Did it Work? Findings from a Flu Pilot Study Using Interactive Voice Response (IVR) and Live Interviewers

Vicki J Pineau, Benjamin Skalland, and
Gillian Lawrence
NORC at the University of Chicago

Abstract

NORC at the University of Chicago funded a pilot survey collecting seasonal flu vaccination information using a redirected in-bound telephone call sample from Reconnect Research, and compared flu vaccination estimates to CDC-published results for the 2016-2017 flu season. This non-probability sample, called MIDI Calls™, comes from intercepted telephone calls from people who mis-dialed a telephone number or experienced an incomplete or disconnected call. MIDI Calls are forwarded from telephone companies to Reconnect Research for research and marketing purposes; this system offers both an opportunity to access a large non-probability based sample of inbound calls and to concomitantly reduce telephone data collection costs by using Interactive Voice Response (IVR) interviewing. In early 2017, NORC fielded a pilot study using a sample of MIDI Calls and experimented with two telephone modes -- IVR alone and IVR routed to live interviews -- to conduct a telephone survey to estimate seasonal flu vaccination coverage rates for four target populations. Pilot survey goals were to: 1) evaluate the cost and quality of using the MIDI Calls for fit-for-purpose surveys of a general population as well as for hard-to-reach subpopulations; and 2) assess whether MIDI Calls could be routed successfully to NORC for telephone interviewer administration. The survey collected seasonal flu vaccination information for adults, children, pregnant women, and health care personnel. We will present seasonal flu estimates from the pilot for these subgroups by experimental mode treatment, along with comparisons to flu estimates published by CDC for the 2016-2017 flu season.

Key Words: Interactive Voice Response (IVR); Live Interviewers; Misdialed, Incomplete, Disconnected, Inbound Call Sampling (MIDI Calls™); Nonprobability Sample

1 Introduction

NORC's MIDI Flu Pilot project was internally-funded in February 2017 to explore the usability and feasibility of a new source of potential survey respondents offered by the company Reconnect Research (RR).[1] RR re-directs incoming telephone calls from telecom companies that are Misdialed, Incomplete, Disconnected, and Inbound (MIDI Calls™) for marketing and research efforts, which includes conduct of surveys. Research has been conducted in recent years to evaluate the representativeness of the MIDI sample as it has the potential to be a cost-effective telephone-based sample source. Published research is mixed in the findings of sample quality for health measures with larger than desired differences in some measured survey outcomes (such as for change in health insurance) relative to trusted benchmarks but promising results for others such as ever smoked cigarettes. (Levine & Krotki, 2017) A Pew study using the MIDI sample for a survey of public opinion found political party ID distributions did not match very well to the General

[1] https://www.reconnectresearch.com/

Social Survey. (Kennedy, Hatley, & McGeeney, 2017). A Westat study found use of the MIDI sample in a health survey for estimates of health status, ever smoked, exercise in past 30 days, sometime/seldom/never seatbelt use to vary between 5 to over 20 percentage points. (Dipko & Jodts, 2017) Notwithstanding the mixed published research results on the quality of the MIDI sample, we were interested in assessing the quality of the MIDI sample for estimating national estimates of flu vaccination of important rare population subgroups as well as whether alternative modes of data collection would be possible with the MIDI phone as an approach to reach rare populations at substantial cost reduction to traditional probability based methods.

If we consider MIDI calls as a sampling frame, then the calls can be sampled and redirected to a phone survey conducted using interactive voice recognition (IVR) and/or using live interviewers. The pilot project had two related goals towards assessing the viability and representativeness of using the MIDI call sample for rare population surveys:

1) Explore the operational feasibility of using RR's MIDI calls, especially the transference of MIDI calls to NORC interviewers for CATI administration using a split sample experimental design; and
2) Evaluate demographic, geographic, and seasonal flu estimates produced using this methodology compared to those based on traditional methods (i.e., NORC telephone interviews) as well as to national benchmark estimates (e.g. NIS-Flu, NHIS, ACS).

This paper begins with an overview of the pilot methodology that relied on the MIDI calls sampling frame and how it was used in this pilot study, including a description of the imbedded experiment to compare IVR mode of interview alone to IVR transferred to live interviewer mode and the associated questionnaire development. We then present the analysis of fielding indicators such as response rates by interview mode treatment and the analysis of survey outcomes for socio-demographics and flu vaccination estimates. We end with an operational discussion of using the MIDI calls frame to conduct a health-related survey focussed on rare population subgroups.

2. Methodology of Flu Pilot Study

RR is a company that provides a non-probability sample of intercepted telephone calls via MIDI calls. The MIDI calls originate from telephone companies that forward the calls to RR for use in research and marketing purposes. As mentioned earlier, these MIDI calls offer a sampling frame from which a sample can be drawn and a telephone survey conducted.

For the Flu Pilot study, a convenience sample of MIDI calls were used to field a survey with questionnaire content focussed on flu vaccinations for the 18+ population and subsets of the population consisting of health care personnel, pregnant women and households with young children. We fielded the survey using two treatment interviewing "modes": interactive voice recording (IVR) and live telephone interviewer to assess the viability of transferring the incoming MIDI calls to NORC live interviewers relative to completing the survey only using IVR.

All sample MIDI calls started out in the IVR mode for screening with a few "softball" questions, and then continued into the main interview in one of two modes. Part of the sample continued with the RR IVR administration, and the remaining sample was routed to NORC for "Live Interview" administration via Computer Assisted Telephone Interviewing (CATI). The mode of administration was determined before the call started in the IVR system, although calls were randomly allocated. We aimed to complete 2,500

cases in each mode: IVR only and IVR to CATI. Cases completed via the IVR had the option to complete in English or Spanish; only English-speaking respondents were transferred to NORC interviewers.

NORC worked collaboratively with RR to determine the number of calls that should be routed to NORC's interviewers per hour during the fielding period, as well as to estimate the number of interviewers that would be required to answer the calls live, while minimizing the number of calls that were not able to be answered with an interviewer.

2.1 Questionnaire Development

To develop the MIDI Flu Pilot questionnaire, we first compiled the list of socio-demographic, health and flu vaccination questions we would need in order to conduct the analyses of the pilot data and embedded experiment. We then compiled a list of available questions from existing studies that would yield the relevant data points. The use of existing measures ensures the items are reliable and valid; additionally, they provide points of comparability when comparing the data obtained from this study from those other nationally administered studies. Key data points included: flu vaccination status and timing for current flu season, healthcare worker status, pregnancy status, demographics, and child flu vaccination status and timing.

When selecting the final questions to include in the survey instrument, the team considered not only the data points of interest, but also the mode of administration and length of the interview. The final set of questions were drawn from the National Immunization Survey (NIS), the Behavioral Risk Factor Surveillance System (BRFSS), and the National Health Interview Survey (NHIS). The questions selected for administration were those that would: 1) be feasible to administer in IVR or with a live interviewer and 2) allow NORC to compare estimates from this dataset against the flu estimates generated by the CDC.

Once we had determined the content of the questionnaire, we worked in collaboration with RR to refine it and ensure the questions would work well for IVR administration. Based on input from RR, we included use of prompts termed "atta-boys," which are short sentences read to respondents during the IVR interview to encourage them to continue, such as "You're doing great!" The initial draft of the questionnaire we developed had several skips, response confirmation questions, and one series of looping questions. These varying types of questions allowed us to explore the robustness of RR's IVR programming capabilities.

Some modifications to question text were necessary for ease of administration via IVR. For example, "Read if necessary" interviewer instructions were folded into the question text itself as the IVR system would not be able to judge if these instructions were needed. Additionally, it was deemed too time-intensive to program the IVR system to capture multiple races for respondents for the pilot test; therefore, instead of looping through the categories to capture each race, an option for "multiple races" was added to the race question for IVR.

The questionnaire was administered in both English and Spanish if conducted via IVR, but only English if sent to a live interviewer for cost reasons, not technical reasons. Cases that started in Spanish but marked for delivery to a live interviewer were routed to a closing script.

2.2 Fielding the Pilot with Embedded Interview Mode Experiment

Surveys across both IVR and Live Interview modes (two experimental treatments) were conducted during regular NORC business hours (9AM-9PM CT). We fielded the survey in two phases—test and production. The test phase, conducted on 2/7/17, tested system set up and looked at overall call volume vs. interviewer staffing. Interviews completed during the test phase are counted in our overall completed interview totals.

The production phase of the survey occurred over a 6-day period from Friday February 17th through Wednesday February 22nd, 2017. By design we included one weekend day (Saturday) in order to include times that outbound CATI calls would normally take place at our call centers. We conducted no interviews in either mode on Sunday 2/19/17 due to the complexity of fielding over the weekend as well as there being a much lower volume of MIDI calls on weekends. Although we aimed to keep the field periods between the two modes as consistent as possible, on the final day of the production phase (2/22/17) the IVR-only mode continued for several hours after all the Live Interviews were completed earlier in the day. We had to extend the IVR-only field period by several hours because RR had lower than expected MIDI call volume both over the weekend and on Monday, 2/20/17, which was a holiday. RR had fewer calls overall in which to field the survey; during this time period they sent more calls to NORC and fewer to the IVR-only survey in order to keep up with planned NORC interviewer staffing.

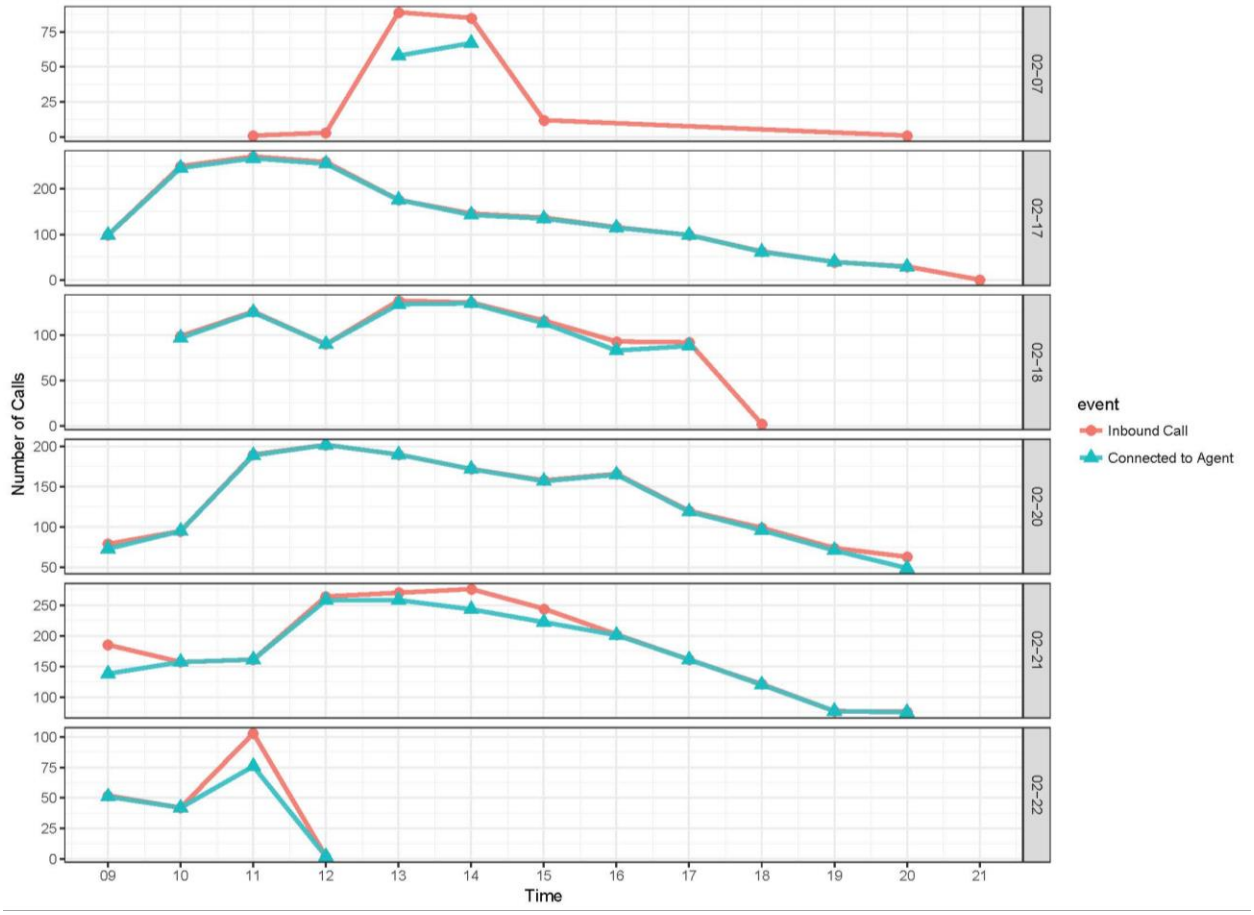
The biggest challenge during fielding was finding the best flow rate of cases from RR to NORC interviewers because we wanted to keep our answer rate (percent of all cases sent to NORC that were answered by an interviewer) as high as possible. In preparation for the field period we worked with RR to create a staffing plan based on the target number of completes and the projected availability of MIDI calls. We then staffed interviewers accordingly. The staffing model for pure inbound calls of this type differed significantly from our staffing model for typical outbound dialing efforts in that the calls peaked during the middle of the day rather than during evening hours.

During the test phase we were able to see that the call volume coming from RR to NORC was still too high and were able to adjust our plans further. Once production began, our call center project manager and team were in frequent contact with the RR team to fine-tune the rate of call delivery to minimize both the number of unanswered calls due to all interviewers being on other calls and the wait time between taking calls.

Exhibit 1, below, shows that during the test phase the number of inbound calls exceeded the number of interviewers available to answer the calls live; as we adjusted over time this gap narrowed significantly.

The percent of calls answered live varied by day, but after the test phase on 2/7/2017 remained high, and overall was 96%. Table 1 shows the live answer rate by day.

Exhibit 1: Inbound Calls vs. Calls Connected To a Live Interviewer by Hour



2.3 Key Indicators from Fielding

In what follows, we divide the data collection process into two stages. Stage 1 refers to the first set of IVR-only questions asked of all callers. Stage 2 refers to data collection after the completion of Stage 1, either through transfer to a live NORC interviewer or continuing in the IVR-only mode. Table 2 presents key indicators for MIDI data collection, including completion of Stage 1 of the interview, completion of the adult and child sections of Stage 2 of the interview, and overall completion. The results are presented for those numbers that would continue in IVR for Stage 2, for those numbers that would be transferred to NORC interviewers for Stage 2, and overall.

As presented in Row A of Table 2, over 140,000 incoming phone calls were made part of this study by RR. Of those, 56 percent were identified as landline phone numbers by RR, 35 percent were identified as cell-phone numbers, and 9 percent could not be determined to be landlines or cell phones. RR's method for determining whether the number is a cell phone vs. landline is based on the area code and prefix of phone number (NPA-NXX); those numbers with area codes and prefixes owned by landline carriers are designated landlines, and those with area codes and prefixes owned by cell-phone carriers are designated cell phones. This determination does not account for the porting of landline numbers to cell phones or vice versa. Of the incoming phone calls, 10.2 percent completed Stage 1 of the interview (Row B of Table 2); that is, 10.2 percent answered the questions about general health and insurance coverage administered via IVR.

Of those completing Stage 1, 5,126 respondents continued in IVR for Stage 2 and 9,260 respondents were routed to NORC for Stage 2 (Row C of Table 2). However, only 69.8 percent of those routed to NORC actually had contact with an NORC interviewer; the remainder either hung up before connecting with an NORC interviewer or an NORC interviewer was not available to take the call (Row D of Table 2). While 84.5 percent of those remaining in IVR for Stage 2 answered the first question in Stage 2, only 49.8 percent of those that had contact with an NORC interviewer answered the first question in Stage 2 (Row E of Table 2).

Of those answering the first question in Stage 2, 82 percent overall completed the adult section of the survey, meaning they answered the adult flu and demographic questions (Row F of Table 2). This percentage was very similar between those in IVR and those speaking with an NORC interviewer.

Of those completing the adult section, 93 percent overall completed the child screener, meaning they indicated whether or not there were children under age 18 in the household (Row G of Table 2). Of those in IVR, 89.9 percent did so, and of those speaking with an NORC interviewer, 97.9 percent did so. Of those completing the child screener in IVR, 32.5 percent indicated they had a child under age 18 versus 23.8 percent of those speaking with an NORC interviewer (Row H of Table 2). Of those with a child under 18, 92.6 percent of those speaking with an NORC interviewer completed the child section of the interview, but only 48.1 percent of those in IVR did so (Row I of Table 2). This large difference in the child section completion rate may be due to the way the IVR questionnaire was programmed. It appears that the IVR required respondents that indicated a child in the household to give an age between 0 and 17 years for the selected child; based on the data, it is likely that there was no way for the IVR respondent to refuse to give the child's age

and still continue with the interview, and there was no way for the respondent to back up and change previous answers. Therefore, much of the IVR nonresponse to the child section may have been due to the limitations of the IVR instrument rather than a real difference in respondents' willingness to answer questions about a child via IVR versus when speaking to an interviewer.

Overall, of those numbers that continued in IVR for Stage 2, 49 percent completed Stage 2, and of those numbers transferred to NORC for Stage 2, 27.2 percent completed Stage 2 (Row J or Table 2).

Of the initial 52,256 incoming calls to RR designated to remain in IVR, 6.8 percent resulted in a completed adult section and 0.9 percent resulted in a completed child section; and of the initial 88,591 incoming calls to RR designated to be transferred to NORC interviewers, 3.0 percent resulted in a completed adult section and 0.6 percent resulted in a completed child section (Rows K and L of Table 2). That is, the yield rates of adult completes and child completes were lower for numbers that would be transferred to NORC, mainly due to a loss of respondents when making the transfer.

3. Analysis of the Pilot Data

Table 3 presents unweighted distributions of survey responses to the health and demographic questions among respondents completing the adult section of the questionnaire, first for those responding via IVR and then for those responding to NORC interviewers. The distributions are presented both including and excluding the missing values, and the distributions excluding missing values are then compared to estimates from the 2015 National Health Interview Survey (NHIS) or the 2015 American Community Survey (ACS).

Before examining the distributions themselves, we first note that the rate of item nonresponse was much higher in IVR than when the questions were asked by NORC interviewers. "Item nonresponse" here includes refusals to answer, "don't know" responses, and responses that are outside of the valid code frame for the question. The IVR did not restrict responses to the code frame; for example, if the response options were "Press 1 for YES, Press 2 for NO, Press 3 for DON'T KNOW, and Press 4 for "REFUSE," the respondent could still press 0, 5, 6, 7, 8, or 9; if the respondent pressed a number outside of the code frame, the question was re-asked. Still, even after incorporating the responses when the question was re-asked, many IVR responses remained outside of the code frame and are counted as item nonresponse. Looking at Table 3, we observe that item nonresponse rates were often 10 times higher (or more) in the IVR than with NORC interviewers.

3.1 Health Distributions and Identification of Rare Subgroups

We now examine the unweighted MIDI response distributions, excluding the missing values due to item nonresponse. The general health and insurance coverage questions were administered via IVR to all respondents in Stage 1 of the interview. The unweighted distributions of responses to these questions are fairly similar to estimates from the 2015 NHIS; 59 percent of MIDI respondents reported they were in Excellent or Very Good Health versus a 60.5 percent estimate from the 2015 NHIS, and 83.5 percent of respondents reported they had insurance versus an estimate of 89 percent from the 2015 NHIS (but note

that here insurance status was established via a single question, whereas the NHIS uses a battery of questions, asking about each type of insurance separately, and therefore NHIS might be expected to result in a higher estimate of insurance coverage.) A higher proportion of MIDI respondents identified their health as Fair or Poor (18 percent) compared to the 2015 NHIS (12.6 percent).

In Stage 2, the unweighted adult demographic distributions were for the most part similar between those completing via IVR and those completing with NORC interviewers. However, there are some exceptions, and the unweighted distributions often differed from the NHIS or ACS benchmarks.

A higher proportion of MIDI respondents identified themselves as health care workers (12 percent) than in the 2015 NHIS (8 percent).

While the overall rate of current pregnancy among females was similar in MIDI and in the NHIS (3.4 percent versus 2.0 percent), the rate differed between IVR (5.5 percent) and live interview respondents (1.1 percent).

3.2 Demographics

3.2.1 Adult Demographic Distributions

The MIDI respondents were more likely to be female (62 percent) than in the population of adults (NHIS estimate of 52 percent).

The MIDI respondents skewed older than the ACS benchmark. Of MIDI respondents, 27 percent were under the age of 40 compared with an ACS estimate of 38 percent, and 31 percent of MIDI respondents were over the age of 65 compared with an ACS estimate of 19 percent. The MIDI age distribution also differed for those completing via IVR versus those completing with NORC interviewers; those completing with live interviewers tended to be older than those completing via IVR, likely due to differential nonresponse across the two modes.

MIDI respondents were much more likely to be non-Hispanic Black alone (21 percent) compared with the ACS estimate (12 percent), and were less likely to be non-Hispanic White alone. IVR respondents were more likely to be Hispanic (17 percent) than those responding with NORC interviewers (11 percent); this difference may be due to the fact that NORC interviews were conducted only in English, were as IVR interviews were conducted in both English and Spanish.

Finally, IVR respondents were more likely to report the presence of children in the household than those responding with NORC interviewers; that difference maybe be due in part to the fact that live interviewer respondents skewed older, but may largely be due to differential item nonresponse for these questions in IVR versus with NORC interviewers. Of those that did report children in the household, those completing via IVR were much more likely to report a higher number of children; how much of this is due to real differences between IVR respondents and those responding with NORC interviewers and how much is due to measurement error within the IVR is uncertain.

3.2.2 Child Demographic Distributions

Table 4 mirrors Table 3 but examines child demographic distributions for those completing the child section of Stage 2 rather than adult demographic distributions. As with the adult questions, item nonresponse was higher for IVR than when NORC interviewers administered the survey. Excluding missing values, we find the child gender distribution was similar between IVR and NORC interviewer administration, and similar to the 2015 ACS estimate.

MIDI children skewed slightly older compared to the ACS benchmark, and the children for those completing with NORC interviewers skewed older compared with the children for those completing via IVR.

As with the adults, the MIDI children were much more likely to be non-Hispanic Black alone and less likely to be non-Hispanic White alone, and the IVR children were more likely to be Hispanic than the children of those completing with NORC interviewers, possibly because the NORC interviews were conducted only in English.

3.3 Flu Vaccination Rate Estimates

3.3.1 Adult Vaccination Rates

Table 5 presents estimates of receipt of 1 or more flu vaccinations since July1, 2016 overall and for demographic subgroups, based on respondents completing the adult section of Stage 2 of the questionnaire. These estimates are presented for those completing via IVR, those completing with NORC interviewers, and overall. All of these estimates exclude those with item nonresponse to the question about receipt of a flu vaccination.

The estimates are first presented on an unweighted basis and then presented using post-stratified weights. The weights were developed by first assigning each respondent to the adult section a base weight of 1, and then post-stratifying these weights to the three-way cross-classification of age category by race/ethnicity by sex. The age categories were 18-29, 30-39, 40-49, 50-64, and 65+; the race/ethnicity categories were Hispanic, non-Hispanic Black alone, and all others. The population controls for the 30 post-stratification cells were derived from the 2015 ACS. The weights were created separately for those completing via IVR, those completing with NORC interviewers, and overall.

External estimates for adult flu vaccination rates come from several different sources. The CDC benchmark for flu vaccination for all adults 18+ is from the BRFSS[2], estimating vaccination since July 2016 with a data collection period from September 2016 – June 2017. CDC flu vaccination rate benchmarks for Health Care Personnel and Pregnant Women are based on non-probability Internet panel survey, estimating vaccination since July 1, 2016 with a data collection period from March 28 – April 19, 2017 for Health Care Personnel, and estimating vaccination since July 1, 2016 with a data collection period from March 28 – April 7, 2017 for Pregnant Women.[3][4]

[2] https://www.cdc.gov/flu/fluvaxview/coverage-1617estimates.htm
[3] https://www.cdc.gov/mmwr/volumes/66/wr/mm6638a1.htm?s_cid=mm6638a1_w
[4] https://www.cdc.gov/mmwr/volumes/66/wr/mm6638a2.htm?s_cid=mm6638a2_w

MIDI estimates of adult flu vaccination rates varied greatly between those based on IVR interviews and those based on interviews with NORC interviewers. On an unweighted basis, the IVR estimate is 22 percent and the NORC interview estimate was 45 percent. The weighted estimates are similar to the unweighted estimates. The difference between the IVR and live interviewer estimate is present for every subgroup examined, but was greater for older respondents and non-Hispanic White respondents. The difference also grows when examining the cumulative vaccination rate estimate by flu season, based on the reported month of vaccination; as of the end of September, there is little difference between the IVR and live interview estimate, but as of the end of January, the difference has grown to over 20 percentage points.

The vaccination coverage rates for health care personnel and pregnant women also differed greatly between IVR and live interviewers, differences of 12.4 and 14.8 percentage points respectively between the two modes of interviewing. Both IVR and live interviewer estimates for health care personnel and pregnant women were substantially lower than the CDC benchmarks.

3.3.1 Child Vaccination Rates

Table 6 mirrors Table 5, but present flu vaccination rate estimates for children based on respondents completing the child section of Stage 2 of the interview. Table 6 presents estimates of 1 or more flu vaccinations since July1, 2016.

As with the adult flu vaccination rate estimates, the child vaccination rate estimates exclude those with item nonresponse to the flu question and are presented on an unweighted and weighted basis. The weights were post-stratified to the three-way cross-classification of age category by race/ethnicity by sex. The age categories were 0-4, 5-12, and 13-17; the race/ethnicity categories were Hispanic, non-Hispanic Black alone, and all others. The population controls for the 18 post-stratification cells were derived from the 2015 ACS, and the weights were created separately for those completing via IVR, those completing with NORC interviewers, and overall.

Whereas the adult flu vaccination rate estimates were much higher when based on those completing with NORC interviewers, for children the estimates are much more similar between the two modes. Sample sizes for the child estimates are much smaller than for the adult estimates, and therefore the child results must be interpreted with more caution, but overall the unweighted child vaccination rate estimate for receipt of 1 or more doses was 44 percent based on those completing with live interviewers versus 38 percent based on those completing via IVR. That difference is present for both males and females and for all age groups, but is not present when looking only at Hispanic or non-Hispanic White children. Weighted estimates differ somewhat from the unweighted estimates, but the differences by mode remain.

Compared to the corresponding NIS-Flu estimates for week ending February 25, MIDI estimates of 1 or more flu vaccinations since July1 are lower based on either the IVR interviews or the live interviews. The overall estimate from CDC is 54 percent compared to 34.5 percent for MIDI IVR interviews, 46.5 percent for MIDI live interviews, and 42 percent for MIDI overall. Those differences in estimates from MIDI IVR versus CDC are present for all age subgroups they are also present from MIDI live interviews versus CDC for the 0-4 and 5-12 age groups. However, the difference is not present when looking at

the estimate for 1 or more flu vaccinations received by the end of September; the difference begins with the end-of-October estimates and grows from there.

3.4 Geographic Location

RR assigned each incoming call to a state, presumably based on the area code of the phone number. Table 7 presents the unweighted distribution of these incoming calls (including both respondents and nonrespondents) across states, as assigned by RR, and compares this distribution to the distribution of U.S. adults across states, as estimated using the 2015 ACS.

First note that 4.4 percent of the incoming calls either could not be assigned to a state by RR, or were assigned to a U.S. territory. Excluding those calls, the distribution of calls that could be assigned to a state closely matches the 2015 ACS distribution of U.S. adults across states, except for Florida. According to the 2015 ACS, 6.6 percent of U.S. adults live in Florida, but 12.2 percent of the incoming calls to RR were assigned to Florida.

4. Summary and Future Considerations

Conduct of the MIDI Flu pilot was a successful endeavor. We found that using the MIDI sample, or more generally an intercept sample of telephone numbers, was proved feasible for the conduct of general population and subgroup populations using either IVR alone or IVR transferred to a live interviewer. We were able to successfully transfer calls using RR's IVR system to NORC interviewers. NORC and RR's staff successfully coordinated the scheduling of incoming CATI sample to NORC to ensure sufficient interviewing staff to minimize dropped calls.

We found that item nonresponse in the IVR was substantially higher than item nonresponse with live interviewers, often 10 times higher. This is an area where more work in developing the questionnaire content and logic for IVR is needed to reduce item nonresponse levels.

When comparing the MIDI sample survey outcomes to benchmarks, we observe:

- Higher proportion of MIDI sample identified themselves as health care workers than the benchmark.
- Higher proportion of MIDI sample were female than the benchmark.
- Higher proportion of MIDI sample were non-Hispanic Blacks than the benchmark.
- Higher proportion of MIDI sample reported excellent general health than the benchmark.
- Lower proportion of MIDI sample had health insurance than the benchmark.

When comparing estimates from the IVR completes to live interviewer completes, the adult demographic distributions were mostly similar (gender, health care worker, age). The adult health outcomes between the IVR and live interviewer completes differed for the following:

- Reported incidence of pregnancy was higher for the IVR completes.
- IVR respondents skewed younger than live interviewer respondents.

- Percent Hispanic was higher for IVR respondents. [Recall that the respondents forwarded to live interviewers did not include those who chose to be administered the interview in Spanish.]
- Percent reporting children in the household was higher for IVR respondents.

- The child gender distributions were similar for the IVR and live interviewer completes and were similar to the benchmark.
    - For age, the IVR yielded more children 0-4 years of age than the live interviewer completes and when compared to the benchmark.
    - And, similar to the adults, the IVR yielded a higher percent of Hispanic children and was comparable to the benchmark.

- MIDI estimates for adult flu vaccination were much higher for live interviewer respondents compared to IVR respondents.
    - The difference is present for every subgroup examined, but greater for older respondents and non-Hispanic White respondents.

- MIDI estimates for child flu vaccination between the IVR and live interviewer completes did not vary as much as the adult flu vaccination estimates.
    - However, the live interviewer flu estimates were higher than the IVR estimates overall and by age group.
    - MIDI child flu estimates were consistently lower than the benchmarks on the order of 2-22 percentage points.

The results, though encouraging, indicate that for the key flu vaccination rate outcomes the MIDI sample produces substantially different results as compared to benchmarks from the NIS-Flu survey. Though the MIDI sample may offer a convenience sample for testing different data collection methods (such as Web options for a phone sample), different instruments, different incentive procedures, it does not appear to be a sufficiently representative sample for official national and subgroup flu vaccination estimates.

The most significant benefit of using an IVR-only system operationally is its efficiency. As IVR does not require interviewers to conduct the survey, there was no need to train any staff, produce training materials, or pay for staff time. The IVR survey required very little management time once it was set up. It can be set to run until the desired number of completes is obtained, including overnight or odd hours (as appropriate) when interviewers are typically unavailable.

Areas for further investigation/improvement include (1) developing the questionnaire content and logic for IVR administration to reduce item nonresponse levels; (2) for transference of calls from IVR to live interviews, improving methods for keeping respondents engaged during the transfer process; (3) development and testing of practices to ask questions via IVR more generally; (4) development of methods for handling multiple language administration for the MIDI sample, especially for transferring IVR calls to live interviewers.

4. References

Dipko, S., & Jodts, E. (2017). Redirected Inbound Call Sampling (RICS) –Pilot Test Results and Caller Reactions. *American Association for Public Opinion Research.* New Orleans: American Statistical Association.

Kennedy, C., Hatley, N., & McGeeney, K. (2017). Measuring Public Opinion with Redirected Calling to IVR. *American Association for Public Opinion Research.* New Orleans: American Statistical Association.

Levine, B., & Krotki, K. (2017). Evaluating Bias in a Survey Using Redirected Inbound Call Sampling Methodology . *American Association for Public Opinion Research.* New Orleans: American Statistical Association.

Table 1: Live Answer Rate by Day

Date	Inbound Calls	Connected to Interviewer	% of Calls Answered Live
2/7/17 (test)	191	125	65%
2/17/17	1687	1666	99%
2/18/17	892	865	97%
2/20/17	1608	1578	98%
2/21/17	2194	2071	94%
2/22/17	199	171	86%
Total	**6771**	**6476**	**96%**

Table 2: Key Indicators

Row	Measure	Continue in IVR	Transfer to NORC	Overall
	Stage 1: Initial IVR Questions			
A	Incoming calls to Reconnect Research	52,256	88,591	140,847
	Landline	28,800 (55%)	49,723 (56%)	78,523 (56%)
	Cell Phone	17,889 (34%)	31,473 (36%)	49,362 (35%)
	Unknown	5,567 (11%)	7,395 (8%)	12,962 (9%)
B	Completed Stage 1	5,126	9,260	14,386
	Stage 1 Completion Rate	***9.8%***	***10.5%***	***10.2%***
	Stage 2: Continue in IVR or Transfer to NORC			
C	Continue in IVR vs. Transfer	5,126	9,260	14,386
D	Contact with "Interviewer"	5,126	6,468	10,583
	Contact Rate	***100.0%***	***69.8%***	***73.6%***
E	Answered 1st Question in Stage 2	4,329	3,219	7,548
	1st Question Answer Rate	***84.5%***	***49.8%***	***71.3%***
F	Completed Adult Section	3,533	2,655	6,188

	Adult Section Completion Rate	***81.6%***	***82.5%***	***82.0%***
G	Completed Child Screener	3,172	2,600	5,772
	Child Screener Completion Rate	***89.8%***	***97.9%***	***93.3%***
H	Had a Child Under Age 18	1,032	620	1,652
	Child Eligibility Rate	***32.5%***	***23.8%***	***28.6%***
I	Completed Child Section	496	574	1,070
	Child Section Completion Rate	***48.1%***	***92.6%***	***64.8%***
J	Completed Stage 2	2,510	2,516	5,026
	Stage 2 Completion Rate	***49.0%***	***27.2%***	***34.9%***
	Overall			
	Yield Rate of Adult Completes	***6.8%***	***3.0%***	***4.4%***
	Yield Rate of Child Completes	***0.9%***	***0.6%***	***0.8%***
K	Call Hours	n/a	262	n/a
	Call Hours per Adult Complete	***n/a***	***0.099***	***n/a***
	Call Hours per Child Complete	***n/a***	***0.457***	***n/a***

Table 3: Unweighted Adult Demographic and Health Distributions*

	Distribution (incl. missings)			Distribution (excl. missings)			
Characteristic							
General Health		**Overall**			**Overall**		**NHIS 2015[a]**
Excellent		28.4%			35.5%		28.9%
Very Good		18.7%			23.5%		31.6%
Good		18.2%			22.8%		26.9%
Fair		10.0%			12.5%		9.7%
Poor		4.6%			5.7%		2.9%
DK/Ref/Missing		20.1%					
Insurance		**Overall**			**Overall**		**NHIS 2015[b]**
Yes		62.0%			83.5%		89.4%
No		12.2%			16.5%		10.6%
DK/Ref/Missing		25.7%					
Health Care Worker	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**NHIS 2015[c]**
Yes	10.9%	11.5%	11.2%	12.7%	11.5%	12.1%	8.2%
No	75.3%	88.4%	80.9%	87.3%	88.5%	87.9%	91.8%
DK/Ref/Missing	13.7%	0.2%	7.9%				
Sex	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**ACS 2015**
Male	34.8%	36.7%	35.6%	38.8%	36.9%	37.9%	48.2%
Female	54.9%	62.7%	58.3%	61.2%	63.1%	62.1%	51.8%
DK/Ref/Missing	10.4%	0.5%	6.1%				

Pregnancy (among females 18+)	IVR	Live	Overall	IVR	Live	Overall	NHIS 2015[d]
Pregnant	5.3%	1.1%	3.3%	5.5%	1.1%	3.4%	2.0%
Not Pregnant	90.5%	98.7%	94.3%	94.5%	98.9%	96.6%	98.0%
DK/Ref/Missing	4.3%	0.2%	2.4%				
Pregnancy Since Aug 1 (among females not currently pregnant)	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	
Pregnant	2.9%	1.7%	2.4%	3.3%	1.8%	2.5%	n/a
Not Pregnant	87.3%	96.9%	91.7%	96.7%	98.2%	97.5%	n/a
DK/Ref/Missing	9.8%	1.3%	5.9%				
Age	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**ACS 2015**
18-29	13.8%	11.4%	12.8%	15.6%	11.6%	13.7%	20.7%
30-39	13.0%	11.1%	12.2%	14.6%	11.3%	13.1%	17.2%
40-49	13.8%	12.8%	13.3%	15.5%	13.0%	14.4%	16.9%
50-64	23.9%	28.5%	25.9%	27.0%	29.0%	27.9%	26.0%
65+	24.1%	34.6%	28.6%	27.3%	35.2%	30.9%	19.3%
DK/Ref/Missing	11.4%	1.7%	7.2%				
Race/Ethnicity	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**ACS 2015**
Hispanic	13.8%	10.8%	12.5%	16.8%	11.1%	14.1%	15.6%
Non-Hispanic White	44.8%	57.4%	50.2%	54.7%	59.1%	56.7%	64.7%
Non-Hispanic Black	17.9%	20.4%	19.0%	21.8%	21.0%	21.4%	11.7%
Non-Hispanic Other or Multiple Races	5.5%	8.5%	6.8%	6.8%	8.8%	7.7%	8.1%
DK/Ref/Missing	18.0%	2.9%	11.5%				
Number of Children in HH	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**ACS 2015**
0	24.2%	73.7%	45.4%	46.7%	76.1%	63.9%	68.6%
1	10.0%	10.3%	10.2%	19.4%	10.7%	14.3%	13.5%
2	9.5%	7.7%	8.7%	18.3%	8.0%	12.3%	11.4%
3	3.3%	2.9%	3.1%	6.4%	3.0%	4.4%	4.6%
4-9	4.8%	2.1%	3.7%	9.3%	2.2%	5.2%	2.1%
10+/ DK/Ref/Missing	48.1%	3.2%	28.9%				

* General health and insurance distributions are computed among respondents completing via IVR in Stage 1, i.e., prior to transferring to a live interviewer or remaining in IVR. The remaining distributions are computed among all adults completing the adult section of Stage 2 of the interview.
[a] Weighted NHIS estimate for adults 18+ using person-level file variable PHSTAT, excluding missing values.
[b] Weighted NHIS estimate for adults 18+ using person-level file variables COVER and COVER65, excluding missing values.
[c] Weighted NHIS estimate for adults 18+ using sample adult file variable WRKHLTH2, excluding missing values.
[d] Weighted NHIS estimate for females 18+ using sample adult file variable PREGNOW, excluding missing values. Women age 50+ are assumed not to be pregnant.

Table 4: Unweighted Child Demographic Distributions*

Characteristic	Distribution (incl. missings)			Distribution (excl. missings)			
Sex	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**2015 ACS**
Male	47.8%	48.8%	48.3%	50.1%	49.5%	49.8%	51.1%
Female	47.6%	49.8%	48.8%	49.9%	50.5%	50.2%	48.9%
DK/Ref/Missing	4.6%	1.4%	2.9%				

Age	IVR	Live	Overall	IVR	Live	Overall	2015 ACS
0-4	33.5%	25.9%	29.5%	33.5%	25.9%	29.5%	26.9%
5-12	39.1%	38.0%	38.5%	39.1%	38.0%	38.5%	44.8%
13-17	27.4%	36.0%	32.0%	27.4%	36.0%	32.0%	28.3%
DK/Ref/Missing	0.0%	0.0%	0.0%				

Race/Ethnicity	IVR	Live	Overall	IVR	Live	Overall	2015 ACS
Hispanic	23.0%	17.1%	19.8%	25.3%	17.3%	20.9%	24.6%
Non-Hispanic White	40.5%	45.3%	43.1%	44.6%	46.0%	45.4%	51.4%
Non-Hispanic Black	21.8%	25.6%	23.8%	23.9%	26.0%	25.1%	13.5%
Non-Hispanic Other or Multiple Races	5.6%	10.5%	8.2%	6.2%	10.6%	8.7%	10.4%
DK/Ref/Missing	9.1%	1.6%	5.0%				

* The distributions are computed among children with a completed child section of Stage 2 of the interview.

Table 5: Adult Flu Vaccination Estimates*

	Unweighted Estimates			Weighted Estimates†			External Estimates
Characteristic							
Overall	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**BRFSS**
Overall	21.8%	45.3%	32.8%	21.2%	41.3%	31.0%	43.3
Health Care Worker	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**CDC**
Health Care Worker	37.0%	59.1%	46.9%	37.1%	59.5%	47.6%	78.6
Sex	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	
Male	19.5%	42.8%	30.3%	19.0%	38.2%	28.2%	
Female	23.3%	46.8%	34.5%	23.2%	44.2%	33.7%	
Pregnancy (among females)	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**CDC**
Pregnant	34.4%	50.0%	36.8%	34.3%	49.1%	38.4%	53.6
Pregnancy Since Aug 1 (among females not currently pregnant)	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	
Pregnant	33.3%	58.6%	42.5%	32.2%	62.2%	44.4%	
Age	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	
18-29	29.4%	33.6%	31.1%	27.6%	34.0%	30.5%	
30-39	22.1%	32.7%	26.5%	21.0%	31.9%	25.8%	
40-49	23.0%	36.0%	28.5%	22.1%	35.3%	28.0%	

50-64	17.8%	44.5%	30.8%	16.4%	44.6%	29.8%	
65+	20.6%	57.3%	40.1%	20.8%	58.2%	40.8%	
Race/Ethnicity	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	
Hispanic	35.1%	42.5%	38.0%	34.4%	39.5%	36.4%	
Non-Hispanic White	18.7%	48.8%	33.7%	18.7%	43.5%	31.2%	
Non-Hispanic Black	21.7%	39.5%	30.2%	22.1%	34.7%	28.6%	
Non-Hispanic Other or Multiple Races	18.6%	43.3%	32.4%	16.5%	40.5%	30.0%	
Cumulative by Month	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**BRFSS**
End of September	10.9%	9.9%	10.5%	11.0%	8.8%	10.4%	7.60%
End of October	15.7%	23.7%	19.3%	15.5%	21.0%	18.4%	22.90%
End of November	18.1%	31.5%	24.1%	17.8%	28.3%	23.0%	27.90%
End of December	19.2%	36.0%	26.8%	18.8%	32.4%	25.4%	32.80%
End of January	19.9%	40.0%	28.9%	19.5%	35.7%	27.3%	35.00%

* Estimates are for receipt of 1 or more doses of flu vaccine since August 1, 2016. Estimates are computed among all adults completing the adult section of Stage 2 of the interview and exclude those for which vaccination status is unknown due to item nonresponse.

† Weights are post-stratified to age category x sex x race/ethnicity category population counts from the 2015 ACS. The age categories were: 18-29, 30-39, 40-49, 50-64, and 65+; the race/ethnicity categories were Hispanic, non-Hispanic Black alone, and all other.

Table 6: Child Flu Vaccination Estimates*: 1 or More Doses

	Unweighted Estimates			Weighted Estimates†			External Estimates
Characteristic							
Overall	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**CDC§**
Overall	37.6%	43.9%	41.1%	34.5%	46.5%	41.7%	54.2±0.6
Sex	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**CDC §**
Male	32.4%	42.2%	37.8%	29.6%	44.4%	38.1%	n/a
Female	42.7%	45.9%	44.5%	38.3%	48.6%	44.9%	n/a
Age	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**CDC §**
0-4	39.4%	44.6%	41.9%	38.9%	42.2%	41.5%	63.7.0±1.1
5-12	37.3%	47.0%	42.4%	33.1%	48.8%	41.5%	55.2±0.9
13-17	35.8%	42.0%	39.6%	32.9%	49.3%	43.3%	45.1±1.3
Race/Ethnicity	**IVR**	**Live**	**Overall**	**IVR**	**Live**	**Overall**	**CDC §**
Hispanic	47.1%	47.3%	47.2%	46.2%	46.2%	47.4%	n/a
Non-Hispanic White	40.2%	38.2%	39.1%	40.5%	38.8%	39.5%	n/a
Non-Hispanic Black	28.4%	47.8%	39.9%	27.3%	47.8%	39.6%	n/a

Non-Hispanic Other or Multiple Races	20.0%	49.1%	40.2%	20.6%	50.9%	41.1%	n/a

Cumulative by Month	IVR	Live	Overall	IVR	Live	Overall	CDC ¶
End of September	21.9%	12.5%	17.0%	19.3%	13.4%	17.1%	16.7±0.5%
End of October	28.1%	19.2%	23.4%	24.5%	19.5%	23.0%	35.6±0.6%
End of November	32.4%	25.1%	28.5%	28.8%	24.5%	27.5%	46.7±0.6%
End of December	33.3%	27.6%	30.4%	29.7%	27.8%	29.7%	51.3±0.6%

* Estimates are for receipt of 1 or more doses of flu vaccine since August 1, 2016. MIDI estimates are computed among all children (including those less than 8 months of age) that have a complete child section of Stage 2 of the interview and exclude those for which the child's age or vaccination status is unknown due to item nonresponse.
† Weights are post-stratified to age category x sex x race/ethnicity category population counts from the 2015 ACS. The age categories were: 0-4, 5-12, and 13-17; the race/ethnicity categories were Hispanic, non-Hispanic Black alone, and all other.
§ NIS-Flu estimates are from the FluVaxView through January, 2017 (https://www.cdc.gov/flu/fluvaxview/reportshtml/reporti1617/reporti/index.html). Note that the NIS-Flu estimates here are for children age 8 months to 17 years, whereas the MIDI estimates are for children age 0-17 years.
§ ¶ Cumulative NIS-Flu estimates by month are from the Cumulative Estimates by Month Tables for January 2017 (https://www.cdc.gov/flu/fluvaxview/reportshtml/reporti1617/reporti/index.html. Note that the NIS-Flu estimates are for children age 6 months (as of Oct-Dec) to 17 years, whereas the MIDI estimates are for children age 0-17 years.

Table 7: Unweighted Distribution of Incoming Calls Across State

State FIPS	State		Based on Phone Number of All Incoming Calls (Including Respondents and Nonrespondents) Counts	Distribution (incl. missings)	Distribution (excl. missings)	ACS 2015
2	AK	Alaska	164	0.1%	0.1%	0.2%
1	AL	Alabama	2,605	1.8%	1.9%	1.5%
5	AR	Arkansas	1,065	0.8%	0.8%	0.9%
4	AZ	Arizona	2,432	1.7%	1.8%	2.1%
6	CA	California	15,791	11.2%	11.7%	12.2%
8	CO	Colorado	2,930	2.1%	2.2%	1.7%
9	CT	Connecticut	1,574	1.1%	1.2%	1.1%
11	DC	District of Columbia	614	0.4%	0.5%	0.2%
10	DE	Delaware	471	0.3%	0.3%	0.3%
12	FL	Florida	16,430	11.7%	12.2%	6.6%
13	GA	Georgia	4,257	3.0%	3.2%	3.1%
15	HI	Hawaii	239	0.2%	0.2%	0.5%
19	IA	Iowa	1,011	0.7%	0.8%	1.0%
16	ID	Idaho	414	0.3%	0.3%	0.5%
17	IL	Illinois	5,113	3.6%	3.8%	4.0%
18	IN	Indiana	2,394	1.7%	1.8%	2.0%
20	KS	Kansas	1,577	1.1%	1.2%	0.9%
21	KY	Kentucky	2,432	1.7%	1.8%	1.4%
22	LA	Louisiana	2,248	1.6%	1.7%	1.4%

25	MA	Massachusetts	2,489	1.8%	1.8%	2.2%
24	MD	Maryland	2,234	1.6%	1.7%	1.9%
23	ME	Maine	393	0.3%	0.3%	0.4%
26	MI	Michigan	3,602	2.6%	2.7%	3.1%
27	MN	Minnesota	1,588	1.1%	1.2%	1.7%
29	MO	Missouri	2,246	1.6%	1.7%	1.9%
28	MS	Mississippi	1,608	1.1%	1.2%	0.9%
30	MT	Montana	244	0.2%	0.2%	0.3%
37	NC	North Carolina	4,391	3.1%	3.3%	3.1%
38	ND	North Dakota	215	0.2%	0.2%	0.2%
31	NE	Nebraska	516	0.4%	0.4%	0.6%
33	NH	New Hampshire	369	0.3%	0.3%	0.4%
34	NJ	New Jersey	3,852	2.7%	2.9%	2.8%
35	NM	New Mexico	834	0.6%	0.6%	0.6%
32	NV	Nevada	1,064	0.8%	0.8%	0.9%
36	NY	New York	9,349	6.6%	6.9%	6.3%
39	OH	Ohio	4,405	3.1%	3.3%	3.6%
40	OK	Oklahoma	1,408	1.0%	1.0%	1.2%
41	OR	Oregon	1,116	0.8%	0.8%	1.3%
42	PA	Pennsylvania	4,446	3.2%	3.3%	4.0%
44	RI	Rhode Island	404	0.3%	0.3%	0.3%
45	SC	South Carolina	1,950	1.4%	1.4%	1.5%
46	SD	South Dakota	225	0.2%	0.2%	0.3%
47	TN	Tennessee	2,780	2.0%	2.1%	2.1%
48	TX	Texas	10,475	7.4%	7.8%	8.2%
49	UT	Utah	979	0.7%	0.7%	0.9%
51	VA	Virginia	2,547	1.8%	1.9%	2.6%
50	VT	Vermont	176	0.1%	0.1%	0.2%
53	WA	Washington	2,312	1.6%	1.7%	2.3%
55	WI	Wisconsin	2,113	1.5%	1.6%	1.8%
54	WV	West Virginia	461	0.3%	0.3%	0.6%
56	WY	Wyoming	125	0.1%	0.1%	0.2%
Missing/Invalid/Territory			6,170	4.4%		