention paid to question components, we will first characterize how respondents process a rating scale and how the processing differs by respondent characteristics. Then we will explore which of the two notions accounts for the scale direction effect. This paper demonstrates how eye-tracking can be used to address theoretical issues related to respondents' use of rating scales.

Presentation Title	**Instructions in Self-Administered Survey Questions: Do They Improve Data Quality or Just Make the Questionnaire Longer?**
Presenter Name	**Chelsea Owens**
Presenter Email	**chelsea.owens@ed.gov**

Pre-testing techniques utilized in the development of production self-administered questionnaires, such as cognitive interviewing, often identify items where respondents misinterpret or are unclear about the meaning of terms in a question. Typically, this finding results in a recommendation to add instructions to an item, which has the detrimental effect of lengthening the questionnaire. Previous experimental research has shown that instructions have an effect on the estimates when the instructions counter the way many people naturally tend to think about a concept. For example, an instruction to exclude sneakers from a count of shoes will reduce the estimate of shoes because many respondents tend to think of sneakers as shoes. However, few studies have looked empirically at whether or not instructions that are the product of actual production pre-testing techniques are similarly effective or useful, and worth the extra length they create. In addition, previous research has shown that instructions placed before questions are more effective than those placed after. To examine these issues further, we report on a study that was administered to a nationally representative sample of teachers (n = 1616 completed questionnaires) by web. Production questions and instructions were selected from a national teacher survey and their effectiveness is compared to questions and instructions that were intentionally created to counter teachers' natural conceptions of terms, as well as a control group with no instructions. Utilizing a factorial experimental design, we also varied three factors that were predicted to alter the effectiveness of instructions: their location, format, and wording. Although the findings of this experiment are clearly generalizable to the web, arguably, these findings extend to mail surveys too.

Presentation Title	**The Effects of Pictorial vs. Verbal Examples on Survey Responses**
Presenter Name	**Hanyu Sun**
Presenter Email	**HanyuSun@westat.com**

Web surveys make it earlier to present images to the respondents than other modes of data collection. A few studies have examined the use of pictorial examples in Web surveys and found that the characteristic of the exemplars (e.g., their frequency or typicality) has an impact on the responses that are collected (e.g., Couper, Tourangeau, and Kenyon, 2004; Tourangeau, Conrad, Couper, and Ye, 2014). Tourangeau et al. (2014) compared visual examples with pictorial examples and found that respondents tended to report more foods consumption when they got verbal examples than when they got pictorial examples. The finding suggests that the pictures may narrow the interpretation of the category of interest. However, the findings also suggested that respondents are more likely to attend to the pictorial examples than to verbal examples. However, no direct evidence of respondent attention was collected to support either argument. Using eye-tracking, the current study compared verbal examples with pictorial examples in a lab setting to examine whether respondents attend to pictures more than words, whether items with verbal examples require more effort to answer than those with pictorial examples, and how the processing of items change over time. To address these research questions, we will examine differences in mean number of fixations and the mean duration for items with pictorial examples and verbal examples. The number of fixation is related to the amount of information that a respondent is processing, while the duration of fixations is related to the amount of difficulty that the respondent is having (Ares et al., 2014). The same food consumption questions with examples used in Tourangeau et al. (2014) will be used in current study.

Presentation Title	**Combining Multiple Questionnaire Testing Methods: The Bento Box Approach in the 2017 Census of Agriculture Testing**
Presenter Name	**Jaki McCarthy**
Presenter Email	**jaki_mccarthy@nass.usda.gov**

A traditional Japanese bento box meal is constructed according to principles of balance -- separate compartments with different colors, flavors, and cooking methods provide an optimum eating experience. As in the bento box, questionnaire testing methods can be combined to provide a balanced questionnaire evaluation. The quinquennial Census of Agriculture (COA) is the largest data collection conducted by the U.S. Department of Agriculture's National Agricultural Statistics Service. Farm size, type, production, economics and demographics of farm operators are collected with a self-administered form, mailed to ~ 3 million potential farms. Similar to the 2012 COA, NASS is using multi-method questionnaire testing for the 2017 COA. Initial activities included internal expert reviews by NASS subject matter and production staff. In addition, an external expert panel provided recommendations regarding the measurement of women and beginning farmers. Item edit and imputation rates in the 2012 COA were reviewed. The 2012 telephone help calls for the form were also reviewed. These initial reviews and proposed new content resulted in initial questionnaire drafts. Multiple rounds of cognitive interviews were then conducted focusing on new material and form changes. In early 2016, a field test with ~30,000 records was conducted. Data from this test will be analyzed to evaluate the forms and identify any

remaining problem areas. The test will also serve to test census processing systems. Respondents in the field test with suspected reporting errors will also be contacted in follow-up cognitive interviews. All of our questionnaire bento box elements will be combined to make final recommendations for the 2017 census questionnaires. Each of the bento box compartments and the complementary information it provides will be discussed with selected results as illustration. Lessons learned to help improve the process for the 2022 COA will also be discussed.

Presentation Title	**Parents and Teens: Proxy versus Self-Report to Measure Teen Disability**
Presenter Name	**Meredith Massey**
Presenter Email	**wnx6@cdc.gov**

Who should answer questions about teen disability: teens or their parents? Parents are generally viewed as suitable proxies for their children since children may have difficulty answering questions due to age, illness, disability or literacy and because it is the parents who most often seek medical care and support services for their children. Therefore, many surveys that measure disability prevalence are designed to be administered to parents. However, older children, teens in particular, may be better able to evaluate their own difficulties especially in domains that are focused on internal processes such as memory or emotions. Using data from a cognitive interview study testing the UNICEF/Washington Group on Disability Statistics Module on Child Functioning and Disability, differences in the ways parents and teens answer and interpret questions on teen disability will be explored. The results of this study, which collected data from 40 parent-teen dyads (n=80), demonstrate that overall, parents and teens answered similarly. However, it was seen that the dyads were more closely aligned in their evaluations of questions categorized as `behavior' questions than they were for those categorized as â€˜ability' questions. Some discordance in teens' and parents' responses was a result of their different interpretations of the questions. Further, for some questions, parents and teens had differing patterns of interpretation even when they answered similarly. Other factors that influence dyad agreement (such as age and gender), will also be explored. Finally, we will outline the implications of these findings on the design and administration of surveys measuring the prevalence of teen disability.

Session 2A: Geography in Surveying and Public Opinion

Presentation Title	**Optimal Routing for Field Interviewers: A GIS-based Approach for Reducing Household Survey Costs**
Presenter Name	**Brad Edwards**
Presenter Email	**bradedwards@westat.com**

Survey costs are increasing, especially in face-to-face surveys with their high labor and travel requirements. Data collection in the field is almost always the cost driver for in-person household surveys. The largest share of the interviewer's time is spent traveling from home to the assigned sampled units. Travel costs are also related to which specific units the interviewer chooses to visit in that block of work time. The most obvious strategy, to visit all units in a cluster, may not be the most efficient. The propensity for a given unit to respond at a specific time of day may be so low that it is not worth the interviewer's time to stop. Other constraints must also be taken into account, such as how long each interview takes, and the outcome of the previous visit to an address. The means to address the field interviewer scheduling/routing problem have been lacking until recently. Today, analytics based on GIS data and paradata such as call attempt outcomes from individual cases enable managers to see much more clearly what is happening in the field in near real time. We report on experiences from two national surveys that have provided interviewers with mobile devices that have GPS capability. Interviewer routing and visits to specific units are observed by supervisors. Paradata analytics inform judgments about optimal case work and routing. Supervisors review the choices interviewers made about cases to visit and routes taken. The supervisors discuss the actual routes with the interviewers the next day and make suggestions for more efficient routing, including use of navigation functions. Routing behavior during the next work block is observed, and additional feedback is given as required. We discuss the impact of this new capability on field survey efficiency and costs.

Presentation Title	**Incorporating Mobile Mapping Instruments in the June Area Survey Data Collection Effort**
Presenter Name	**Denise A Abreu**
Presenter Email	**denise.abreu@nass.usda.gov**

The National Agricultural Statistics Service (NASS) conducts the June Agricultural Survey (JAS), which is based on an area sampling frame. Segments of land comprise the sampling units for the JAS. Traditionally, for the JAS, field enumerators use a hard copy aerial photo to locate and interview all operators within the segment boundary. Then, they draw off all the fields by hand on the aerial photography. In an effort to incorporate newer technologies in it data collection effort, NASS is evaluating the use of a mobile mapping instrument that would replace the aerial photo and paper questionnaire.

Research conducted in 2014, using the mobile mapping prototype, indicated that drawing fields during the interview takes longer than is operationally feasible even when enumerators are proficient with use of the instrument. Testing in 2015 focused on using JAS segments with pre-delineated fields in the mobile mapping instrument. Fields were delineated using Common Land Units (CLUs) from the Farm Service Agency (FSA) whenever available. For areas without FSA CLU coverage, cartographers used various sources to delineate fields within the selected JAS segments. Enumerators were then provided with prepared field boundaries in the mobile mapping instrument to compare completion times to current procedures using the paper aerial photo. Enumerators recorded previous year's JAS data using a mock interview format. Research results and the future direction of NASS's data collection activities are described in this presentation.

Presentation Title	**Innovative Quality Assurance Use of GIS Data in Field Surveys**
Presenter Name	**Marsha Hasson**
Presenter Email	**MarshaHasson@Westat.com**

As the ability to collect geospatial data has become more accurate and ubiquitous, researchers are turning attention to other uses of this information to support operations management and quality assurance. Data exploration and visualization tools can support improvements in efficiency, level of effort, accuracy and time to discovery for project management staff, quality assurance specialists, and analysts. This approach is based on importing GIS data from mobile devices and field laptops and combining it with other study information (e.g., visit logs, questionnaires, time and expense). A visual review of the interviewer's day superimposed on a map shows the tasks performed in location/time context, among other features. This rich data can be used to review efficiency of routes taken, confirm time and expense reports, and other quality assurance information. One example is the use of this data to identify interviewer falsification of study activities. Previous studies have explored various statistical or pattern approaches to identify at-risk interviewers based on various indices. Other common approaches include reviewing recorded interviews and re-interviewing respondents. These approaches may take days to weeks to finalize and fraudulent activity may continue during that time. The cost to detect, confirm, remediate the data, and re-field cases can be substantial. Geospatial data offers a quick, inexpensive way to detect falsified interviews across 100% of the sample. Similarly, expenses can be verified by comparison to time and distance traveled. Analysis of two major field surveys demonstrates that the use of geospatial information succeeded immediately in reducing the time to discovery and level of effort associated with falsification detection and remediation. A data exploration and visualization tool for geospatial information is a valuable asset in ensuring the quality of survey data and reducing costs associated with interviewer falsification.

Presentation Title	**Dwellings Enumeration and Estimation: Geo-Listing or Drones?**
Presenter Name	**Safaa Amer**
Presenter Email	**samer@rti.org**

In many countries, the census of population and dwellings is the basis for implementing household surveys. In between censuses, estimates are updated using projection models. Censuses are large- scale, resource-intensive efforts that in many countries suffer from under-coverage and other sources of error. In addition, projection models are based on specific assumptions which might not apply. In an attempt to reduce error in projection models and provide planning tools which can reduce the cost of dwelling and household census, we conducted a case study on the use of drones to improve estimates of dwellings in between census and provide planning tools for upcoming household survey not to mention the next census data collection. In this paper, we discuss the use of drone imagery combined with point cloud technology to generate 3D images of buildings and help in the enumeration of household without having to use field staff to do a traditional listing. We also describe a system for extracting information from these images to provide more accurate information for updating dwelling and household counts. We investigate the use of the enhanced imagery to support the planning of household surveys and census operations. We compare the drone imagery with information available from aerial satellite photography and in some limited cases the result of traditional listing operations

Session 2B: New Methods in Gathering and Linking Data

Presentation Title	**Wearables: Passive Measurement Tool of the Future**
Presenter Name	**Adam Gluck**
Presenter Email	**Adam.Gluck@nielsen.com**

Wearable technology includes a myriad of devices that possess advanced functionality, connect to the Internet (typically via smartphone), and are usually designed to be worn on the body persistently (Endeavour Partners, 2014; Khan & Marzec 2014). As researchers become increasingly focused on more passive ways to measure people's behavior, wearable technology provides a comprehensive method of data collection. The appeal of wearable technology for research

purposes lies in the fact that it gives up-to-the-minute data without distracting the user from other responsibilities and activities. Users expect wearable technology to be multi-functioning, with a fast response time, and very little interaction between the user and device (Cognizant, 2014). Nielsen uses a variety of devices to passively measure media consumption, including wearable meters that measure media exposure to encoded television and radio signals. People are recruited and asked to wear or carry the meter wherever they go throughout their day. The current form of the wearable meter, a pager-like device, is a significant barrier to compliance particularly for key subgroups such as young women and children. Our research explores challenges with carrying or wearing a meter, people's interest and use of wearable devices, preferences in form and additional functionality, and willingness to wear for sustained periods of time. The goal is to evaluate whether people can better integrate a more wearable meter that leverages these new designs into their everyday lives, improving overall measurement and compliance. We will present findings from an iterative, mixed method approach to exploring alternative form factors, including an extensive literature review, a survey and focus groups with former panelists, a nationally representative survey of the U.S. population, and focus groups with field staff who interact with panelists.

Presentation Title	**Evaluating Record Linkage Software for Agricultural Surveys**
Presenter Name	**Michael Bellow**
Presenter Email	**mbellow@nass.usda.gov**

Reducing duplication and matching records lacking unique identifiers are common practices associated with the building and maintenance of a list sampling frame. The National Agricultural Statistics Service (NASS) employs a record linkage system built using AutoStan and AutoMatch (originally developed by MatchWare Technologies) to maintain its list frame of farm operators and agribusinesses. The overall process consists of four steps: 1) reformatting, 2) standardizing (AutoStan), 3) matching (AutoMatch) and 4) review. Because the current match engine is no longer supported and becoming increasingly obsolete, NASS has recently begun to explore alternative software options such as Statistics Canada's G-Link package. In this paper, we describe the results of a preliminary study comparing G-Link with AutoMatch/AutoStan using list frame data from a national survey of organic farmers and discuss issues associated with upgrading the agency's record linkage system.

Presentation Title	**Best practices for integration of alternative data sources into petroleum programs**
Presenter Name	**Shawna Waugh**
Presenter Email	**shawna.waugh@eia.gov**

The Office of Petroleum and Biofuels Statistics designs and implements the Petroleum Marketing and Petroleum Supply programs, publications, and surveys. PBS integrates administrative records and third-party (A3P) data sources to supplement or to replace survey data with the intention to reduce respondent burden and program costs, to improve quality, and/or to reduce data gaps. This session will focus on various uses of A3P data, pros and cons of A3P data, challenges and opportunities of integrating A3P data sources, and best practices throughout the survey lifecycle.

Presentation Title	**Use of Eye-tracking to Measure Response Burden**
Presenter Name	**Ting Yan**
Presenter Email	**tingyan@westat.com**

Concerns about the burden that surveys place on respondents have a long history in the survey field. A review of the existing literature shows that the term "burden" is defined loosely and that researchers measure response burden in many different ways. Some measure response burden through properties of surveys/tasks that are believe to impose response burden, such as the length of an interview. Some measure response burden through respondents' attitudes and beliefs toward surveys, such as interest in and perceived importance of the survey. Others measure response burden through respondent behaviors (e.g., willingness to be re-interviewed) or direct respondent measures (e.g., feelings of burden). All three types of measurement are based on self-reports to survey questions, which are subject to the usual sources of reporting error due to misunderstanding of the survey questions, partial retrieval of information, biased or inaccurate judgment strategy, problems in mapping to the given response options, and more or less deliberate misreporting. In this paper, we will examine the use of eye-tracking equipment to measure burden. In eye-tracking research, task-evoked pupillary responses have been shown to be a consistent index of cognitive load and difficulty. For instance, dilated pupils are found to be indicative of higher levels of cognitive load and difficulty, in essence burden. We will create three measures -- mean pupil dilation, peak pupil dilation, and latency to peak -- as indicators of burden. These alternative measurements of response burden are free from errors in self-reports and are potentially stronger indicators of burden. We will evaluate the feasibility of using task evoked pupillary responses to measure burden in the survey context by comparing the three indicators to respondents' self-reports about burden.

Session 3A: Approaches to Coverage, Sampling, and Weighting

Presentation Title	**Conducting a Telephone Survey Using an ABS Sample: A Case Study of the California Health Interview Survey**
Presenter Name	**Jennifer Kali**
Presenter Email	**jenniferkali@westat.com**

Geographically targeting samples in small areas is difficult in telephone surveys that include a cell phone component. There have been several solutions practitioners have used to increase the efficiency of targeting telephone surveys to small areas. One solution is the use of Address Based Sampling in which the sample unit is an address allowing for very precise geographic targeting. Because the sample unit is an address, an ABS frame is best suited for mail surveys. In some surveys, the flow of questions and skip patters in the instruments are too complex to be conducted on paper. Thus, there has been interest in using ABS as a way to recruit respondents for telephone surveys. To do so, a telephone number must be obtained for the addresses sampled via ABS. A telephone number can either be matched to the sampled address or obtained via a letter is mailed to the sampled address requesting respondents to provide their telephone number. The California Health Interview Survey is a dual-frame RDD telephone survey of the population of California. In 2013-2014, ABS was used to supplement the RDD sample in one rural county in which geographic targeting for the cell phone sample required a large screening effort. This presentation describes the CHIS experiences with using ABS to conduct a telephone survey and examines the differences between the ABS and RDD samples. Methods to obtain telephone numbers, data collection outcomes, including response rates, disposition codes, and accuracy of the telephone matching, and estimates of health and demographics are compared between the RDD and ABS samples.

Presentation Title	**The Effect of Sending Advance Letters to Cellphone Respondents Using Appended Addresses**
Presenter Name	**Kyley McGeeney**
Presenter Email	**kmcgeeney@pewresearch.org**

One of the key challenges with cell phone random-digit-dial surveys has been the "cold call" nature of the recruitment. The researcher has little if any opportunity to explain the purpose and legitimacy of the survey before sampled adults start deciding whether they will answer the calls, much less participate. Thanks to a new advancement in cell RDD sampling, this dynamic is changing. Telephone sample vendors now offer the ability to sample on or append billing address and demographics to the cell phone sample. Researchers can use this information to send advance letters and incentives to future cell phone respondents. While promising, this possibility raises a number of research questions. What percent of the cell phone sample has a matched address and what percent has matched demographics? What would be the effect of using these addresses and demographics to send a targeted advance mailing to cell phone sample addresses in terms of response rate, resulting sample composition and cost? To answer these questions the Pew Research Center conducted an experiment using a monthly political survey. Addresses and demographics were matched to the entire cell phone sample. Cell phone numbers with a matched addresses and demographics were randomized to a treatment or a control group. The treatment group was sent an advance mailing containing a small prepaid cash incentive and a letter explaining that they would soon receive a phone call for a survey. The match rate of addresses and of demographics was examined. Researchers also analyzed the response rate of the treatment and control groups, the demographics composition of the resulting samples and the cost associated with this design.

Presentation Title	**Calibration Weighting for Nonresponse with Proxy Frame Variables**
Presenter Name	**Phillip S. Kott**
Presenter Email	**pkott@rti.org**

A calibration weight-adjustment function frequently used when adjusting for nonresponse assumes that the unit response/nonresponse mechanism is a function of variables known either for the entire sample before unit response or at the aggregate level for the frame. Often, however, some of the variables governing the response/nonresponse mechanism can only be proxied by variables on the frame while they are measured (more) accurately on the survey itself. For example, an address-based sampling frame may contain area-level estimates for the median income and the fraction home ownership in a census block group, while a household's income category and ownership status are reported on the survey itself, but only for housing units that respond to the survey. A relatively new calibration-weighting technique employed by the WTADJX procedure in SUDAAN 11 allows a statistician to calibrate the sample using proxy variables while assuming the response/nonresponse mechanism is a function of the analogous survey variables. We will demonstrate how this works with data from the Residential Energy Consumption Survey National Pilot, a nationally representative web-and-mail survey of U.S. households sponsored by the U.S. Energy Information Administration.

Presentation Title	**The utility of non-probability samples for assessing population-level health information seeking behavior**
Presenter Name	**Reanne Townsend**
Presenter Email	**reannetownsend@westat.com**

The goal of this paper is to help to inform the literature on non-probability sampling by comparing estimates produced from two non-probability surveys to estimates from a national probability sample. Surveys utilizing probability samples of a target population are widely used in a variety of research disciplines. However, survey participation rates have been declining for several years, leading to increased cost to recruit sample members and concerns about the representativeness of the data collected. As the amount of effort required to make valid inferences from probability samples has increased, non-probability samples have become an increasingly used alternative. Although there have been a variety of previous comparisons between non-probability and probability samples, the field of survey research is still struggling to understand the circumstances under which non-probability samples yield valid inferences of population estimates. Further, there is no single approach to non-probability sampling, as this general term subsumes a range of specific practices that may influence the cost, representativeness, and usefulness of the collected information. We will examine estimates produced from these three non-probability samples which represent alternative approaches to recruiting sample members, and vary in the cost of obtaining the sample. These estimates will be compared with a national probability sample mail survey of households about general health and health information seeking behavior. We will examine differences including point estimates and associations between variables. Finally, we will suggest conclusions concerning the trade-offs involved in selecting fundamentally different approaches to the collection of self-report information through non-probability sampling.

Presentation Title	**Covering our Most Mobile Users: Identifying Which States are Most Susceptible to Coverage Error**
Presenter Name	**Stephanie Marken**
Presenter Email	**stephanie_marken@gallup.com**

The percentage of cellphone only households and cellphone mostly users continues to increase nationally, and researchers are continually challenged to arrive at a sampling strategy that appropriate covers this increasingly mobile population. The percentage of cellphone users that reside in a different state than the state in which they purchased their cellphone poses a significant obstacle to properly sampling at the state level using an RDD cellphone frame. In their 2012 publication, Skalland et. al. demonstrated how this inaccuracy in the sample state and current state of residency varies tremendously by state for cellphone only users, and how undercoverage and overcoverage can impact final survey estimates. In 2015 Gallup built upon this research by sharing estimates of inaccuracy for all cellphone users by state, and information about how these differences can introduce bias into final survey estimates. In this paper, Gallup researchers provide updates of these estimates from a nationally representative survey of all U.S. adults. We also provide researchers with more information about which states with the highest undercoverage rates may benefit from sampling adjoining rate centers in neighboring states to minimize undercoverage error in final survey estimates. We also estimate total cost and productivity losses that result from sampling rate centers in adjoining states.

Session 3B: Mode and Interviewer Effects

Presentation Title	**The Future of Telephone Surveys**
Presenter Name	**David Dutwin**
Presenter Email	**ddutwin@ssrs.com**

I would propose to provide a summary of the panel on the Future of Telephone Surveys as presented by me, Trent Buskirk, and Tom Guterbock. I believe I can keep this brief, despite the volume of new data that we presented at AAPOR! The outline would be:
1. Trends in Telephone Dispositions in the Past Decade: a summary of data I compiled on behalf of the AAPOR Future of Telephone Surveys task force, showing a dire picture with regard to productivity of landline samples, but cell phone samples continuing to hold steady over the past decade. Detailed results of the data will be shown for trends in refusals, no answers, answering machines, and non-working numbers, by landline and cell frames.
2. Trends in Telephone Sample Data Quality, 1996 - 2015: A summary of an analysis showing that bias in telephone samples was basically flat from 1996-2005, increased from 2005 - 2010, and has been on a downward trend since then. We find it likely that the downward trend is due to firms increasing their cell phone samples.
3. Trends in Cell Phone Sample Data Quality, 2009 - 2015: A summary of how well cell phone samples have represented the ever-moving cell phone population in the U.S. 4. Trends in Telephone Survey Costs: A synopsis of Tom Guterbock's 3 studies on telephone survey costs, 2010, 2013, and 2015.

Presentation Title	**Item Administration Times and Item Nonresponse to the National Health Interview Survey Sexual Identity Question: An Examination of Interviewer Effects**
Presenter Name	**James M. Dahlhamer**
Presenter Email	**fzd2@cdc.gov**

In 2013, the National Health Interview Survey (NHIS) added a question on sexual identity to the sample adult interview. The question was the culmination of an extensive testing effort, including a large-scale split-ballot field experiment where estimates of sexual minorities (i.e., gay/lesbian, bisexual) were compared by mode of administration: computer-assisted personal interviewing (CAPI) versus audio computer-assisted self-interviewing (ACASI). Because no significant differences by mode were identified, the decision was made to include the question as part of the standard CAPI interview. After two years of CAPI administration, however, concerns have emerged that some interviewers may not be asking the question, and/or eliciting an above-average number of non-substantive responses (e.g., don't know, refused). Hence, some interviewers may adversely affect the quality of the sexual identity data. Using 2013-14 NHIS data, interviewer effects on (1) item administration times and (2) item nonresponse to the sexual identity question are explored via multilevel logistic regressions. Item administration times in the bottom quintile of the time distribution are compared to longer times in the first analysis. In the second analysis on item nonresponse, don't know, refused, and responses coded as "sexual minority" including gay/lesbian and bisexual, are compared to responses of straight. For both dependent variables, unconditional models are run to produce baseline estimates of the variance attributable to interviewers. Sequential models are then run, starting with the addition of respondent/household characteristics (e.g., age, education, number of household members), followed by the addition of interviewer characteristics (e.g., years working for the Census Bureau, education, gender). Significant fixed effects are highlighted. The paper concludes by discussing the implications of the results for asking about sexual identity in interviewer-administered survey modes.

Presentation Title	**SMS vs. Mobile Web in Emerging Markets**
Presenter Name	**Max Richman**
Presenter Email	**Max@GeoPoll.com**

Over the past few years mobile penetration has greatly increased in emerging markets in Africa, Asia, and Latin America. Although mobile surveys themselves are still a new method of data collection there is a need for researchers to keep up to date with the latest trends in mobile. For example, self-completed surveys can be administered through several channels including SMS, Interactive Voice Recording (IVR), and mobile web. Each of these methods has advantages and disadvantages. While SMS and IVR can be administered on any type of phone, with or without internet access, they have limitations regarding message and questionnaire length. Mobile web mitigates the issues with question length and complexity, however web-enabled phones have a lower penetration in emerging markets and there is a question of the representatively of a mobile web survey. In this paper we will compare two methods of remote mobile data collection: SMS and mobile web, looking at completion rates on each, and any differing characteristics in each sample. Very little research has been done on remote mobile surveys in any capacity, and this paper aims to provide concrete findings on the characteristics of SMS and mobile web. Research will be conducted in South Africa, which has a relatively high percentage of mobile web users comparative to other emerging markets, and Mozambique and Uganda, which have lower mobile web penetration. Initial findings suggest that there are differences between mobile web and SMS, especially when it comes to reaching specific age groups or locations. For example, respondents over the age of 24 may be more likely to respond to mobile web surveys than SMS surveys. This paper will do a comprehensive analysis of how each method can best be utilized to collect data from emerging markets.

Presentation Title	**A Test of Web/PAPI Protocols and Incentives for the Residential Energy**
Presenter Name	**Paul Biemer**
Presenter Email	**ppb@rti.org**

The Residential Energy Consumption Survey (RECS), sponsored by EIA, collects data on householder behaviors and housing characteristics that affect current and long-term energy usage and cost. To build a more responsive and cost-effective data platform, EIA is considering moving the RECS from the current personal interview mode to a Web/PAPI mixed-mode survey design. This paper describes an experiment to test two incentive options and four data collection protocols arrayed in 2x4 factorial design. A national epsem sample of 9,650 households was selected and divided equally among the eight treatments. The two incentive treatments included a $5 prepaid incentive in the first questionnaire mailing but varied the promised incentive amount from $10 to $20 paid after the completed questionnaire is received. The four protocol treatments are: (a) CAWI (or Web only), (b) Choice (which gives respondents the simultaneous choice of responding by either PAPI or Web), (c) Choice+ (similar to the Choice protocol but adds a $10 bonus incentive if the

respondent chooses to respond by Web) and (d) CAWI/PAPI (which initially offers the Web only option, but allows the choice of Web or PAPI in nonresponse follow-up invitations). The paper reports the results of the study including which data collection protocol/incentive structure combination provides the highest quality estimates based upon a total survey error analysis.

Presentation Title	**Are households with unusable phone numbers more likely to respond by mailing than calling, and do we really want them in our sample?**
Presenter Name	**Vrinda Nair**
Presenter Email	**vrinda.nair@nielsen.com**

Whenever we switch the mode of a survey from phone to mail, or from mail to phone we introduce bias. As survey methodologists we want to study the characteristics of the respondents that we attract because of a mode-switch, in order to draw better inference about the population that we represent in our sample. There can be various reasons for switching the mode of an established survey design. It is possible that calling a home is no longer a feasible option because of unusable phone numbers, and hence we want to mail instead. New technologies on the vendor side introduce ways of screening the phone numbers ahead of beginning the survey. At Nielsen we use both phone and mail as part of a multi-mode survey design. The phone numbers appended to our samples are from large consumer databases and the age of the data can vary depending on the source. On average, we phone over 900,000 households four times a year. We normally find that ~40% of the phone numbers are categorized as unusable (not in service, out of order, or no answers) after recruitment. In a 2014 diary test we found that the Cell WINS technology effectively identified unusable phone numbers. In February 2016, to further improve our sample performance, we are investigating removing the bad phone number linked to the address and mailing the household a screener pre-pack instead. The diary pre-pack is usually provided to our ABS sample where we do not have a phone number to introduce ourselves, collect household demographics, TV data, and collect a phone number. Our goal with this test is to identify the shift in household characteristics captured in the respondent data. Does it add more desired demographics to our sample? Are we reaching the right households when we shift the first point of contact mode to mailing from calling for unusable phone numbers? This paper will review the results.

Session 4A: Data Analysis and Dissemination

Presentation Title	**Predicting Pathologic Complete Response Using Microarray Data in the Presence of a Dominating Class**
Presenter Name	**Jonathon Abernethy**
Presenter Email	**abernet1@uic.edu**

Pathologic complete response (PCR) is defined as the complete removal of residual invasive disease in the breast at the completion of chemotherapy. Breast cancer patients who achieve PCR often have a better chance of avoiding disease recurrence. PCR is relatively rare, occurring in around 20% percent of patients. If the goal is to predict PCR when considering chemotherapy, the case where non-achievers form a dominating class in a data set must be considered. In this paper, we used micro array data to predict PCR status using a sample of 271 training cases and 64 testing cases. Roughly 81% of the patients in this data set do not achieve PCR. Modeling without adjustment for the dominating class will lead to a top classification accuracy of 83% by over representing non-achievers (25% sensitivity, 96% specificity). We adjust for the dominant class using three approaches: multiple undersampling of the dominant class, thresholding the SVM decision function, and a modified classifier similar to DLDA. Using the latter approach we significantly increase the sensitivity to 83%, while still keeping a specificity of 79%.

Presentation Title	**U.S. Roadways**
Presenter Name	**Mike Jadoo**
Presenter Email	**mike_jadoo@yahoo.com**

Every year many lives are either lost or life changing injuries are experienced that could have been prevented while driving on the U.S. roadways. With the use of the National Highway Traffic Safety Administration National Automotive Sampling System General Estimates System databases, an analysis was conducted to reveal which areas in the United States is the deadliest place to drive and what factors predicate the occurrence of accidents. The databases were used to conduct a statistical analysis on the maximum injury severity variable and other associated variables in the data set, over time. The report presents the state and area that is prone to accidents and what are the contributing factors of these crashes. As such, the report offers a practical framework to preventing accidents by analyzing what predicates them.

Presentation Title	**Using Text Analytic Techniques to Create Efficiencies in Analyzing Qualitative Data: A Comparison between Traditional Content Analysis and A Topic Modeling Approach**
Presenter Name	**Andrew Stavisky**
Presenter Email	**staviskys@gao.gov**

The U.S. Government Accountability Office (GAO) conducts independent, non-partisan policy studies for Congress. Many of the studies include qualitative research components, such as case studies, in-depth interviews, focus groups, and expert forums. The methodology for analyzing qualitative data for these studies typically uses comprehensive content analysis, an extremely time-and resource-heavy approach. Therefore, GAO has begun to explore ways to find efficiencies in the analytic process while still maintaining the rigor and quality of the analyses. Topic modeling refers to a family of computational techniques that extract latent topics or themes from collections of text, bringing to the surface underlying structure that is not likely to be apparent via traditional techniques such as term-frequency analysis. Variations of topic models have been used in the social sciences for diverse purposes, ranging from analyzing trends in the scientific literature, to characterizing patterns of agenda-setting and agenda-framing in political communication, to identifying thematic categories in psychotherapy transcripts. In this study, traditional content analysis was compared to a topic modeling approach using 12 focus group transcripts from a completed GAO investigation that employed comprehensive content analysis. Recommendations to a Federal Agency about both cost savings and safety improvements were made based on the content analysis from these focus groups. In order to create a practical method combining strengths of both data-driven modeling and human expertise, we developed an iterative process for refining the computational modeling using feedback from a human domain expert. We explore the questions of whether efficiencies in the analytic process can be gained and assess the effects of our combined human-machine process on quality and rigor. Our findings show promise for the use of a topic modeling approach, which has implications for both public and private-sector firms that use qualitative methods such as focus groups and in-depth interviews.

Presentation Title	**Large Scale Reproducibility**
Presenter Name	**Richard Schwinn**
Presenter Email	**richard.schwinn@sba.gov**

The presentation discusses the history of reproducible research, its potential outside academia, and its implementations. The audience is provided with real-time demonstrations of reproducibility techniques that culminate in the production of high quality final products built on government APIs. Reproducible research refers to analyses published with both source data and code so that others can easily verify the findings and build upon them. The SBA Office of Advocacy's Small Business Profiles is an annual report on the state of America's small businesses. Its reproducible 2016 redesign illustrates why now is the time for researchers to begin using the tools of reproducibility. The languages of R, Knitr, LATEX, and Markdown make calling government APIs, repeating analyses, and embracing aesthetic principles straightforward and easy. The suggested workflow offers a template for future researchers who want to meet the highest standards in transparency and aesthetics while efficiently managing their resources.

Session 4B: Survey Design Issues

Presentation Title	**Influence of Multiple Factors on Response Rates**
Presenter Name	**Brad Chaney**
Presenter Email	**BradChaney@westat.com**

Using built-in experiments and paradata, we examine the relative influence of multiple factors on response rates: prepaid incentives ($0, $2, and $5), variations in contact approaches (e.g., regular versus priority mail), web vs. paper surveys, quality of the contact data, relevance of the survey topic to respondents, questionnaire design (formatting and the ordering of questions), and length of the questionnaire. These factors are considered both alone and in combination, allowing us to examine: (1) which factors have the greatest impact, (2) the additive impact when all conditions are optimal, and (3) how some factors influence others (e.g., the quality of the contact data may influence the effect of providing financial incentives). Some results were surprising. For example, while most respondents preferred the web format over the mail format, and while more web responses were received than mail responses, the web format received higher response rates if people first received a paper questionnaire. Also, skip patterns appeared to lower the perceived relevance of the survey, resulting in lower response rates for mail surveys but not for web surveys. Based on these findings, we conducted a second round of data collection using a reserve sample, getting much higher response rates. This paper includes data from both rounds. The National Survey of Small Businesses, sponsored by the National Center for Science and Engineering Statistics within the National Science Foundation, was of 5,000 businesses with between one and nine employees, using a sample stratified by business organization type, industry, and number of employees. This audience is a difficult one to reach because of high mobility and sensitivity concerning the time required to complete the

survey. Data collection was conducted by mail, with mail, email, and telephone follow-up.

Presentation Title	**Language Negotiation on a Bilingual (English/Spanish) Telephone Survey in the U.S.**
Presenter Name	**Casey Langer Tesfaye**
Presenter Email	**casey.tesfaye@nielsen.com**

The negotiation of language in telephone interviews where there is potential for a bilingual respondent is often debated. Should telephone interviews begin in Spanish if data indicates the possibility of reaching a Hispanic respondent or should they always begin in English? Each strategy has its advantages and disadvantages. Starting in Spanish can help Spanish speakers feel comfortable more quickly but may drive some English speakers away, and starting in English can have a wider appeal but draw less Spanish respondents. However, the language negotiation process is more nuanced and less conscious than these debates would have us believe. Not only is the first conversational turn taken by the interviewee and not the interviewer, language choice is ultimately a negotiation that involves both speakers. For this study, we examine a corpus of over 300 phone calls from a bilingual study to behavior code the mechanics of the language negotiation process. For each call the language of the first and subsequent conversational turns are coded as well as whether there was any shift in language during the conversation. We then take observational notes about interactions that involve more nuanced negotiation or any shift in language. In this paper, we will first present the overall findings about language choice and language shifting and then explore a selection of phone calls in greater depth to discuss the effect of language negotiation strategies on the progress and outcome of the call. The goal of this research is to develop data-driven, practical recommendations regarding language negotiation in telephone interviews.

Presentation Title	**When can we call? Experiment to assess SMS text to prompt response across cultures**
Presenter Name	**John Lee Pratt Holmes**
Presenter Email	**johnleeholmes@gmail.com**

This paper examines a split sample experiment that alternated a Short Message Service (or SMS) advance message across seven language and three socio-economic groups surveyed by telephone in Qatar. Qatar's unique advantages include: a small size that facilitates population enumeration, universal cellular coverage, a multicultural respondent base, and a legal environment that permits such public contact techniques. Two years ago Qatar University's Social and Economic Survey Research Institute (SESRI) began conducting its own quarterly survey by phone to measure variation in quality of life and consumer confidence among its Qatari, high income expatriate ("white collar" and low income migrant ("blue collar") populations. Conducted using a random sample drawn from cellular telephones in Qatar, the survey has high response rates (over 50%) compared to surveys in the U.S. and Europe. Nevertheless concern about survey fatigue within a small population and the increasing use of avoidance features on smart phones has led SESRI to pre-emptively encourage positive response. To appreciate whether such a tactic would have the desired effect, advance text contact was launched via a split sample experiment. Half of the sample received an advance text sent out using Qatar University SMS system and half did not. We measured whether this increased blocking of SESRI's number, variance in contact and cooperation rates, and differences in the composition of the respondent pool by language group and respondent type (Qatari, white collar, blue collar). We believe this experiment has value in verifying whether such a method changes the composition of those who participate in the survey to better match the actual population. Because of the multiple language groups, it is applicable to broader cross-cultural survey quality concerns for other countries that may also seek to leverage widespread use of cellular and smart phone technology to understand popular preferences.

Presentation Title	**Differences in Mortality Across NHANES Survey Periods**
Presenter Name	**Yutaka Aoki**
Presenter Email	**yaoki@cdc.gov**

  We report on mortality of adult participants in National Health and Nutrition Examination Survey (NHANES) 1999-2010 followed up through 2011. Survival curves for all-cause deaths for each 2-year NHANES survey period drawn with age as a time scale appeared to diverge over age, with older adults from later survey periods having better survival than older adults from earlier survey periods. This divergence disappeared with data restricted to follow-up < 3 years. Survival curves redrawn by follow-up duration revealed that shorter follow-up is associated with better survival, indicating the observed survey period-mortality association may be confounded by follow-up duration. We further investigated whether the observed pattern was explained by improved mortality trends for older adults or by follow-up duration using a Cox model. Preliminary analyses revealed that differences by survey period may be explained by the association between follow-up duration and mortality found in older, not in younger, ages, which is independent of survey period or calendar time of death event.

Poster Session 1

Poster Title	**Effectiveness of the Think-Aloud Method in Children of Different Age Groups**
Presenter Name	**Mila Sugovic**
Presenter Email	**sugovicm@eurekafacts.com**

This study evaluates the effectiveness of the think-aloud method used during cognitive interviewing for survey-item testing with children. This study examines the application of a think-aloud technique during testing of a computer-based assessment, across two different grades and age-groups: grade 4 (age 9-10) and grade 8 (age 14-15). We compared the effectiveness of the think-aloud method between the two grades by evaluating the verbal reports length, time of response, and the quality and the completeness of the student responses. We found that the think-aloud method does not vary in the quality of the verbal output obtained from children in the two age groups. Although research suggests that young children typically find the think-aloud method to be difficult (Someren et al, 1994), this research shows that there is no difference in the effectiveness of the think-aloud method between students in Grade 4 and in Grade 8.

Poster Title	**Targeting Asian Subpopulations in a Cell Phone Random Digit Dial Survey: Using Census Data to Improve the Usefulness of Billing ZIP Codes**
Presenter Name	**Michael Jacobsen**
Presenter Email	**mjacobsen@rti.org**

Geographic targeting of racial and ethnic populations with cell phone samples is difficult due to the relative lack of geographic information on the sampling frames. The billing ZIP code of the cell phone account assigns cell phone numbers to a geographic area, and are a potential way to target subpopulations, thereby reducing coverage bias. However, about 50% of records with billing ZIP codes report living in their billing ZIP codes (Dutwin, 2014). This limits the ability of billing ZIP codes to target subpopulations. By combining billing ZIP codes to Census geographic regions, using American Community Survey (ACS) data may improve the utility of billing ZIP codes in targeting subpopulations. This study will examine the ability of Census data in augmenting the efficacy of billing ZIP codes in targeting a subgroup in a cell RDD survey. Specifically, this study will examine how appending ACS data can improve the ability of billing ZIP codes in targeting the Asian subpopulation in a cell phone survey. The USPS ZIP Crosswalk Files produced by the Department of Housing and Urban Development estimate the proportion of residential addresses in a Census region that belong to each ZIP code that crosses that region. The proportion of Asian households in each ZIP code were calculated using these files and the 2013 5-Year ACS Asian household tract-level counts. The ZIP codes were then categorized into quartiles based on their proportions of Asian households. Cell phone respondents to the 2015-2016 California Health Interview Survey who both indicated they belonged to an Asian ethnicity and had a billing ZIP code were matched by billing ZIP code to the quartiles. Initial findings show that 75% of these Asian respondents were in the top 50% of ZIP codes. However, non-Asian respondents with billing ZIP codes were evenly distributed among the quartiles.

Poster Session 2

Presentation Title	**Men's Attitudes Towards Abortion: Are Men More Likely To Be Pro-Choice Or Pro-Life In The U.S.?**
Presenter Name	**Rachel Hassan**
Presenter Email	**rachel_34@msn.com**

Men make up the majority of the elected representatives in the U.S. and have a direct hand in the passing of new abortion laws or making changes to current ones. Men also make up about half of the voting population who elect those representatives. Therefore, it is important to know if men are more or less supportive of abortion because men can influence current or future abortion laws and rights for women. Although many studies have included the opinions of both men and women or of women alone, very few studies have been dedicated to men's opinions on the issue of abortion in the United States. The few studies that do exist are outdated. Data from the 2014 General Social Survey (GSS) and previous GSS data were used to examine if men are more pro-choice or pro-life, the characteristics that account for their different opinions, and how male views on this topic have changed over time. As demographic and socioeconomic factors are major determinants of people's attitudes and opinions, this research specifically focused on: race, age, education, region of residence, political party affiliation, religious affiliation, and frequency of religious service attendance as independent variables. When the seven abortion opinion questions that respondents were asked were pooled together, it was found that men were more pro-choice than pro-life. Men who were white, aged 46-60, educated (associates degree or higher), lived in the West region, had no religious affiliation, and who attended religious services infrequently (nearly every week or less often) were more likely to be pro-choice. Men of other races (nonwhite and

nonblack), who were aged 18-30, had a high school diploma or less, lived in the South region, were Republican, and attended religious services frequently (every week or more) were more likely to be pro-life. When the abortion questions were again pooled together, it was found that men's pro-choice stance on abortion has decreased over the past 40 years.

Presentation Title	**Evaluation of Algorithms Used to Determine Diabetes Type**
Presenter Name	**Sarah Lessem**
Presenter Email	**slessem@cdc.gov**

Most surveys that provide population-level estimates of diabetes either do not ask diabetes type or rely solely on self-reported diabetes type with no ability to verify type. Instead, researchers using these survey data rely on algorithms to categorize respondents as having type 1 diabetes (T1DM) or type 2 diabetes (T2DM) based on responses to survey questions. These algorithms incorporate some combination of respondent's age at diabetes diagnosis, insulin use, time between diabetes diagnosis and insulin use, use of oral agents, body mass index (BMI), and self-reported diabetes type. This study used 2005-2014 multi-year data from the National Health Interview Survey (NHIS), National Health and Nutrition Examination Survey (NHANES), and California Health Interview Survey (CHIS) to examine how the weighted percent of respondents identified as having T2DM changes by algorithm and calculated kappas to measure agreement between algorithms both overall and by race. Significant variation exists in the percent identified as having T2DM by algorithm (85.6%-96.2%) and in every combination of algorithms in all three datasets showing much poorer agreement for minorities (.12-.33) than for Whites (.30-.62). An algorithm in which respondents who report T1DM and using insulin are typed as having T1DM whereas those who report T2DM or report T1DM and no insulin are typed with T2DM appears to be most accurate. This algorithm has higher kappas (.36-.52) with all other algorithms than any other combination of algorithms (.22-.29). Compared with other algorithms, agreement is higher both for minorities and whites and the weighted percent of people with diabetes that this algorithm identifies as having T2DM (94.0%) is in line with what the literature states (90%-95%). Further, this algorithm allows adults over age 30 to be typed with T1DM which is not possible using algorithm that restrict typing with T1DM to those diagnosed under 30. It also acknowledges that not all people with T2DM are overweight nor are all those with T1DM at normal weight.