FOR RELEASE OCTOBER 24, 2016

# Cellphone Activity Flags

## A trade-off between efficiency and coverage

BY *Kyley McGeeney* **AND** *Courtney Kennedy*

# About Pew Research Center

Pew Research Center is a nonpartisan fact tank that informs the public about the issues, attitudes and trends shaping America and the world. It does not take policy positions. The center conducts public opinion polling, demographic research, content analysis and other data-driven social science research. It studies U.S. politics and policy; journalism and media; internet, science and technology; religion and public life; Hispanic trends; global attitudes and trends; and U.S. social and demographic trends. All of the center's reports are available at www.pewresearch.org. Pew Research Center is a subsidiary of The Pew Charitable Trusts, its primary funder.

# Cellphone Activity Flags
*A trade-off between efficiency and coverage*

As telephone interviewing costs continue to rise and cellphones represent an increasing share of survey samples, survey researchers are exploring approaches to make these designs more cost-effective. A recent advancement in telephone random-digit-dial (RDD) studies[1] is the ability to identify nonworking cellphone numbers using what are known as "activity flags." Activity flags allow researchers to remove flagged nonworking cellphone numbers from the sample before they are dialed, thereby helping to contain rising interview costs.

One concern with using these flags, however, is that they might also erroneously flag eligible cellphone numbers for removal from the sample and in turn reduce the survey population coverage rate. This reduction in coverage can lead to bias in the resulting survey estimates.

Pew Research Center recently investigated the performance of these flags in the Center's 10,000 case, dual-frame RDD survey of U.S. adults conducted in early 2014. The primary finding is that while there was a measurable reduction in the population coverage rate, the effect on final weighted estimates was small.

## A large portion of cellphone RDD number are nonworking

Telephone survey researchers are very interested in using activity flags as a tool to combat rising data collection costs. One factor driving up the cost of cellphone interviewing is the substantial share of nonworking numbers in the sample dialed by interviewers. This results in wasted interviewer time. The incidence of nonworking numbers in the U.S. cellphone RDD frame was approximately 38% in 2015, based on results from Pew Research Center surveys.

Survey sample vendors began offering activity flags in 2012 as a way to increase interviewing efficiency. How the flag is created varies across vendors. Depending on the vendor, the flag reflects either past calling activity of a given number or a real-time test of the number's status as in service or not. Pew Research Center evaluated both types of activity flag by appending them to the cellphone sample drawn for a national RDD survey. Consistent with other research,[2] researchers

---

[1] When Pew Research Center undertakes these studies, they involve, in the case of cellphones, placing manually dialed calls to a random sample set of cellphone numbers to ensure compliance with applicable law.

[2] Dutwin, David and David Malarek. 2014. "The Use of Recent Activity Flags to Improve Cellular Telephone Efficiency." Survey Practice. 7 (1). http://www.surveypractice.org/index.php/SurveyPractice/article/view/311/html_34#r2

found that the flag relying on the real-time test performed better than the flag based on past calling activity. Analysis in the report is based solely on the real-time test flag.

## Activity flags lead to removal of nonworking cellphone numbers but also some working ones

Among the 124,811 cellphone numbers dialed in the 2014 Pew Research Center survey studied here, roughly six-in-ten (62%) were flagged as active, nearly one-third (32%) were flagged as inactive and the remainder (6%) had unknown status.

A comparison of the activity flags to the final dispositions assigned by the interviewers reveals that while the flag is accurate for the majority of cases, there are both false positives (numbers flagged as active that are actually nonworking or nonresidential) and false negatives (numbers flagged as inactive that are actually working and residential).

Among the flagged-inactive cases, 15 percent were found to be working and residential (false negatives). Among the flagged-active cases, 10 percent were found to be nonworking or nonresidential (false positives).

**One-third of cellphone numbers flagged as inactive**

*Among 124,811 dialed cellphone numbers ...*
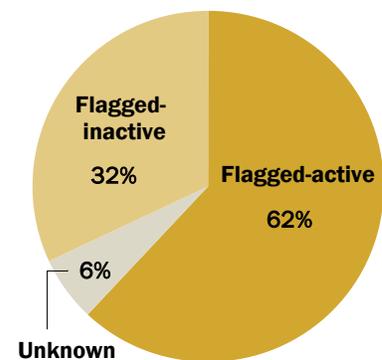


Flagged-inactive 32%

Flagged-active 62%

6%

Unknown

Source: Survey conducted Jan. 23-March 16, 2014.
"Cellphone Activity Flags"

**PEW RESEARCH CENTER**

## Activity flags have both false positives and false negatives

*% of cellphone numbers based on dialing*
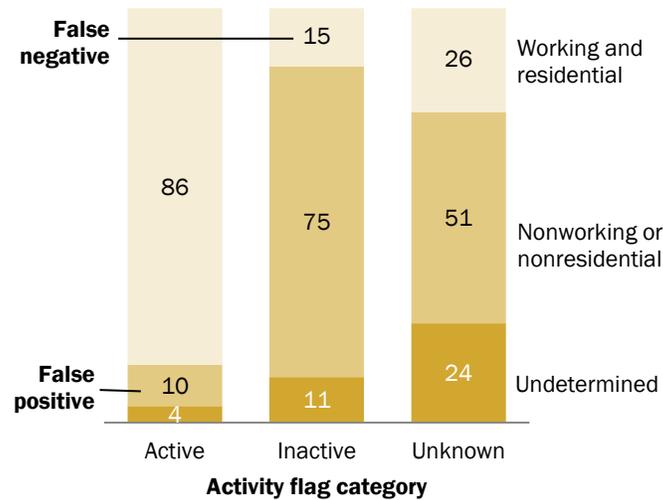


Note: Sample size=124,811 cellphone numbers dialed.
Source: Survey conducted Jan. 23-March 16, 2014.
"Cellphone Activity Flags"

**PEW RESEARCH CENTER**

# Coverage rate of dual-frame telephone sample designs reduced

False positives are basically harmless, aside from the loss in efficiency. False negatives, however, have the potential to reduce the survey coverage rate. This increases the risk of noncoverage error, which is when the people who have a chance to be sampled for a survey differ from those who have no chance of being selected, leading to bias in the survey's estimates.

Pew Research Center estimates that 8% of all working, residential cellphones in the U.S. are erroneously flagged inactive, which is several points higher than a previous study reported.[3] Taken together with the fact that 91% of U.S. adults own a cellphone, the result is a 7-percentage-point

---

[3] As in virtually all cellphone RDD samples, interviewers could not determine working and residential status for all dialed numbers. Some 7.5% of cellphone numbers dialed were assigned a final disposition of always busy, no answer or call blocking. The estimated impact on coverage is sensitive to the researcher's assumption of how many of the undetermined numbers are in fact working and residential. The estimated proportion of working and residential cellphone numbers that would be excluded using the activity is 7.8% under the AAPOR RR1 assumption that none of the undetermined numbers are working and residential, 8.2% under the RR3 assumption that some proportion of the numbers are working and residential, and 11.9% under the RR5 assumption that all of the numbers in the undetermined pool are working and residential. Prior work in this area indicates that the RR3 assumption is the most plausible.

reduction in sample coverage of U.S. adults. For a national cellphone sample, this reduces the estimated net coverage rate of U.S. adults from about 91% to 83%.[4]

These net rates are based on estimates from the Centers for Disease Control and Prevention's January to June 2015 National Health Interview Survey that 3% of adults have no telephone and another 6% have a landline but no cellphone. For national RDD surveys of adults, the inclusion of a landline sample substantially mitigates the reduction in coverage from excluding flagged-inactive numbers (from 97% coverage to 93%, rather than 83%).

### Excluding flagged-inactive cellphone numbers reduces sample coverage of U.S. adults by seven points

*% of U.S. adults with ...*

| **Only a landline** | | **A flagged-inactive cellphone** |
|---|---|---|

| 6 | A cellphone flagged-active or unknown 83 | 7 | 3 |

**No phone**
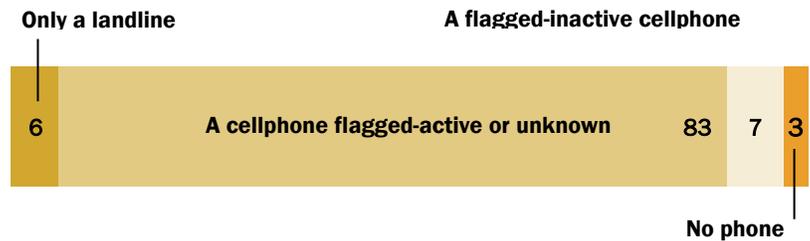
Note: 0.1% of adults in NHIS with landline but unknown wireless status included as landline only. Figures do not add to 100% because of rounding.
Source: Survey conducted Jan. 23-March 16, 2014, and CDC's National Health Interview Survey, January-June 2015.
"Cellphone Activity Flags"

**PEW RESEARCH CENTER**

It is possible to avoid any coverage reduction by using the activity flags to subsample flagged-inactive numbers rather than to wholly exclude them. In other words, despite their being flagged as inactive, the researchers would dial 50% of these numbers and then weight up their results to compensate for the lower sampling rate. Decreased precision and increased cost are two downsides of subsampling with these flags – at least in theory. The weighting adjustment to correct for the subsampling may increase the variance (reflected in an increased design effect[5]) and thus reduce the precision of survey estimates. Also, subsampling involves retaining some fraction of the flagged-inactive numbers in the sample, which reduces interviewer productivity relative to the exclusion approach. Considering that the activity flag currently costs 7 cents per number to append, the net cost savings may be marginal at best under the subsampling approach.

---

[4] The unrounded values are 90.6% and 83.2%, for a difference of 7.4%.
[5] The design effect is the ratio of the variance under the current sample design to the variance if a simple random sample had been used.

## Respondents with flagged-inactive numbers tended to be younger, less educated, more racially diverse

Flagged-inactive respondents are younger, less educated and lower income than flagged-active respondents. They are also more likely to be black non-Hispanic and less likely to be registered to vote or married than flagged-active respondents.

All of these demographic groups are considered hard to reach in surveys because they are less likely to respond. Typically, they are either not found in the resulting respondent sample at the same levels they are found in the population or researchers take great pains in terms of time or cost to have them adequately represented. Excluding or undersampling flagged-inactive cases from the sample means reducing the very respondents that researchers have the hardest time interviewing.

## Removing flagged-inactive numbers generally does not affect survey estimates

While excluding flagged-inactive cellphones reduces the survey coverage rate and disproportionately excludes hard-to-reach demographic groups, it does not appear to have a meaningful effect on bias, at least for public opinion surveys. Adults with cellphones erroneously flagged as inactive do differ systematically on several dimensions from those with flagged-active cellphones, as described above, but they constitute too small a fraction of the population to meaningfully move estimates. In the survey, 2% of all cellphone sample interviews and 1% of the combined (cellphone plus landline) sample interviews were with adults reached on numbers erroneously flagged as inactive using the activity flag.

### Respondents with flagged-inactive numbers more likely to be from hard-to-reach demographic groups

*Demographic profiles of flagged-active cellphone and flagged-inactive cellphone respondents*

|  | Flagged-active | Flagged-inactive |
|---|---|---|
|  | *Unweighted* | *Unweighted* |
|  | % | % |
| White, Non-Hispanic | **66** | 49 |
| Black, Non-Hispanic | **11** | 25 |
| Hispanic | 12 | 16 |
| Other, Non-Hispanic | 8 | 8 |
| 18-29 | **24** | 35 |
| 30-49 | 32 | 35 |
| 50-64 | 28 | 17 |
| 65+ | 16 | 13 |
| Registered to vote | **73** | 55 |
| High school grad or less | **30** | 50 |
| Some college/Associate degree | 29 | 33 |
| Bachelor's degree or more | **40** | 17 |
| *Family income* |  |  |
| Less than $30,000 | **31** | 51 |
| $30,000-$74,999 | 32 | 26 |
| $75,000 or more | **29** | 16 |
| Unweighted n | 4,881 | 101 |

.

Note: 21 cases flagged as "unknown" activity are not shown. Significant differences in **bold.**
Source: Survey conducted Jan. 23-March 16, 2014.
"Cellphone Activity Flags"

**PEW RESEARCH CENTER**

To assess whether excluding or subsampling flagged-inactive phones would have changed the study estimates, researchers created experimental weights simulating each scenario. The weight simulating exclusion precisely followed the survey weighting protocol, but discarded the 101 interviews with flagged-inactive cellphones. The weight simulating subsampling (at a rate of 50%) also followed the survey weighting protocol but discarded a random half of the 101 interviews with flagged-inactive cellphones and weighted up the cases in the non-discarded half by the inverse of the simulated subsampling rate (1/50%=2).

## Excluding or subsampling flagged-inactive cellphone respondents does not change weighted estimates

*Weighted estimates from full landline and cellphone sample, estimates from full sample with flagged-inactive cellphone respondents excluded and estimates from full sample with flagged-inactive cellphone respondents subsampled*

|  | Full sample | Flagged-inactive excluded | Flagged-inactive subsampled |
|---|---|---|---|
|  | *Weighted* | *Weighted* | *Weighted* |
|  | % | % | % |
| Internet user | 88.6 | 88.5 | 88.5 |
| Republican/Lean Republican | 39.0 | 39.1 | 39.2 |
| Conservative | 35.9 | 35.9 | 35.9 |
| Unaffiliated (religion) | 20.0 | 20.0 | 19.9 |
| Attend religious services weekly or more | 35.2 | 35.3 | 35.4 |
| Always vote | 49.1 | 49.5 | 49.4 |
| Very unfavorable view of Republican party | 23.9 | 24.0 | 24.0 |
| Very unfavorable view of Democratic party | 22.6 | 22.7 | 22.8 |
| Government should do more to help needy | 43.5 | 43.3 | 43.4 |
| Immigrants strengthen country | 56.7 | 56.7 | 56.6 |
| Follow government most of the time | 48.4 | 48.6 | 48.6 |
| Average (abs) difference from full sample |  | 0.1 | 0.1 |
| Unweighted n | 10,013 | 9,912 | 9,963 |

Note: Respondents who said they "Don't know" or refused to answer are not shown.
Source: Survey conducted Jan. 23-March 16, 2014; for adult population.
"Cellphone Activity Flags"

**PEW RESEARCH CENTER**

For the 11 common opinion survey questions analyzed, the weighted estimate simulating exclusion or subsampling of flagged-inactive cellphones were virtually indistinguishable from the final

weighted survey estimates. The average difference from the final survey estimate was near zero under both the exclusion and the subsampling simulations.

While full sample estimates were basically immune to use of the activity flags, it was not clear if this result would hold for subgroups – particularly subgroups that are more likely to have cellphones erroneously flagged as inactive. To test this, the Center repeated the analysis for blacks and adults ages 18-29. For estimates based on blacks, the average difference from the full sample figures ticked up slightly to 0.4 and 0.2 percentage points in the exclusion and subsampling simulations, respectively.

### Excluding or subsampling flagged-inactive cellphone respondents does not change weighted estimates for non-Hispanic blacks

*Among non-Hispanic blacks, weighted estimates from full landline and cellphone sample, estimates from full sample with flagged-inactive cellphone respondents excluded and estimates from full sample with flagged-inactive cellphone respondents subsampled*

| | Full sample | Flagged-inactive excluded | Flagged-inactive subsampled |
|---|---|---|---|
| | *Weighted* | *Weighted* | *Weighted* |
| | % | % | % |
| Internet user | 82.6 | 81.9 | 82.3 |
| Republican/Lean Republican | 9.8 | 9.4 | 9.9 |
| Conservative | 30.3 | 30.1 | 30.1 |
| Unaffiliated (religion) | 13.2 | 12.8 | 13.0 |
| Attend religious services weekly or more | 49.1 | 49.7 | 50.1 |
| Always vote | 53.5 | 54.0 | 53.6 |
| Very unfavorable view of Republican party | 37.7 | 38.2 | 37.9 |
| Very unfavorable view of Democratic party | 6.8 | 6.6 | 7.0 |
| Government should do more to help needy | 62.0 | 62.1 | 62.0 |
| Immigrants strengthen country | 56.0 | 56.1 | 56.2 |
| Follow government most of the time | 41.2 | 41.5 | 41.4 |
| Average (abs) difference from full sample | | 0.4 | 0.2 |
| Unweighted n | 998 | 973 | 986 |

Note: Respondents who said they "Don't know" or refused to answer are not shown.
Source: Survey conducted Jan. 23-March 16, 2014; for adult population.
"Cellphone Activity Flags"

**PEW RESEARCH CENTER**

For estimates based on adults ages 18 to 29, the average difference from the full sample figures ticked up slightly to 0.2 and 0.1 percentage points in the exclusion and subsampling simulations, respectively. The overall conclusions, however, did not change. For many public opinion surveys,

including most work conducted by Pew Research Center, differences of that magnitude (i.e., less than half a percentage point) are too small to trigger serious concerns about bias.

## Excluding or subsampling flagged-inactive cellphone respondents does not change weighted estimates for adults ages 18-29

*Among adults ages 18-29, weighted estimates from full landline and cellphone sample, estimates from full sample with flagged-inactive cellphone respondents excluded and estimates from full sample with flagged-inactive cellphone respondents subsampled*

|  | Full sample | Flagged-inactive excluded | Flagged-inactive subsampled |
|---|---|---|---|
|  | *Weighted* | *Weighted* | *Weighted* |
|  | % | % | % |
| Internet user | 99.0 | 98.9 | 99.0 |
| Republican/Lean Republican | 34.7 | 34.9 | 34.6 |
| Conservative | 25.8 | 25.5 | 25.5 |
| Unaffiliated (religion) | 30.0 | 29.9 | 29.8 |
| Attend religious services weekly or more | 27.5 | 27.4 | 27.6 |
| Always vote | 30.4 | 30.4 | 30.4 |
| Very unfavorable view of Republican party | 18.6 | 18.3 | 18.5 |
| Very unfavorable view of Democratic party | 13.8 | 13.9 | 13.8 |
| Government should do more to help needy | 46.6 | 46.1 | 46.4 |
| Immigrants strengthen country | 69.5 | 69.4 | 69.5 |
| Follow government most of the time | 27.4 | 27.5 | 27.2 |
| Average (abs) difference from full sample |  | 0.2 | 0.1 |
| Unweighted n | 1,405 | 1,370 | 1,389 |

Note: Respondents who said they "Don't know" or refused to answer are not shown.
Source: Survey conducted Jan. 23-March 16, 2014.

**PEW RESEARCH CENTER**

## Cost

Survey designers can exclude or subsample flagged-inactive cases from their samples and, in turn, save money by reducing the amount of time that interviewers spend manually dialing unproductive numbers. One team of researchers[6] estimated that excluding flagged-inactive numbers reduced the amount of labor hours needed for telephone interviewing by as much as 20 percent. Some survey research firms now routinely exclude flagged-inactive cellphone numbers from their samples in order to increase efficiency.

Based on this analysis, Pew Research Center decided to forego using activity flags as standard practice in its polling – not out of concern about bias (researchers found very little) but out of concern for the population coverage rate. Maintaining the highest possible coverage rate has always been a priority in the Center's work. For the time being, the potential efficiency gains do not justify the coverage rate reduction in our studies.

---

[6] Dutwin, David and David Malarek. 2014. "The Use of Recent Activity Flags to Improve Cellular Telephone Efficiency." Survey Practice. 7 (1). http://www.surveypractice.org/index.php/SurveyPractice/article/view/311/html_34#r2

# Acknowledgements

This report is a collaborative effort based on the input and analysis of the following individuals:

# Methodology

The analysis in this report is based on telephone interviews conducted Jan. 23-March 16, 2014, among a randomly selected national sample of 10,013 adults, ages 18 and older, living in all 50 U.S. states and the District of Columbia (5,010 respondents were interviewed on a landline, and 5,003 were interviewed on a cellphone, including 2,649 who had no landline telephone). The survey was conducted under the direction of Abt SRBI. A combination of landline and cellphone random-digit-dial samples were used; both samples were provided by Survey Sampling International. Interviews were conducted in English and Spanish. Respondents in the landline sample were selected by randomly asking for the youngest adult male or female who was at home at the time of the call. Interviews in the cell sample were conducted with the person who answered the phone, if that person was an adult 18 years of age or older. All cellphone numbers were manually dialed. For detailed information about our survey methodology, see http://www.pewresearch.org/methodology/u-s-survey-research/.

Data collection was divided equally into three phases (A, B, and C) with independent samples, non-overlapping interview dates and separate weighting. The questionnaire for each phase contained a core set of measures of political attitudes and values, political engagement and demographic characteristics, along with a set of unique questions about issues, lifestyle, media use and other topics covered in this series of reports. Additionally, most respondents to the survey were invited to join Pew Research Center's American Trends Panel.

The combined landline and cellphone sample is weighted using an iterative technique that matches gender, age, education, race, Hispanic origin and nativity and region to parameters from the 2012 Census Bureau's American Community Survey and population density to parameters from the 2010 U.S. Census. The sample also is weighted to match current patterns of telephone status (landline only, cellphone only, or both landline and cellphone), based on extrapolations from the 2013 National Health Interview Survey. The weighting procedure accounts for the fact that respondents with both landline and cellphones have a greater probability of being included in the combined sample and adjusts for household size among respondents with a landline phone. Three sets of weights were created for this analysis. Weights were calculated for the entire sample; the entire sample with 101 flagged-inactive cellphone cases excluded; and the entire sample with flagged-inactive cellphone cases randomly subsampled at a rate of 50%, where the remaining flagged-inactive cases where weighted up by the inverse of the simulated subsampling rate (1/50%=2). Sampling errors and statistical tests of significance take into account the effect of weighting.

The following table shows the unweighted sample sizes and the error attributable to sampling that would be expected at the 95% level of confidence for different groups in the survey:

| Group | Unweighted sample size | Plus or minus … |
|---|---|---|
| Total sample | 10,013 | 1.1 percentage points |
| Total sample with flagged-inactive subsampled | 9,963 | 1.1 percentage points |
| Total sample with flagged-inactive excluded | 9,912 | 1.1 percentage points |
| Flagged-active cellphone sample | 4,881 | 1.6 percentage points |
| Flagged-inactive cellphone sample | 101 | 11.2 percentage points |

Sample sizes and sampling errors for other subgroups are available upon request.

In addition to sampling error, one should bear in mind that question wording and practical difficulties in conducting surveys can introduce error or bias into the findings of opinion polls.