Pew Research Center's American Trends Panel
Wave 17
Methodology Report

Submitted to:

The Pew Research Center
1615 L Street, NW
Suite 700
Washington, DC 20036

Prepared by:
Abt SRBI
Marci Schalk
Nick Bertoni
Allison Ackermann
Raphael Nishimura
Dean Williams
Chintan Turakhia

275 Seventh Avenue, Suite 2700,
New York, New York 10001
(212) 779-7700
www.srbi.com

June 13, 2016

I. SUMMARY

The American Trends Panel (ATP) is a national, probability-based online panel of adults in the United States living in households. Adults who use the internet participate in the panel via self-administered Web surveys, and adults who do not use the internet participate via mail. The seventeenth wave of the panel survey was fielded for the Pew Research Center by Abt SRBI from May 10 to June 6, 2016. In total, 4,563 ATP members completed the survey, with 4,091 participating by Web and 472 participating by mail. The survey was administered in English and Spanish. Survey weights are provided to account for differential probabilities of selection into the panel as well as differential nonresponse to the panel recruitment survey, the panel invitation, and the panel survey itself (Wave 17). The margin of sampling error for full sample weighted estimates is ± 2.31 percentage points.

II. SAMPLE DESIGN

The target population for Wave 17 was non-institutionalized persons age 18 and over, living in the US, including Alaska and Hawaii. The sample consisted of 5,670 members of the ATP, which is a probability-based online panel of adults in the United States. The ATP originally consisted of 8,314 members, however, 371 members requested to be removed from the panel prior to the start of Wave 17. An additional 2,273 panelists were removed prior to the start of Wave 17 because they had not responded to any of the panel surveys since their recruitment or have since been unable to be reached by email, phone or mail. This number reflects 25 panelists that were inactivated prior to the beginning of Wave 17 due to non-response. The Wave 17 Survey featured a simultaneous mixed-mode design. Panelists who use the internet and provided an email address participated via self-administered Web survey, and adults who do not use the internet (or do but did not provide an email address) participated via a mail survey. Abt SRBI conducted the survey but is not reporting on results based on part of the sample.

III. QUESTIONNAIRE DEVELOPMENT AND TESTING

The questionnaire was developed by the Pew Research Center in consultation with Abt SRBI. In order to improve the quality of the data, the Web program was rigorously tested by the Abt SRBI project management team and PRC researchers. Test scenarios were created for all possible combinations of sample variables and question filters to ensure all skip logic was functioning correctly. Test data was then analyzed to determine that all logic was correct. After initial testing and corrections, the questionnaire was tested through our autopilot program which is a simulation of the data collected for the survey. The autopilot program tests the survey programming by initiating the Web interview and generating random responses as the questions appear. The Abt SRBI project management team checks the simulated answers against the questionnaire logic. This permitted Abt SRBI to track the response patterns and compare them to the questionnaire in order to identify programming errors. Pew Research Center has a copy of the final instruments in English and Spanish.

IV. DATA COLLECTION PROTOCOL FOR WAVE 17

Currently all ATP panelists have been recruited from two large (n=10,013 and n=6,004) national overlapping dual frame landline and cell phone random digit dial (RDD) surveys conducted for the Pew Research Center. At the end of each RDD survey, respondents were invited to join the panel. The first RDD survey was conducted from January 23rd to March 16th, 2014, and the second RDD survey was conducted from August 27th to October 4th, 2015, both in English and Spanish. Sample for the RDD surveys was obtained from SSI. Please refer to the Pew Research Center Political Typology/Polarization Survey Methodology Report and 2015 Pew Governance Survey Methodology Report for additional information on the sample design for the RDD surveys.

ATP panelists who reported using the internet and for whom we had an email address were invited to participate in Wave 17 via a self-administered Web survey. The data collection for the Web survey was conducted from May 16 to June 6, 2016. Advance postcards were mailed to all Web mode panelists with a known residential address. One hundred panelists were included in a soft launch of the Web survey which began with an initial email invitation sent on May 16, 2016. The Web panelists chosen for the soft launch were known responders to previous ATP surveys who had completed their surveys within two days of receiving their invitation. The remaining panelists assigned to the Web mode were included in the full launch and were sent an initial email invitation on May 17. All Web mode panelists who previously provided consent to receive text messages received a text message invitation in addition to their email invitation. Panelists who were known to have short codes blocked from previous survey wave SMS invitations were not included in this group. A total of 1,885 SMS invitations were sent. Up to

four reminder emails and SMS messages were sent to those who did not respond to the Web survey. Table S1 shows the field dates of mailings. The Web survey was closed June 6, 2016 at 9 a.m. Eastern.

Table S1. Mailing Dates for Wave 17 Panelists in Web Condition.

Mailing Dates for Wave 17 Panelists in Web Condition		
	Soft Launch	**Full Launch**
Advance Post Card	May 16, 2016	May 16, 2016
Initial email invitation	May 16, 2016	May 17, 2016
1st reminder email	May 19, 2016	May 20, 2016
2nd reminder email	May 25, 2016	May 25, 2016
3rd reminder email	May 31, 2016	May 31, 2016
Final reminder email	June 3, 2016	June 3, 2016

ATP Web panelists who completed their survey in Spanish were offered a $20 post-paid incentive for completing the Wave 17 Web survey. Panelists who were age 18-25, African American, with high school education or less, were not registered to vote or reported being Hispanic but taking the survey in English in the RDD recruitment survey were offered a $10 post-paid incentive for completing the Wave 17 Web survey. All other panelists who completed the Web survey were offered a $5 post-paid incentive. Web respondents could choose to receive the post-paid incentive in form of a check or a gift code to Amazon.com. The differential incentive amounts were designed to increase panel survey participation among groups that traditionally have low survey response propensities.

At the start of the Wave 17 Survey, 617 ATP members belonged to the non-Internet arm of the panel. The data collection for both English and Spanish language mail surveys was conducted May 10 to May 31, 2016 following the timeline shown in Table S2. The first packet of English language surveys was mailed using first class mail and Spanish language surveys were mailed using Priority Mail. ATP mail panelists were offered a cash pre-incentive that was included in their first survey packet. The same differential amounts that Web panelists received were also used for mail panelists. A reminder postcard was sent to all mail mode respondents. The second survey packet was mailed to 363 mail mode respondents who did not have a survey returned by May 20, 2016. The cutoff date to process returned mail surveys was May 31, 2016 which allowed a week for final data entry and quality checking.

Table S2. Mailing Dates for Wave 17 Panelists in Mail Mode.

Mailing	Date
First Packet	May 10, 2016
Reminder Post Card	May 13, 2016
Second Packet	May 20, 2016

V. WEIGHTING

Survey weights are needed to support reliable inference from the panel to the target population of US adults. The final survey dataset contains a full sample weight (WEIGHT_W17). The design of the full sample weight is described below.

The final full sample weight was computed in three main stages:

- Base weight adjusting for differential probabilities of selection
- Propensity adjustment for attrition

- Calibration to demographic distributions for the target population

Base Weight

A base weight is computed for all ATP members. The base weight adjusts for factors affecting the probability that the individual was selected for the panel. This probability comes from the survey in which the respondent was recruited. Currently, all ATP members were recruited through two probability-based, national overlapping dual-frame landline and cell phone RDD surveys: the Typology Survey, in 2014, and the Governance Survey, in 2015. The target population for the RDD surveys was identical to the target population for the ATP (adults living in households in the US). Both RDD surveys were administered in English and Spanish. In the landline sample of the RDD surveys, one adult was randomly selected from within the household. Interviewers asked to speak with either the youngest male or youngest female at home at the time of the call. In the cell sample of the RDD surveys, interviews were conducted with the person who answered the phone, provided they were age 18+ and spoke English or Spanish.

All respondents to the RDD surveys were invited to join the panel, except some individuals who do not use the internet, as this group was initially subsampled for the panel in the Typology Survey prior to all non-internet households being invited. The panel invitation featured a $10 post-paid incentive for agreeing to join and a fixed post-paid incentive for each panel survey completed. Hispanics/Latinos and adults age 18 to 25 recruited in the Typology Survey were offered $10 per panel survey. Hispanics/Latinos taking the survey in Spanish recruited in the Governance Survey were offered $20 per panel survey. Respondents who were Hispanics/Latinos taking the survey in English, African Americans, age 18 to 29, with high school education or less, or not registered to vote, recruited in the Governance Survey were offered $10 per panel survey. All other respondents in both surveys were offered $5 per survey. The differential incentives were designed to preemptively offset anticipated differential response rates across these groups.

For most ATP members, their base weight was computed using single frame estimation to adjust for the probability that the respondent's phone number was selected from the sampling frame, the overlap in the landline and cell phone frames, and the within household selection in the landline sample. For most panel members, the base weight can be expressed as:

$$BASEWT = \frac{1}{\left(\frac{S_{ll}}{U_{ll}} \times \frac{LL}{AD}\right) + \left(\frac{S_{cp}}{U_{cp}} \times CP\right) - \left(\frac{S_{ll}}{U_{ll}} \times \frac{LL}{AD} \times \frac{S_{cp}}{U_{cp}} \times CP\right)}$$

Where:

LL =1 if respondent has a landline phone
=0 if respondent has no landline phone

CP =1 if respondent has a cell phone
=0 if respondent has no cell phone

S_{ll}= number of cases released in the landline sample

S_{cp}=number of cases released in the cell phone sample

U_{ll}=size of the landline RDD frame

U_{cp}=size of the cell phone RDD frame

AD=number of adults in the household (1, 2, 3 or more)

For a subset of the ATP members, an additional adjustment is included in the base weight to account for the fact that they belong to a group that was subsampled for invitation to the panel. In the Typology Survey, non-internet users were subsampled at a rate of 25% from January 23, 2014, through February 5, 2014, but they were not subsampled (100% invited) from February 6, 2014 through the end of the field period. Internet users who agreed to join the panel in the Typology Survey but did not have an email address were taken at 100% from January 23, 2014, through February 5, 2014, but they were subsampled at a rate of 25% from February 6, 2014, through the end of the field period. The base weight of the affected cases was multiplied by the inverse of the subsampling rate (1 / .25 = 4).

Adjusting for Attrition

In total, 15,814 RDD survey respondents were invited to join the ATP and 8,314 accepted, yielding a panel acceptance rate of 52.6%. A majority of those who agreed to join the panel were still active at the start of Wave 17 (5,670/8,314=68.2%). To the extent that active panel members may be different from individuals who are not active (either because they declined to join or because they dropped out), there is a risk that estimates from the panel could be subject to nonresponse bias. A propensity score adjustment was computed to adjust for this attrition.

Most of the information available for individuals who either declined the panel invitation or have been dropped from the panel comes from the recruitment surveys. A logistic regression model was estimated in which being an active panel member was regressed on recruitment survey sampling frame, incentive amount ($10/20 vs $5 per survey), internet user, race, child in the household, age, education, religious service attendance, household income, frequency of voting, opinion of the Tea Party (agree vs disagree), political ideology, survey recruitment (Typology vs Governance), and statistically significant 2-way interactions ($p < .05$). The model was estimated using the respondents in the recruitment surveys who were invited to join the panel. Hispanic ethnicity was excluded from the model because it was collinear with the incentive variable. Number of adults in the household and incentive were not predictive and ultimately excluded from the model. The set of predictors considered for the model are variables that are routinely measured in surveys conducted for the Pew Research Center for the People & the Press. The significant predictors used in the final model are presented in Table S3.

The estimated propensities were used to divide cases into approximately equal size groups using the quintiles of the estimated propensity score. Quintiles have been found to be effective in capturing most of the variation. The propensity score adjustment was computed as the inverse of the active status rate in each quintile. This approach helps to protect against model misspecification, relative to using the inverse of the propensity score.

Table S3. Parameter Estimates from the Attrition Propensity Model^.

Parameter Estimates from the Attrition Propensity Model				
Variable (reference group)	**Estimate**	**s.e.**	**p-value**	
Intercept	-1.677	0.279	<.001	***
Frame (landline)	-0.215	0.125	0.085	*
Gender (male)	0.981	0.129	<.001	***
Internet User (non-user)	-0.997	0.130	<.001	***
Race (other race)			<.001	***
White	0.309	0.080	<.001	***

African American	0.259	0.094	0.006	**
Asian	-0.076	0.132	0.562	
Multi-racial	0.261	0.118	0.027	*
Tea Party (disagree)	0.499	0.130	<.001	***
Voting Frequency	0.171	0.019	<.001	***
Age	-0.008	0.002	<.001	***
Education	0.108	0.039	0.005	**
Religious Attendance	0.096	0.032	0.003	**
HH Income	-0.179	0.034	<.001	***
Ideology	0.129	0.062	0.039	*
Recruitment Survey (Typology)	0.723	0.141	<.001	***
Frame x Age	0.008	0.002	<.001	***
Internet User x HH Income	0.237	0.031	<.001	***
Gender x Tea Party	0.284	0.091	0.002	**
Gender x Age	-0.011	0.002	<.001	***
Gender x HH Income	-0.056	0.015	<.001	***
Tea Party x Ideology	-0.148	0.052	0.005	**
Education x HH Income	-0.009	0.004	0.046	*
Education x Ideology	0.037	0.009	<.001	***
Religious Attendance x Ideology	-0.043	0.010	<.001	***
Recruitment Survey x Age	-0.006	0.002	0.004	**
Recruitment Survey x Education	-0.053	0.019	0.007	**

*** $p<.001$, ** $p<.01$, * $p<.05$

^Variables are coded such that the model predicts active status in the panel. Positive coefficients are associated with a higher probability of being active. Negative coefficients are associated with lower probability of being active.

Calibration to Target Population Controls

In the final stage of weighting, the attrition-adjusted base weights for the panelists responding to a particular panel survey are calibrated to population benchmarks using raking, or iterative proportional fitting. This adjustment is designed to reduce the risk of nonresponse bias stemming from nonresponse at the various stages of the panel design. The raking dimensions and the source for the population parameter estimates are reported in Table S4. All raking targets are based on the non-institutionalized U.S. adult (age 18+) population. Moreover, the raking was conducted within each of the three forms of Wave 17.

Table S4. Raking Dimensions and Source for Population Parameter Estimates.

Raking Dimensions and Source for Population Parameter Estimates	
Raking Dimension^	**Source**
Gender(2) x Age(6)	2014 American Community Survey
Gender(2) x Education (3)	2014 American Community Survey
Age(3) x Education(3)	2014 American Community Survey
Race/Ethnicity(4)	2014 American Community Survey
Census Region(4)	2014 American Community Survey
Population Density(5)	2010 Decennial Census
Telephone Service(3)	July-December 2015 National Health Interview Survey, projected to 2016
Internet Usage(2)	2015 Pew Governance Study
Party Affiliation(3)	Average from the three most recent monthly surveys conducted for the Pew Research Center for the People

	& the Press
Volunteerism(2)	September 2013 Current Population Survey Volunteer Supplement
Daily Internet Usage(2)	July 2013 Current Population Survey Computer and Internet Supplement
^ The numbers of categories (prior to any collapsing from small cell size) are shown in parentheses.	

Most of the dimensions are commonly observed in weighting protocols for general population household surveys in the US. One exception is the raking for internet usage. This is included in the algorithm so that the panel survey estimates reflect the target population with respect to the proportion of people who use the internet and the proportion who do not. The large majority of ATP interviews are completed via self-administered Web survey. There is, therefore, a concern that internet users could be over-represented in the survey estimates if this dimension is not controlled for in the raking. Currently, the estimated population parameter for the percent of U.S. adults who use the internet is 89.3%, based on the 2015 Governance Survey conducted for the Pew Research Center. It would have been preferable to use a large, federal in-person survey (such as ACS or CPS) to obtain this parameter estimate, but unfortunately the federal government does not routinely measure internet access from any location.[1,2] Another two dimensions that are not typically used in weighting protocols for general population household surveys in the US are volunteering and daily Internet use. These are included in the calibration to adjust for some potential bias due to the over-representation of more politically and civically engaged adults of the panel identified in some recent analysis.

Trimming

The distribution of the raked weights was then evaluated and checked for extreme values. For Wave 17, the weights were trimmed at the 2nd and 98th percentiles.

VI. DESIGN EFFECT AND MARGIN OF ERROR

Weighting and survey design features that depart from simple random sampling tend to result in an increase in the variance of survey estimates. This increase, known as the design effect or *deff*, should be incorporated into the margin of error, standard errors, and tests of statistical significance. The overall design effect for a survey is commonly approximated as the 1 plus the squared coefficient of variation of the weights. For this survey, the margin of error (half-width of the 95% confidence interval) incorporating the design effect for full sample estimates at 50% is ± 2.31 percentage points. Estimates based on subgroups will have larger margins of error. It is important to remember that random sampling error is only one possible source of error in a survey estimate. Other sources, such as question wording and reporting inaccuracy, may contribute additional error. A summary of the weights and their associated design effect is reported in Table S5 below.

Table S5. Design Effect and Effective Sample Size.

[1] The July 2011 Current Population Survey estimated that 73% of US residents age 15 and older access the internet from some location. Given the increasing trends in internet access, particularly on mobile devices, this 2011 CPS estimate was deemed too out-of-date to be helpful in the ATP weighting.

[2] Starting in 2013 the American Community Survey is measuring internet access, but it only measures access inside the sample household. Members of the ATP are permitted to complete the surveys from any location. So the more relevant parameter for the ATP is the proportion of adults who can access the internet from any location, not just at home.

Weight Variable	Completed Interviews	Approximate Design Effect	Effective Sample Size	Margin of Error (95% confidence level)
WEIGHT_W17	4,563	2.54	1,798	± 2.31

VII. DISPOSITIONS

The final dispositions and AAPOR rates from the Web and mail components are reported in Tables S6a and S6b, respectively. The response rate to Wave 17 itself was 80.5%. Table S7 reports the cumulative response rate for Wave 17 when all of the stages of recruitment into the panel are taken into account.

Table S6. a. Final Dispositions from the Web Component of the Wave 17 Survey.

Final Dispositions from the Web Component of the Wave 17 Survey		
Final Disposition	**AAPOR Code[1]**	**Cases**
Completed interview	1.1	4,091
Logged onto survey; broke-off	2.12	41
Logged onto survey; did not complete any items	2.1121	24
Never logged on (implicit refusal)	2.11	897
Total Panelists in Web Sample of the Wave 17 survey		**5,053**
Completed interviews	I	4,091
Partial interviews	P	
Refusals	R	962
Non-contact	NC	
Other	O	
Unknown household	UH	
Unknown other	UO	
Not eligible	NE	
Total		**5,053**
AAPOR RR1 = I / (I+P+R+NC+O+UH+UO)		81.0%

Table S6. b. Final Dispositions from the Mail Component of the Wave 17 Study.

Final Dispositions from the Mail Component of the Wave 17 Study		
Final Disposition	**AAPOR Code[1]**	**Cases**
Completed mail survey during field period	1.10	472
Completed mail survey after field period	2.27	23
Refusal	2.11	0
Other: Known eligible but nothing ever returned	2.30	107
Other: Known eligible but undeliverable address	2.30	15
Total Panelists in Mail Sample of Wave 17 Study		**617**
Completed interviews	I	472
Partial interviews	P	
Refusals	R	107
Non-contact	NC	38
Other	O	
Unknown household	UH	
Unknown other	UO	
Not eligible	NE	
Total		**617**
AAPOR RR1 = I / (I+P+R+NC+O+UH+UO)		76.5%

[1] These codes are modified for reflect the fact that this survey was a panel survey, not an RDD survey. All sample members were eligible.

Table S7. Cumulative Response Rate.

Cumulative Response Rate	
Weighted Response Rate to Recruitment Surveys^	10.0%
Percent of Recruitment Survey Respondents Who Agreed to Join the ATP, Among Those Invited	52.6%
Percent of Those Agreeing to Join Who Were Active Panelists at Start of Wave 17	68.2%
Response Rate to ATP Wave 17 Survey	80.5%
Cumulative Response Rate for the Wave 17 Survey	**2.9%**
^ Weighted by the total phone numbers used in each survey	