

UnderStandingAmericaStudy

UAS 73: VOTING INTENTIONS OF SOCIAL CONTACTS



Survey author(s): Mirta Galesic and Wandí Bruine de Bruin

Fielded November 3, 2016 - November 7, 2016

Contents

1	Introduction	3
1.1	Topics	3
1.2	Experiments	3
1.3	Citation	3
2	Survey Response And Data	4
2.1	Sample selection and response rate	4
2.2	Timings	4
2.3	Sample & Weighting	5
3	Standard Variables	6
4	Background Demographics	9
5	Data conventions	13
6	Routing Syntax	14
7	Survey with Routing	15
	Election	15
	Closing	17

1 INTRODUCTION

This UAS panel survey, titled "UAS73: Voting intentions of social contacts" asks respondents questions about the voting intentions of the respondents' circle of friends. Previous administrations of these questions were asked in the 2016 election wave series file, asked in pollwaves 2, 6, and 11. A version was also asked again in the post-election poll UAS 71. This survey is no longer in the field. Respondents were paid \$2 to complete the survey. *For more information about all of the Understanding America Study's 2016 Election data files and documentation, visit the election data page at <https://uasdata.usc.edu/data/election-data>.*

1.1 Topics

This survey contains questions (among others) on the following topics: Politics, Social Networks. A complete survey topic categorization for the UAS can be found [here](#).

1.2 Experiments

This survey includes experiment(s) of the following type(s): Auxiliary Randomization. Please refer to explanatory comments in the Routing section for detailed information. A complete survey experiment categorization for the UAS can be found [here](#).

1.3 Citation

Each publication, press release or other document that cites results from this survey must include an acknowledgment of UAS as the data source and a disclaimer such as, 'The project described in this paper relies on data from survey(s) administered by the Understanding America Study, which is maintained by the Center for Economic and Social Research (CESR) at the University of Southern California. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of USC or UAS.' For any questions or more information about the UAS, contact Tania Gutsche, Project and Panel Manager, Center for Economic and Social Research, University of Southern California, at tgutsche@usc.edu.

2 SURVEY RESPONSE AND DATA

2.1 Sample selection and response rate

The sample selection for this survey was:

All active respondents to election series 2,6, or 11.

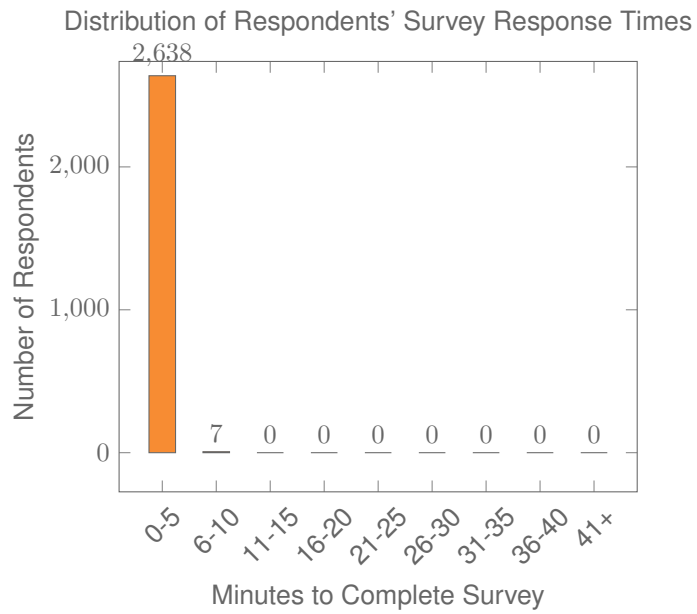
As such, this survey was made available to 3459 UAS participants. Of those 3459 participants, 2645 completed the survey and are counted as respondents. Of those who are not counted as respondents, 17 started the survey without completing and 797 did not start the survey. The overall response rate was 76.47%.

The detailed survey response rate is as follows:

UAS73 - Response Overview	
Size of selected sample	3459
Completed the survey	2645
Started but did not complete the survey	17
Did not start the survey	797
Response rate	76.47%

2.2 Timings

The survey took respondents an average of 1 minutes, and the full distribution of survey response times is available in the figure below. Times per question are available upon request.



2.3 Sample & Weighting

Weights are included in the data set for this survey. For details on the UAS weighting procedures please refer to the UAS Weighting Procedures V1. Please contact UAS staff with any questions.

3 STANDARD VARIABLES

Each Understanding America Study data contains a series of standard variables, consisting of individual, household and sample identifiers, language indicator, time stamps and a rating by the respondent of how much he or she liked the survey:

- **uasid**: the identifier of the respondent. This identifier is assigned to a respondent at recruitment and stays with the respondent throughout each and every survey he/she participates in. When analyzing data from multiple surveys, the 'uasid' can be used to merge data sets.
- **uashhid**: the household identifier of the respondent. Every member is assigned a household identifier, stored in the variable 'uashhid'. For the primary respondent this identifier equals his or her 'uasid'. All other eligible members of the primary respondent's household (everyone who is 18 or older in the household) who become UAS respondents receive the 'uasid' of the primary respondent as their household identifier. The identifier 'uashhid' remains constant over time for all respondents. Thus it is always possible to find the original UAS household of an UAS panel member (even after they, for example, have moved out to form another household).
- **survhhid**: uniquely identifies the household a UAS panel member belongs to in a given survey. For instance, if the primary respondent and his/her spouse are both UAS members at the time of a given survey, they both receive the same 'survhhid' identifier for that survey. If they subsequently split, they receive two different 'survhhid' in subsequent surveys. They, however, always share the same 'uashhid'. The identifier 'survhhid' is set to missing (.) if no other household members are UAS panel members at the time of the survey. Since individuals can answer the same survey at different points in time (which can be relatively far apart if the survey is kept in the field for a prolonged time), it may be possible that, within the same data set, household members have different 'survhhid' reflecting different household compositions at the time they answered the survey. For instance, suppose that the primary respondent and his/her spouse are both UAS members. If the primary respondent answers the survey when he/she is living with the spouse, but the spouse answers the survey when the couple has split, they receive different 'survhhid'. Hence, the variable 'survhhid' identifies household membership of UAS panel members, at the time the respondent answers the survey. Note: in the My Household survey 'survhhid' is set to unknown (.u) for respondents who last participated in the My Household survey prior to January 21, 2015.
- **uasmembers**: is the number of other household members who are also UAS panel members at the time of the survey. Since individuals can answer the same survey at different points in time (which can be relatively far apart if the survey is kept in the field for a prolonged time), it may be possible that, within the same data set, the primary respondent of a household has a value of '0', whereas the second UAS household respondent has a value of '1'. Therefore 'uasmembers' should be interpreted as the

number of household and UAS panel members at the time the respondent answers the survey. Note: in the My Household survey 'uasmembers' is set to unknown (.u) for respondents who last participated in the My Household survey prior to January 21, 2015.

- **sampletype**: indicates the sampling frame from which the household of the respondent was recruited. All UAS recruitment is done through address based sampling (ABS) in which samples are acquired based on postal records. Currently, the variable 'sampletype' takes on three values reflecting three distinct recruitment categories (in future data sets the number of categories may increase due to the incorporation of new recruitment categories):
 1. Nationally Representative Sample
 2. Native Americans: recruited through ABS, where the probability of drawing a zip-code is a function of the percentage of Native Americans in the zip-code. Primary respondents in these zip-codes who are not Native Americans are not invited to join the UAS.
 3. LA County: recruited through ABS drawing from zip-codes in Los Angeles County.
- **batch**: indicates the batch from which the respondent was recruited. There are currently the following values this variable takes (in future data sets the number of categories may increase due to the usage of new recruitment samples):
 1. ASDE 2014/01 Nat.Rep.
 2. ASDE 2014/01 Native Am.
 3. ASDE 2014/11 Native Am.
 4. LA County 2015/05 List Sample
 5. MSG 2015/07 Nat.Rep.
 6. MSG 2016/01 Nat.Rep. Batch 2
 7. MSG 2016/01 Nat.Rep. Batch 3
 8. MSG 2016/01 Nat.Rep. Batch 4
 9. MSG 2016/02 Nat.Rep. Batch 5
 10. MSG 2016/03 Nat.Rep. Batch 6
 11. MSG 2016/04 Nat.Rep. Batch 7
 12. MSG 2016/05 Nat.Rep. Batch 8
 13. MSG 2016/08 LA County Batch 2
 14. MSG 2017/03 LA County Batch 3
 15. MSG 2017/11 California Batch 1
 16. MSG 2018/02 California Batch 2
 17. MSG 2018/08 Nat.Rep. Batch 9

- 18. MSG 2019/04 LA County Batch 4
- 19. MSG 2019/05 LA County Batch 5
- 20. MSG 2019/11 Nat. Rep. Batch 10
- 21. MSG 2020/08 Nat. Rep. Batch 11
- 22. MSG 2020/10 Nat. Rep. Batch 12

- **primary_respondent**: indicates if the respondent was the first person within the household (i.e. to become a member or whether s/he was added as a subsequent member. A household in this regard is broadly defined as anyone living together with the primary respondent. That is, a household comprises individuals who live together, e.g. as part of a family relationship (like a spouse/child/parent) or in context of some other relationship (like a roommate or tenant).
- **hardware**: indicates whether the respondent ever received hardware or not. Note: this variable should not be used to determine whether a respondent received hardware at a given point in time and/or whether s/he used the hardware to participate in a survey. Rather, it indicates whether hardware was ever provided:
 - 1. None
 - 2. Tablet (includes Internet)
- **language**: the language in which the survey was conducted. This variable takes a value of 1 for English and a value of 2 for Spanish.
- **start_date (start_year, start_month, start_day, start_hour, start_min, start_sec)**: indicates the time at which the respondent started the survey.
- **end_date (end_year, end_month, end_day, end_hour, end_min, end_sec)**: indicates the time at which the respondent completed the survey.
- **cs_001**: indicates how interesting the respondent found the survey.

4 BACKGROUND DEMOGRAPHICS

Every UAS survey data set includes demographic variables, which provide background information about the respondent and his/her household. Demographic information such as age, ethnicity, education, marital status, work status, state of residence, family structure is elicited every quarter through the “My Household” survey. The demographic variables provided with each survey are taken from the most recent ‘MyHousehold’ survey answered by the respondent. If at the time of a survey, the information in “My Household” is more than three months old, a respondent is required to check and update his or her information before being able to take the survey.

The following variables are available in each survey data set:

- **gender**: the gender of the respondent.
- **dateofbirth_year**: the year of birth of the respondent.
- **age**: the age of the respondent at the start of the survey.
- **agerange**: if the respondent’s age cannot be calculate due to missing information, ‘agerange’ indicates the approximate age. Should a value for both the ‘age’ and ‘agerange’ be present, then ‘age’ takes precedence over ‘agerange’.
- **citizenus**: indicates whether the respondent is a U.S. citizen.
- **bornus**: indicates whether the respondent was born in the U.S.
- **stateborn**: indicates the state in which the respondent was born. This is set to missing (.) if the respondent was not born in the U.S.
- **countryborn**: indicates the country in which the respondent was born. This is set to missing (.) if the respondent was born in the U.S.
- **countryborn_other**: indicates the country of birth if that country is not on the drop down list of countries shown to the respondent’.
- **statereside**: the state in which the respondent is living.
- **immigration_status**: indicates whether the respondent is an immigrant. It takes one of the following values: 0 Non-immigrant, 1 First generation immigrant (immigrant who migrated to the U.S), 2 Second generation immigrant (U.S.-born children of at least one foreign-born parent), 3 Third generation immigrant (U.S.-born children of at least one U.S.-born parent, where at least one grandparent is foreign-born), or 4 Unknown immigrant status.
- **maritalstatus**: the marital status of the respondent.
- **livewithpartner**: indicates whether the respondent lives with a partner.

- **education**: the highest level of education attained by the respondent.
- **hisplativo**: indicates whether the respondent identifies him or herself as being Hispanic or Latino. This variable is asked separately from race.
- **hisplatinogroup**: indicates which Hispanic or Latino group a respondent identifies him or herself with. This is set to missing (.) if the respondent does not identify him or herself as being Hispanic or Latino.
- **white**: indicates whether the respondent identifies him or herself as white (Caucasian).
- **black**: indicates whether the respondent identifies him or herself as black (African-American).
- **nativeamer**: indicates whether the respondent identifies him or herself as Native American (American Indian or Alaska Native).
- **asian**: indicates whether the respondent identifies him or herself as Asian (Asian-American).
- **pacific**: indicates whether the respondent identifies him or herself as Native Hawaiian or Other Pacific Islander.
- **race**: indicates the race of the respondent as singular (e.g., '1 White' or '2 Black') or as mixed (in case the respondent identifies with two or more races). The value '6 Mixed' that the respondent answered 'Yes' to at least two of the single race categories. This variable is generated based on the values of the different race variables (white, black, nativeamer, asian, pacific). This composite measure is not conditional on hisplativo, so an individual may identify as Hispanic or Latino, and also as a member of one or more racial groups.
- **working**: indicates whether the respondent is working for pay.
- **sick_leave**: indicates whether the respondent is not working because sick or on leave.
- **unemp_layoff**: indicates whether the respondent is unemployed or on lay off.
- **unemp_look**: indicates whether the respondent is unemployed and looking for a job.
- **retired**: indicates whether the respondent is retired.
- **disabled**: indicates whether the respondent has a disability.
- **If_other**: specifies other labor force status.
- **laborstatus**: indicates the labor force status of the respondent as singular (e.g., '1 Working for pay' or '2 On sick or other leave') or as mixed (in case the respondent selects two or more labor statuses). The value '8 Mixed' indicates that the respondent answered 'Yes' to at least two of the single labor force status variables. This variable is generated based on the values of the different labor status variables (working, sick_leave, unempl_layoff, unempl_look, retired, disabled, If_other).

- **employmenttype**: indicates the employment type of the respondent (employed by the government, by a private company, a nonprofit organization, or self-employed). This is set to missing (.) if the respondent is not currently working or currently on sick or other leave.
- **workfullpart**: indicates whether the respondent works full or part-time. This is set to missing (.) if the respondent is not currently working or currently on sick or other leave.
- **hourswork**: indicates the number of hours the respondent works per week. This is set to missing (.) if the respondent is not currently working or currently on sick or other leave.
- **hhincome**: is the total combined income of all members of the respondent's household (living in their household) during the past 12 months.
- **anyhhmember**: indicates whether there were any members in the respondent's household at the time he/she answered the survey as reported by the respondent.
- **hhmembervnumber**: indicates the number of household members in the respondent's household at the time of the survey as reported by the respondent. It may be that 'anyhhmember' is 'Yes', but 'hhmembervnumber' is missing if the respondent did not provide the number of household members at the time of the survey.
- **hhmemberin_#**: indicates whether a household member is currently in the household as reported by the respondent. Household members are never removed from the stored household roster and their information is always included in survey data sets. The order of the roster is the same order in which household members were specified by the respondent in the 'MyHousehold' survey. The order is identified by the suffix _# (e.g., _1 indicates the first household member, _2 the second household member, etc.).

As an example, if the first household member is in the household at the time of the survey, 'hhmemberin_1' is set to '1 HH Member 1 is in the HH'; if he/she has moved out, 'hhmemberin_1' is set to '0 HH member 1 is no longer in the HH'. Since information of other household members (stored in the variables listed below) is always included in survey data sets, information about 'hhmemberin_1' is available whether this person is still in the household or has moved out.

- **hhmembergen_#**: indicates the gender of another household member as reported by the respondent.
- **hhmemberage_#**: indicates the age of another household member. The age is derived from the month and year of birth of the household member as reported by the respondent.
- **hhmemberrel_#**: indicates the relationship of the respondent to the other household member as reported by the respondent.

- **hhmemberuasid_#**: is the 'uasid' of the other household member if this person is also a UAS panel member. It is set to missing (.) if this person is not a UAS panel member at the time of the survey. Since this identifier is directly reported by the respondent (chosen from a preloaded list), it may differ from the actual (correct) 'uasid' of the UAS member it refers to because of reporting error. Also, this variable should not be used to identify UAS members in a given household at the time of the survey. This is because the variables 'hhmemberuasid_#' are taken from the most recent 'My Household' and changes in household composition involving UAS members may have occurred between the time of the respondent answered 'My Household' and the time the respondent answers the survey. To follow UAS members of a given household, it is advised to use the identifiers 'uashhid' and 'survhhid'.
- **lastmyhh_date**: the date on which the demographics variables were collected through the 'My Household' survey.

5 DATA CONVENTIONS

Data files provide so-called clean data, that is, answers given to questions that are not applicable anymore at survey completion (for example because a respondent went back in the survey and skipped over a previously answered question) are treated as if the questions were never asked. In the data files all questions that were asked, but not answered by the respondent are marked with (.e). All questions never seen by the respondent (or any dirty data) are marked with (.a). The latter may mean that a respondent did not view the question because s/he skipped over it; or alternatively that s/he never reached that question in the survey due to a survey break off.

If a respondent did not complete a survey, the variables representing survey end date and time are marked with (.c). Household member variables are marked with (.m) if the respondent has less household members (e.g. if the number of household members is 2, any variables for household member 3 and up are marked with (.m).

Formatting wise, in the STATA data sets all questions come with short descriptions (not available in the CSV files). 'Please select one' questions come with value labels for each answer option. In STATA these labels will include the labels 'Not asked' and 'Not answered' for (.a) and (.e), and will show in tabulations such as 'tab q1, missing'. For 'select all that apply' questions a binary variable is created for each answer option indicating whether the option was selected or not. A summary variable is also provided in the format '1-3-2' reflecting which options were selected and in which order. For example, if a question asked about favorite animals with options cat, dog, and horse, then if a respondent selected horse and then cat, the binary variables for horse and cat will be set to yes, while the overall variable would have a string value of '3-1'. If no answer was given, all binary variables and the summary variable will be marked with '.e'.

Questions that are asked multiple times are often implemented as so-called array questions. Supposing the name of such question was Q1 and it was asked in 6 different instances, your data set would contain the variables Q1_1_ to Q1_6_. To illustrate, if a survey asked the names of all children, then child_1_ would contain the name of the first child the respondent names and so on.

More information about the UAS data can be found in the UAS Data Guide available on the UAS Data Pages web site.

6 ROUTING SYNTAX

The survey with routing presented in the next section includes all of the questions that make up this survey, the question answers when choices were provided, and the question routing. The routing includes descriptions of when questions are grouped, conditional logic that determines when questions are presented to the respondent, randomization of questions and answers, and fills of answers from one question to another.

If you are unfamiliar with conditional logic statements, they are typically formatted so that **if** the respondent fulfills some condition (e.g. they have a cellphone or a checking account), **then** they are presented with some other question or the value of some variable is changed. If the respondent does not fulfill the condition (e.g. they are not a cellphone adopter or they do not have a checking account), something **else** happens such as skipping the next question or changing the variable to some other value. Some of the logic involved in the randomization of questions or answers being presented to the respondent is quite complex, and in these instances there is documentation to clarify the process being represented by the routing.

Because logic syntax standards vary, here is a brief introduction to our syntax standards. The syntax used in the conditional statements is as follows: '=' is equal to, '<' is less than, '>' is greater than, and '!= ' is used for does not equal. When a variable is set to some number N, the statement looks like 'variable := N'.

The formatting of the questions and routing are designed to make it easier to interpret what is occurring at any given point in the survey. Question ID is the bold text at the top of a question block, followed by the question text and the answer selections. When a question or variable has associated data, the name links to the appropriate data page, so you can easily get directly to the data. Text color is used to indicate the routing: **red** is conditional logic, **gold** is question grouping, **green** is looping, and **orange** is used to document randomization and other complex conditional logic processes. The routing is written for a computer to parse rather than a human to read, so when the routing diverges significantly from what is displayed to the respondent, a screenshot of what the respondent saw is included.

The name of the randomization variables are defined in proximity to where they are put into play, and like the question ID the names of the randomization variables can be used to link directly to the associated data page.

7 SURVEY WITH ROUTING

Start of section **Election**

WE001 (percentage of social contacts likely to vote or voted already in section Election)
Now we would like you to think of your friends, family, colleagues, and other acquaintances of 18 years of age or older that you have communicated with at least briefly within the last month, either face-to-face, or otherwise. We will call these people your social contacts.

What percentage of your social contacts are likely to vote (or have already voted) in the 2016 presidential election? For instance, 0% means that you think none of your social contacts will vote, and 100% means that all of your social contacts will vote (or have already voted). If you are not sure, just try to give your best guess.

Percentage of my social contacts who are likely to vote (or have already voted) is:
RANGE 0..100

/* In the below respondents are asked the percentage of their social contacts voted for Hillary Clinton or Donald Trump. The order of the two candidates is randomized per variable **WE002_order** with one of two values:

- o 1 Trump, then Clinton
- o 2 Clinton, then Trump

*/

IF WE002_order = EMPTY THEN

| WE002_order := mt.rand(1,2)

END OF IF

IF WE002_order = 1 THEN

GROUP OF QUESTIONS PRESENTED ON THE SAME SCREEN

WE002_intro (intro percentage likely to vote or voted in upcoming election in section Election)

For the next question, please consider only those of your social contacts who are likely to vote or have voted in the upcoming election for U.S. President. Of all of your social contacts who are likely to vote or have voted, what percentage do you think will vote/have voted for Hillary Clinton, Donald Trump, or someone else?

For instance, 0% would mean that you think no voters in your social circle will vote/have voted for that candidate, and 100% means that all voters in your social circle will vote/have voted for that candidate. Again, if you are not sure, just try to give your

best guess.

WE002a (Donald Trump in section Election)

Donald Trump
RANGE 0..100

WE002b (Hillary Clinton in section Election)

Hillary Clinton
RANGE 0..100

WE002c (Someone else in section Election)

Someone else
RANGE 0..100

END OF GROUP

IF WE002a + WE002b + WE002c != 100 THEN

WE002.check (check adds up to 100 in section Election)

Please go back and make sure the percentages add up to 100%

END OF IF

ELSE

GROUP OF QUESTIONS PRESENTED ON THE SAME SCREEN

WE002.intro (intro percentage likely to vote or voted in upcoming election in section Election)

For the next question, please consider only those of your social contacts who are likely to vote or have voted in the upcoming election for U.S. President. Of all of your social contacts who are likely to vote or have voted, what percentage do you think will vote/have voted for Hillary Clinton, Donald Trump, or someone else?

For instance, 0% would mean that you think no voters in your social circle will vote/have voted for that candidate, and 100% means that all voters in your social circle will vote/have voted for that candidate. Again, if you are not sure, just try to give your best guess.

WE002b (Hillary Clinton in section Election)

Hillary Clinton
RANGE 0..100

WE002a (Donald Trump in section Election)

Donald Trump
RANGE 0..100

WE002c (Someone else in section Election)

Someone else
RANGE 0..100

END OF GROUP

IF WE002a + WE002b + WE002c != 100 THEN

WE002_check (check adds up to 100 in section Election)

Please go back and make sure the percentages add up to 100%

END OF IF

END OF IF

End of section **Election**

Start of section **Closing**

CS_001 (HOW PLEASANT INTERVIEW in section Closing)

Could you tell us how interesting or uninteresting you found the questions in this interview?

1 Very interesting

2 Interesting

3 Neither interesting nor uninteresting

4 Uninteresting

5 Very uninteresting

CS_003 (comments in section Closing)

Do you have any other comments on the interview? Please type these in the box below.(If you have no comments, please click next to complete this survey.)

STRING

End of section **Closing**

/* Please note that although question CS_003 is listed in the routing, the answers are not included in the microdata in the event identifiable information is captured. Cleaned responses are available by request. */