

Methodology of the ANES 2016 Non-Response Follow-Up Study

Matthew DeBell
Natalya Maisel

American National Election Studies
The University of Michigan and Stanford University
November 20, 2018

Suggested Citation

DeBell, Matthew, and Natalya Maisel. 2018. *Methodology of the ANES 2016 Non-Response Follow-Up Study*. Ann Arbor, MI and Palo Alto, CA: the University of Michigan and Stanford University.

Acknowledgments

This study was funded by the National Science Foundation under Grants SES-1444721 and SES-1444910. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the University of Michigan, Stanford University, or other individuals who worked on the study.

The ANES Principal Investigators for this Non-Response Follow-Up (NRFU) study were Shanto Iyengar at Stanford University and Vincent Hutchings and Ted Brader at the University of Michigan. Matthew DeBell was Co-Investigator at Stanford and study director for the NRFU. Staff contributing to the study were Natalya Maisel and Darrell Donakowski.

Data collection was performed by Westat, Inc. Michelle Amsbary, Brad Edwards, Vanessa Meldener, Roger Tourangeau, and others at Westat contributed to the study design and implemented the data collection.

See the documentation for the ANES 2016 Time Series Study for more information on the ANES 2016 Times Series Study.

Contact

The ANES website address is <http://www.electionstudies.org>

ANES sends occasional updates on Twitter @electionstudies

Any questions not answered on the ANES website or by this report may be directed to ANES staff by email at anes@electionstudies.org

Introduction and Purpose of the Study

A non-response follow-up (NRFU) study is designed to collect data about survey non-respondents so that a non-response error analysis can compare respondents to non-respondents and measure the extent of non-response error in a sample survey.

The present study collected data about sample members from the ANES 2016 Time Series study, including both respondents and non-respondents to that study. The ANES 2016 Time Series study used a dual-frame design, with independent samples selected for interviewing using face-to-face in-person interviews and using computerized questionnaires administered on the Internet. This NRFU study included sampled dwelling units from both modes. The study was conducted by mail with a one-page questionnaire and achieved a response rate of 45 percent.

ANES 2016 Non-Response Follow-Up Study at a Glance

Title:	ANES 2016 Non-Response Follow-Up Study
Purpose:	To measure non-response bias in both the face-to-face and Internet samples of the ANES 2016 Time Series Study.
Sample:	The study was done with a subset of ANES 2016 Time Series sample.
Data location:	The NRFU study data are part of the ANES 2016 Methodology Dataset, with additional documentation in that dataset's codebook and user's guide.
# of cases in this study:	4,725, of which 2,026 completed the NRFU questionnaire
Field dates:	March 10 to August 23, 2017
Mode of administration:	Mail
Incentives:	\$5 cash prepaid; non-respondents were offered \$20 postpaid incentive.
Questionnaire:	The questionnaire was a one-page paper questionnaire with 15 questions, one of which was a 4-item grid.
Response rate:	45 percent (AAPOR response rate 1)
Postage experiment:	The study incorporated an experiment testing the effect of mailing the invitation letters using First Class compared to Priority Mail.
Visible cash experiment:	The study incorporated an experiment testing the effect of making prepaid \$5 cash incentives visible through a window envelope, compared to a control group that used a standard enclosure.
Field firm:	Westat, Inc., performed data collection under contract, and in collaboration, with Stanford University.

Methodology

Sample: Address Selection

The sample for the NRFU was a subset of the sampled addresses for the ANES 2016 Time Series study. In the Time Series study, 10,680 addresses were selected, of which 7,800 were in the Internet sample and 2,880 were in the face-to-face sample. Due to budgetary limits it was not possible to include all 10,680 addresses in the NRFU.

There were 4,725 addresses selected for the study. The number of cases included in the NRFU sample by ANES Time Series mode and non-response status is shown in Table 1. The NRFU study was designed to include all addresses of face-to-face respondents, all addresses of face-to-face non-respondents that had not been determined to be ineligible for the Time Series study, 1,200 addresses from among the Internet respondents, and 1,300 addresses from among eligible or potentially eligible Internet non-respondents.

There were 2 face-to-face respondents not included in the NRFU sample due to a discrepancy in the operational definition of a completed case at the time the NRFU sample was drawn. There were 653 face-to-face non-respondents not included, of whom 652 were ineligible for the study (for reasons such as these sampled address not being an occupied residence or the household containing no US citizens) and 1 was an unusually firm refusal to the Time Series study who was excluded from all further contacts. There were 1,200 Internet respondents selected at random from among all 3,090 Internet respondents. There were 1,300 Internet non-respondents selected at random from among 3,641 eligible or potentially eligible Internet non-respondents.

Table 1. Number of cases included in NRFU sample, by ANES Time Series Status

Time Series status	NRFU sample status		Time Series total
	Included	Not included	
Respondents			
Face-to-face mode	1,179	2	1,181
Internet mode	1,200	1,890	3,090
Respondent subtotal	2,379	1,892	4,271
Non-respondents			
Face-to-face mode	1,046	653	1,699
Internet mode	1,300	3,410	4,710
Non-respondent subtotal	2,346	4,063	6,409
Total	4,725	5,955	10,680

Note: in this table, "non-respondents" include ineligible cases.

Sample: Person Selection

Data collection was matched to the ANES 2016 Time Series study at the level of the address, not the respondent. In the Time Series study, an individual respondent was randomly selected from among all

eligible adults at a sampled address. In the NRFU study, an adult at the sampled address was selected to complete and return the questionnaire, and no effort was made to contact the individuals who had been selected for ANES. The quasi-random person selection procedure used for the NRFU was to state in the invitation letter that we sought a response from the oldest or youngest male or female in the household. Letters were randomized between oldest and youngest and between male and female. The number of addresses to which invitations were addressed to each of these four types of possible respondents is shown in Table 2.

Table 2. Number of NRFU invitations by invitee

Invitee	Number
Youngest male	1,181
Oldest male	1,182
Youngest female	1,181
Oldest female	1,181
Total	4,725

In statistical expectation, the NRFU sampling procedure should result in the NRFU sample representing the Time Series sample, subject to random error, but in many households the NRFU respondent will not be the same person as the Time Series respondent. In households with two or more adult US citizens, the probability of the NRFU respondent and the Time Series respondent being the same is approximately $1/n$, where n is the number of adult US citizen residents. This assumes household composition did not change between the administration of the Time Series and NRFU studies, and that person selection was correctly implemented in both studies. The NRFU study did not identify changes in household composition between the Time Series and NRFU, but given the months elapsed between the studies it is inevitable that changes occurred in some households, such as residents at the time of the Time Series study moving away and new residents moving in.

Questionnaire

The NRFU used a one-page paper questionnaire in English with 15 numbered questions, of which one had four parts. Several questions collected data that was collected on the ANES Time Series survey, for comparison between the ANES respondents and the ANES non-respondents who completed the NRFU. Other questions were not previously asked of ANES respondents, and were designed for comparison between NRFU respondents who had completed the ANES and NRFU respondents who had not completed the ANES. All questions were considered possible correlates of non-response, and this was the primary basis for their inclusion in the study.

The questions were as follows:

1. When is the last time you talked to one of your neighbors?
2. Generally speaking, how often can you trust other people?
3. How much do you like or dislike talking about politics?
4. How much do you worry about your personal privacy?
5. Do you like or dislike each of the following?
 - College professors
 - News reporters
 - Surveys

National government

6. How much free time do you have?
7. How many children under 18 live with you?
8. Generally speaking, do you usually think of yourself as a Democrat, a Republican, an independent, or what?
9. Did you vote for president in November, and if so, who did you vote for?
10. Does anyone in this household connect to the Internet from home?
11. What is the highest level of school you have completed or the highest degree you have received?
12. Are you... [male, female?]
13. What year were you born?
14. Including you, how many adult U.S. citizens (age 18 or older) live in this household? If none, write 0.
15. Are you Spanish, Hispanic, or Latino?

Data Collection Dates and Mailings

The study was conducted entirely by mail. The study began on March 10, 2017 and ended August 23, 2017. Up to six mailings were sent, as follows.

1. On March 10, 2017, an advance postcard was sent, describing the study and announcing that, "In a few days, you'll receive a letter with the study and \$5 enclosed."
2. On March 21, an invitation letter was sent by First Class or Priority Mail (see Mailing Experiment & Visible Cash Experiment, below), with the questionnaire, pre-addressed postage-paid return envelope, and \$5 cash enclosed.
3. On March 28, a reminder/thank you postcard was mailed.
4. On April 10, a reminder/thank you letter was mailed by First Class mail with a second copy of the questionnaire and return envelope.
5. On June 21, a final letter was mailed by First Class mail, with a third copy of the questionnaire and return envelope, which stated that the study was "ending in a few days, on June 30." This letter offered a post-paid incentive of \$20 to complete the questionnaire.
6. On June 23, a final reminder postcard was mailed.

Respondents returned 2,026 questionnaires.¹ Each questionnaire mailed out to respondents was uniquely marked to identify the respondent and the mailing. There were 1,687 questionnaires returned from mailing #2, 247 returned from mailing #4, and 92 returned from mailing #5.

Branding

In order to differentiate the study from the ANES Time Series study, and thereby minimize the correlation between the response propensity to the ANES Time Series and the response propensity to

¹ After the study ended on August 23, 2017, and before the first draft of this report was completed on October 19, three completed surveys were received and four pieces of mail addressed to three addresses were returned by the Postal Service as undeliverable. These surveys received after the final end date of August 23 were not opened and were not added to the dataset because they were received after the formal closure of data collection. These respondents were not paid the \$20 incentive. The final letter to respondents stated that the study was ending on June 30. The letters from the three would-be respondents who were not paid were postmarked August 21, August 23, and September 6.

the NRFU, the NRFU was described to respondents as the Study of Attitudes and Lifestyles. In the first two mailings, its institutional sponsor was described as the Institute for Research in the Social Sciences (IRiSS) in Palo Alto, California. The IRiSS name and logo had not been previously used in any correspondence for the ANES Time Series study, which had emphasized its origins from the University of Michigan, Stanford University, the National Science Foundation, and Westat. In subsequent mailings, the Institute’s affiliation with Stanford University was included. All correspondence was signed by staff member Dr. Natalya Maisel, who had not signed any correspondence for the ANES Time Series study.

Mailing Experiment & Visible Cash Experiment

Addresses were randomly assigned to receive the invitation letter by First Class mail or Priority Mail. First class letters were sent in a #10 full-face window envelope and Priority Mail letters were sent in a 5 x 10 inch USPS Priority Mail Window Flat Rate Envelope (SKU EP 14-H) with a USPS tracking label. Both envelopes had full-face windows to reveal their contents from outside the envelope.

A second experiment (in a 2x2 experimental design with the postage mailing experiment) randomized the presentation of the prepaid \$5 cash incentive. This experiment varied whether the cash was or was not visible through a clear window in the envelope. Addresses were randomly assigned to a “visible cash” or “not visible cash” condition. In the “visible cash” condition, the \$5 bill was paper-clipped to the front of the invitation letter and placed inside the envelope so it was plainly visible from the front of the sealed envelope. In the “not visible cash” condition, the cash was paper-clipped behind the invitation letters so it could not be seen until the envelope was opened.

The number of cases per experiment group is shown in Table 3.

Table 3. Number of cases per NRFU experiment group

Window condition	Mailing condition		
	First Class	Priority	Total
Visible cash	1,181	1,181	2,362
Not visible cash	1,182	1,181	2,363
Total	2,363	2,362	4,725

Dispositions and Response Rates

There were 4,725 cases selected for the NRFU study and to which letters were mailed. Of these, at 2,440 addresses, we mailed all six mailings without receiving any response. At 2,026 addresses we received a completed survey returned by mail. At 41 addresses, a resident refused to participate: at 9 of these addresses the refusal was implicit, by returning a blank questionnaire (sometimes also returning the \$5 incentive), and at 32 of these addresses the refusal was a direct communication such as a phone call to Westat. At 217 addresses, mail was returned by the US Postal Service as undeliverable, for reasons such as the dwelling unit being vacant or the address not being valid. At one address we were told the address was not a residence. These dispositions, along with weighted numbers that adjust the NRFU sample to represent the full eligible Time Series sample, are shown in Table 4.

Table 4. NRFU case dispositions

Disposition	Unweighted number	Weighted number
Complete	2026	3831
Eligible or presumptively eligible nonresponse		
No response & no return	2440	4653
Refused		
Blank questionnaire returned	9	18
Other refusals (e.g., by phone)	32	54
Ineligible or presumptively ineligible		
Postal non-delivery	217	401
Non-residential address	1	1
Total	4725	8959

Note: weighted numbers do not sum to total due to rounding.

Response rate

A survey response rate is the proportion of eligible sample units completing the survey. The numerator of the response rate is the number of completed questionnaires, 2,026. The denominator is the total number of sampled cases minus the cases determined to be ineligible to participate. In this study the number of eligible cases is 4,725 minus the 217 cases where mail was returned minus the 1 case that was non-residential. The unweighted response rate is $2,026 / (4,725 - 217 - 1) = 45$ percent. In this calculation the cases where eligibility is unknown are assumed eligible, making this the minimum response rate, also known as AAPOR Response Rate 1.

The response rate noted above is unweighted. A weighted response rate would reflect the different selection probabilities of different subgroups of the sample. Based on the weighting factors given in the Weights section, below, the weighted response rate for the study is also 45 percent.

Response rates differed by Time Series study mode and Time Series response status. The weighted response rates were 53 percent for face-to-face respondents and 65 percent for Internet respondents (that is, for Time Series respondents selected to complete the Time Series study in those modes), and 62 percent for Time Series respondents overall. The weighted response rates were 33 percent for face-to-face non-respondents and 27 percent for Internet non-respondents, and 29 percent for Time Series non-respondents overall.

The response rates and numbers of responses by subgroup are shown in Table 5.

Table 5. Response numbers and rates by subgroup, weighted and unweighted

Subgroup	Unweighted			Weighted		
	Responses	Eligible	Rate (%)	Responses	Eligible	Rate (%)
Time Series respondents	1,369	2,304	59	2,565	4,138	62
Face-to-face mode	610	1,141	53	611	1,143	53
Internet mode	759	1,163	65	1,954	2,995	65
Time series non-respondents	657	2,203	30	1,266	4,419	29
Face-to-face mode	319	973	33	319	974	33
Internet mode	338	1,230	27	947	3,445	27
Entire NRFU sample	2,026	4,507	45	3,831	8,557	45

Note: weighted numbers may not sum to totals due to rounding.

Weights

As described in the Sample section, addresses were selected for inclusion in the NRFU study using unequal probabilities among sample groups of face-to-face and Internet respondents and non-respondents. To make the NRFU data representative of the full Time Series study sample, weighting factors must be applied to account for this sampling. These weighting factors are equal to the inverse of the selection probability for each case. Weighting factors can also account for differences in rates of non-response among these groups and for cases found after selection to be ineligible to respond. These weighting factors are shown in Table 6.

Table 6. Selection probability and weight factors for NRFU cases

Subgroup	Eligible to select	Selected	Selection factor	Eligible to respond	Response	Response factor	Weight factor
Time Series respondents							
Face-to-face mode	1,181	1,179	1.00170	1,141	610	1.87049	1.87366
Internet mode	3,090	1,200	2.57500	1,163	759	1.53228	3.94562
Time series non-respondents							
Face-to-face mode	1,047	1,046	1.00096	973	319	3.05016	3.05307
Internet mode	3,641	1,300	2.80077	1,230	338	3.63905	10.19215

In Table 6, the “Eligible to select” column shows the number of sample cases in each category: 1,181 people responded to the face-to-face survey and 3,090 responded to the Internet survey, and 1,047 addresses selected for a face-to-face interview had an eligible or potentially eligible respondent who did not respond, and 3,641 addresses selected for an Internet interview had an eligible or potentially eligible respondent who did not respond. The “Selected” column shows the number of these addresses selected for inclusion in the NRFU study. The design called for all eligible face-to-face cases to be included in the NRFU. Two respondents were not included because of a discrepancy between the operational definition of a completed case at the time of NRFU case selection and data release. One non-responding case was not included because this person contacted the Stanford University IRB to refuse to participate and the IRB directed us to make no further contact with this address.

The “Selection factor” column is the inverse of the selection probability, that is, $1/(\text{Selected}/\text{Eligible to select})$. This is the weighting factor for the calculation of the study’s weighted response rate, reflecting the sampling of cases for NRFU. (On the data file, this variable is `nrfu_selection_weight`.) The “Eligible to respond” column is the number of cases in the “Selected” column minus the number of cases that were determined during the NRFU study to be ineligible addresses, usually by returned mail indicating no one was living at the address. The “Response” column is the number of addresses that completed the NRFU questionnaire. The “Response factor” is the inverse of the probability of response, i.e. $1/(\text{Response}/\text{Eligible to respond})$. The weight factor is the product of the selection factor and response factor. When cases in the indicated subgroups are weighted by the weight factor, the NRFU sample represents the full ANES sample, adjusted for the cases found during NRFU to be ineligible. On the data file, this variable is `nrfu_weight`.

Variables, Data and Codebook

There are 29 variables from the NRFU study, as follows.

<i>Variable name</i>	<i>Description</i>
<code>version_nrfu</code>	Dataset version of ANES 2016 NRFU data file
<code>nrfu_flag</code>	Flag: included in NRFU study
<code>nrfu_group</code>	Experiment group, mail and cash
<code>nrfu_dispo</code>	Disposition
<code>nrfu_latepnd</code>	NRFU disposition 4 postal non-delivery returned July-August
<code>nrfu_selection</code>	Household adult invited to complete NRFU questionnaire
<code>nrfu_mailnum</code>	NRFU mailing returned to complete questionnaire
<code>nrfu_q1</code>	Q1. When is the last time you talked to one of your neighbors?
<code>nrfu_q2</code>	Q2. Generally speaking, how often can you trust other people?
<code>nrfu_q3</code>	Q3. How much do you like or dislike talking about politics?
<code>nrfu_q4</code>	Q4. How much do you worry about your personal privacy?
<code>nrfu_q5a</code>	Q5a. Do you like or dislike each of the following? College professors
<code>nrfu_q5b</code>	Q5b. Do you like or dislike each of the following? News reporters
<code>nrfu_q5c</code>	Q5c. Do you like or dislike each of the following? Surveys
<code>nrfu_q5d</code>	Q5d. Do you like or dislike each of the following? National government
<code>nrfu_q6</code>	Q6. How much free time do you have?
<code>nrfu_q7</code>	Q7. How many children under 18 live with you?
<code>nrfu_q8</code>	Q8. PID. Generally speaking, ... a Dem, a Repub, an ind., or what?
<code>nrfu_q9</code>	Q9. Did you vote for president in November, and if so, who did you vote for?
<code>nrfu_q10</code>	Q10. Does anyone in this household connect to the Internet from home?
<code>nrfu_q11</code>	Q11. EDU. Highest level of school or degree
<code>nrfu_q12</code>	Q12. SEX. Are you...
<code>nrfu_q13</code>	RESTRICTED. Q13. What year were you born?
<code>nrfu_q13x</code>	NRFU: Q13. What year were you born? - topcoded
<code>nrfu_q14</code>	Q14. Including you, how many adult U.S. citizens...live in this household?
<code>nrfu_q15</code>	Q15. Are you Spanish, Hispanic, or Latino?
<code>nrfu_mainstatus</code>	Time Series response status of NRFU cases
<code>nrfu_weight</code>	NRFU weight accounting for selection and nonresponse
<code>nrfu_selection_weight</code>	NRFU weight accounting for selection probability

The NRFU data are part of the ANES 2016 Methodology Dataset. The *User's Guide for the ANES 2016 Methodology Dataset* describes the data file, and the codebook information for the NRFU data is part of that study's codebook.