**ONLINE APPENDICES FOR**

# Measuring Voter Registration and Turnout in Surveys:

# Do Official Government Records Yield More Accurate Assessments?

Matthew K. Berent, Stanford University

Jon A. Krosnick, Stanford University

Arthur Lupia, University of Michigan

June, 2015

**Online Appendix 1 - Government Records**

CALIFORNIA STATE VOTER FILE

The California state government file, obtained June 19, 2009, contains records for N=17,094,209 people.[1] The file contains each person's full name, current residential address, previous residential address, mailing address, date of birth, sex, telephone number, language, registration date, registration status, party registration, and turnout for the previous eight elections.

Each county maintains its own records and receives some information from the California Secretary of State about listings that need to be updated or changed. The California Secretary of State receives information from the California Department of Motor Vehicles (CDMV), the California Department of Public Health (CDPH), the California Department of Corrections (CDCR), and the United States Postal Service (USPS). Change of address information is received from the CDMV daily, and from the USPS monthly. Death records are obtained from the CDPH on an intermittent basis, and state felony conviction information is received from the CDCR monthly. Any information requiring a change in an person's registration listing is transmitted electronically to the appropriate county.

FLORIDA VOTER FILES

The Florida records are maintained separately for each of the 67 counties. Registration records and turnout histories are maintained in separate files. The combined government files,

[1] The California Secretary of State (2009) reported that 17,304,091 people were registered to vote in California for the 2008 general election.

obtained February 3, 2009, contain records for a total of N=12,570,868 people.[2] The registration records included listings for N=12,558,413 people, and the turnout records contained listings for an additional N=12,544 records for people who voted in the 2008 general election but were no longer registered at the time that the file was created. Each county's registration records contain a person's full name, residential and mailing addresses, date of birth, sex, race, date of registration, registration status, registered party affiliation, and telephone number. Each county's turnout history records included turnout for every election since at least 1998.[3]

Registration records are updated based on information from a variety of sources. New listings are added as the county receives registration requests from people. The counties also update information (e.g. new addresses or name changes) in existing registration listings when that information is provided by registered people. Each month, the state receives (from the Department of Health, clerks of the circuit court, United States Attorney, Department of Law Enforcement, Board of Executive Clemency, Department of Corrections, and Department of Highway Safety and Motor Vehicles) the names of people whose eligibility to vote has changed due to death, felony convictions, determination that a person lacks the mental capacity to vote, or relocation to another state. The state distributes this information to the county supervisors of elections, who update the listings in their records. County supervisors are also responsible for updating records based on change of address information supplied by the United States Postal Service.

---

[2] The Florida Division of Elections (2009) reported that 11,247,634 people were registered to vote in Florida as of October 6, 2008.

[3] Several counties included vote histories for elections prior to 1990, and one county included vote histories for an election apparently conducted in 1900.

## NEW YORK STATE VOTER FILE

The New York state government file, obtained June 11, 2009, contains records for N=12,718,771 people.[4] The records contain each person's full name, current residential address, mailing address, date of birth, sex, date of registration, method of registration, registration status, registered party affiliation, and turnout history since 2002.

New York counties maintain individual registration records and obtain information about records that need to be changed from the State Board of Elections. The State Board of Elections receives death notifications from the New York State Department of Health and the New York City Department of Health and Mental Health. The Board of Elections receives notifications of felony convictions and persons determined mentally incompetent from the New York State Office of Court Administration. The Board of Elections forwards this information to the counties, which have 25 days to update their registration records. The state also compares its statewide list of registered people to the National Change of Address registry managed by the United States Postal Service at least once a year and forwards information about address changes to the counties.

## NORTH CAROLINA VOTER FILES

The North Carolina government records, obtained on February 12, 2009, are maintained separately by each of its 100 counties. Each county maintains registration and turnout history information in separate records. The 100 counties' records contain listings for 6,230,749

[4] The New York State Board of Election (2009) reported that 12,031,312 people were registered to vote in New York as of November 1, 2008.

people.[5] The registration records included listings for N=6,223,974 people, and the turnout records contained listings for an additional N=6,776 records for people who voted in the 2008 general election but were no longer registered at the time that the file was created. Although the counties differ in the types of information maintained, all county records contain each person's full name, residential and mailing addresses, current age, sex, date of registration, registration status, registered party affiliation, and race. Each county's records contain turnout histories for elections during the previous 14 years.

County records are updated based on information from a variety of sources. New listings are added as the county receives registration requests from people. The counties also update information (e.g. new addresses or name changes) in existing registration listings when that information is provided by registered people. The state receives the names of deceased people who resided in the state from the Department of Health and Human Services once a month. The state distributes this information to the counties, which then remove the listings from their records. Each month, the state also receives the names of people who have committed felonies from the United States Attorney, and distributes that information to the counties so that registration listings may be updated. The counties also receive information monthly from the State Board of Elections regarding people who have been convicted of a felony within the state. Change of address information is acquired from the United States Postal Service and distributed to the counties every three months.

[5] The North Carolina State Board of Elections (2009) reported that 6,277,007 people were registered to vote in North Carolina as of November 29, 2008.

OHIO STATE VOTER FILE

The Ohio state government file, obtained April 25, 2009, contains records for N=8,246,881 people.[6] The records contain each person's full name, residential and mailing addresses, year of birth, registration date, turnout history for elections during the previous 7 years, and the partisan primary in which registered people turned out.

Ohio registration records are managed almost entirely by county boards. Each month, county boards obtain death information from the Board of Health. Counties also receive from the courts information about residents who have been convicted of felonies or deemed incompetent to vote. Counties update registration records and send them to the Ohio Secretary of State, which serves as a repository of those records. Every two years, the state checks information about registered people against the United States Postal Service's National Change of Address registry.

PENNSYLVANIA STATE VOTER FILE

The Pennsylvania state voter file, obtained June 1, 2010, contains records for N=9,444,317 people. Of these, N=9,246,579 were registered prior to the 2008 general election.[7] Government records are maintained separately for each of the 67 counties in that state. Each county maintains registration and turnout history information in separate records. Each county's

---

[6] The Ohio Secretary of State (2009) reported that 8,287,665 people were registered to vote for the 2008 general election.

[7] The Pennsylvania Department of State (2009) reported that 8,755,588 people were registered to vote in Pennsylvania as of November 4, 2008 (available at (http://www.dos.state.pa.us/portal/server.pt/community/voter_registration_statistics/12725).

registration records contain each person’s full name, residential and mailing addresses, date of birth, sex, date of registration, registration status, registered party affiliation, and telephone number.  Each county’s turnout history records include turnout and voting method for the last 40 elections in the county.

**Online Appendix 2 – Matching Methods**

To use a government record to evaluate a survey respondent's registration or turnout self-report, it is necessary to develop a method by which a respondent is matched to a record. Matching records from different sources presents challenges to researchers (McDonald and Levitt, 2008). Different challenges may be addressed through different criteria by which one record is considered "matched" to another. Strict criteria (e.g., coding as "matched" only those respondents for whom a record shares identical first and last names, address, and date of birth) will depress the number of successful "matches" relative to other methods – a single error in a single variable is sufficient to prohibit a match. Less strict criteria, such as one that requires only that a survey respondent's name be similar to a name in a government record, will generate more matches. However, less strict methods can match respondents to records that are not theirs. Therefore, *a priori*, it is not possible to specify which approach will produce the fewest errors. We chose to evaluate a range of matching procedures that differ in strictness to provide a broader view of the types of estimates that different TV procedures produce.

Matching was executed in two phases. First, we used Link Plus to identify potential matches in the government records for each survey respondent. Link Plus is a probabilistic record linkage software program developed at the CDC's Division of Cancer Prevention and Control in support of CDC's National Program of Cancer Registries (NPCR). Link Plus identified potential matches by linking records with the most similarities across names, address, and birth date (or year of birth or age). Similarities between last names were determined by shared sequences of characters (e.g., "Robertson" and "Roberson" share the character sequences "Rober" and "son", but not "t").[8] Shared character sequences were also used to identify similar

[8] Last name matching also included suffixes such as Sr., Jr., or III.

addresses. Name similarity was determined first by exact matches, then by matching formal names to common informal names, and last by the number of shared character sequences.[9] Birth date similarities were determined by how much of the month, day, and year information was the same. For states that provided only birth year or age, similarity was determined by the numerical proximity of the two numbers (e.g., 1975 is numerically more proximal to 1976 than is 1972). The potential matches for a survey respondent included listings with identical name, address, and birth date (or year or age) in both the ANES and government records, as well as those with an exact match on only one of those fields.[10]

---

[9] We did not observe any issues in the ability of Link Plus to parse and match address strings.

[10] Link Plus identifies which record from one data file has the highest probability of being the match for a record from a different data file. It does not identify all potential matches. However, Link Plus is limited to matching records from files with less than 5 million cases. We obtained more than 5 million registration and turnout history records from each state. To accommodate the Link Plus limitation, we divided government records from a state into smaller files with less than 5 million cases. The records making up the smaller files were determined by the first character of the last name associated with the record. For example, we divided the more than 17 million records from California into 5 smaller files. The first contained all records for a person whose last name started with A, B, C, or D. Records for people with a last names starting with one of the letter E through K were placed in the second file. The third file included last names starting with one of L through Q, the fourth R through S, and the fifth T through Z. Each California panel respondent was matched to one government record from each of the five California files. We used the match with the most similarities across first and last names,

The second phase of the matching process involved scoring potential matches based on increasingly strict sets of criteria. A potential match was scored as meeting the least strict criterion (which we call "LEAST") if the information provided by a survey respondent and the information in a government record listing had the same or similar first and last names and one of the following; (1) the same or similar residential address with the same or similar birth date information, or (2) different addresses with identical birth information.[11] Additional potential matches were scored as meeting the least strict criteria if they had identical last names, addresses, and dates of birth. Potential matches were scored as meeting the moderately strict criterion (which we call "MOD") if information provided by a survey respondent and a government record listing had the same or similar first and last names, had the same or similar addresses, and had identical birth date information. Matches scored as meeting the most strict

---

addresses, and birth dates in the analyses we report. We observed no case in which two or more government records matched a respondent equally well.

[11] Similar last names included minor differences that could easily be explained by a typographical error (e.g., Brown and Brow) and one name being a portion of a hyphenated name in the other (e.g. Brown and Brown-Smith). Similar first names included minor differences that could be explained by a typographical error (e.g., Christopher and Christophere) and one formal and a comparable informal name (e.g., Christopher and Chris). Similar addresses included minor differences that could be explained by a typographical error (e.g., 123 Elm St and 123 Elm Ct). Similar birth information included minor differences that could be explained by a typographical error (e.g., 01/01/1990 and 01/11/1990).

criterion (which we call “STRICT”) had identical first and last names, identical addresses, and identical birth date information.[12]

The three criteria produced strikingly different matching rates. Table OA2.1 compares match rates for LEAST, MOD, and STRICT with those of earlier ANES TV studies, and the match rate reported by Ansolabehere and Hersh (2008) in their 2006 Cooperative Congressional Election Study’s (CCES) TV effort.[13] Of our three methods, LEAST produced a match rate of 77.4%, and STRICT produced a match rate of 45.6%. LEAST matched a higher proportion of respondents to government records than did all of the other methods described in Table OA2.1.

Previous ANES TV projects attempted to locate records only for respondents who said they were registered. When looking only at respondents who said they were registered, LEAST produced a match rate (85.3%), which falls within the range of match rates found in previous ANES projects (low=77.6% in 1976 and high=90.8% in 1984). All ANES match rates were higher than that found in the CCES TV effort.

---

[12] All of the potential matches that met one criterion also met the criterion or criteria that were less strict. Respondents matched to a government record that did not meet the least strict criterion were categorized as not matched.

[13] The 2006 CCES was a survey of 30,000 respondents that was intended to permit projection to Congressional districts in the United States. CCES respondents were drawn from a non-probability sample of people who volunteered to participate in surveys in exchange for payment. Ansolabehere and Hersh limited matching to respondents living in one of the 26 states with the highest quality government records. The CCES validation attempt is described at http://web.mit.edu/polisci/portl/cces/commoncontent/Voter%20Validation%20Variables.pdf.

Finally, we tested whether a set of matching criteria generated turnout estimates that were closer to official turnout rates in some states than in other states. We tested this by first computing the difference between each respondent's turnout according to a matching method (where 1=turned out according to the method and 0=did not turn out according to that method) and the official turnout rate (see Table OA2.2 for the average difference scores). Then, we regressed difference scores on dummy codes for the six target states separately for each set of matching criteria. Omnibus tests of variation in the unstandardized coefficients indicated no significant differences among the states for the STRICT matching criteria [$F(5,798)=1.42$, ns.], the MOD criteria [$F(5,798)=.62$, ns.], or the LEAST criteria [$F(5,798)=.41$, ns.]. Thus, any set of criteria used to match respondents to government records yielded equally accurate turnout estimates in every target state.

**Online Appendix 3 - Sample Weighting**

We constructed two sets of weights for the survey respondents. Both sets used base weights to adjust for unequal probability of selection. One set of weights were created for all respondents who provided substantive answers to the registration and turnout questions to match the distributions of sex, age, race, education, income, and marital status among U.S. adult citizens residing in a U.S. household served by a landline. A second set of weights was created only for respondents who: (1) had a residential address in any of the six target states; (2) said they were registered in the state of their residential address or said they were not registered; and (3) provided a substantive answer to the turnout question. These respondents were weighted to reflect the distributions of sex, age, race, education, income, and marital status among adult citizens living in a household served by a landline in that state.[14] National and statewide estimates of these distributions were obtained using data from the 2008 American Community Survey (ACS), which is administered by the U.S. Census Bureau (2009).[15] The 2008 ACS collected data from 2,437,226 U.S. citizens 18 years or older, and estimates were computed with those data after they were weighted to reflect national distributions of residency location, marital status, race, ethnicity, sex, and age. The following sections describe the process of computing the weights.

---

[14] Respondents who reported registration in a state other than the state of their residential address are excluded from the analyses we report in the paper.

[15] The ACS is an ongoing survey conducted by the U.S. Census Bureau that collects data for community planning.

WEIGHTING RESPONDENTS TO REFLECT NATIONAL DISTRIBUTIONS

To compute weights for respondents who provided substantive answers to the registration and turnout questions, we followed the methodology recommended by DeBell and Krosnick (2010). The first step in this process involved acquiring the distributions of sex, age, race, ethnicity, education, income, and marital status in the U.S. adult population living in a household served by a landline in the 50 U.S. states. We obtained this information from the 2008 ACS and the 2008 National Health Interview Survey administered by the U.S. Center for Disease Control and Prevention (2009).[16]

The NHIS data indicate that about 83% of the U.S. population lives in households served by a landline. However, this percentage varies across demographic categories. For example, older people are more likely to live in a household served by a landline (65 years or older = 96.9%, 18-24 years = 69.2%). We used the NHIS data to estimate the percent of U.S. citizens living in the 50 states in a household served by a landline for as many of the demographic categories listed above that the data allowed.[17] We then used the NHIS landline rates to adjust the ACS distributions of the target demographics in each state.[18]

[16] The ACS does not weight respondents to reflect population distributions of education or income. We used the unweighted ACS distributions of these demographics as our estimates of the distributions of education and income in the U.S. adult population living in one of the 50 states.

[17] With the exception of income, NHIS demographic variables used the same categories as did the ACS. We estimated the percentages of people living in the 50 states in a household served by a landline for each of the ACS income categories by linearly interpolating using the NHIS data. For example, NHIS data indicated that every additional dollar of income between $17,500

According to DeBell and Krosnick (2010), a demographic variable should be used if the sample distribution deviates from the known population distribution by more than a small amount. They offered 5% and 2% as possible cut-offs for determining whether a difference between the sample and population distributions is small. None of the differences between the sample and population distributions for any of the demographic variables we examined were

---

and $42,500 ($17,500 and $42,500 are the midpoints of the two smallest NHIS income categories) corresponded to a .0002% percentage point increase in the probability of living in a household with a landline. To estimate the landline rate at an income of $32,500 (the midpoint of the ACS $25,000 to $39,999 income category), we multiplied .000002 by the difference between $32,500 and $17,500. We then added this product to the landline rate for people with incomes of $17,500 to estimate the landline rate at the midpoint of the $25,000 to $39,999 ACS income category. However, this method yielded a landline rate greater than 100% for the $175,000 or more ACS category. For this ACS category, we assigned the interpolated landline rate for the $85,000 to $174,999 ACS income category.

[18] For example, imagine the ACS data indicate that the distribution of sex in a population is 50% female and 50% male. Also imagine that the NHIS data indicate that 85.5% of adult females lived in a household served by a landline, whereas only 82.4% of adult males lived in such households. This means that of the 500,000 females in a population of 1 million adults, 427,300 live in a household served by a landline (500,000 × .8546). Of the 500.000 males in that same population, only 411,950 live in a household served by a landline (500,000 × .8239). This means that of the 839,250 adults living in a household served by a landline (427,300 females + 411,950 males), approximately 51% were female and 49% were male.

small by either of these standards (see Table OA2.1). Therefore, all of the demographic variables were used to compute weights for the total sample weighting procedure.

Another factor that determines the use of demographic variables in weighting is the number of people in each category in the sample and in the population. Samples or populations with a fewer than 5% of people in a demographic category can disproportionately influence weights (DeBell & Krosnick, 2010). For example, the weights assigned to cases in a low frequency demographic category may increase the deviations between weighted sample percentages and population percentages for other demographic categories. In the data we analyzed, four demographic categories contained sample percentages less than 5% (see Table OA2.1). The unweighted panel respondents in the 18-24 years old category made up only 3.3% of all respondents who provided substantive responses to the registration and turnout questions. Similarly, Hispanic respondents made up 4.7% of the respondents, only 2.9% of the respondents fell in the no high school diploma education category, and 1.2% of respondents reported a marital status of separated.[19] Effective weighting required modifying or eliminating demographic variables with low frequency categories for the weighting procedure.

We modified the categories for age, education, and marital status variables to eliminate problems that could be caused by these low frequencies. Specifically, we combined people ages 18-24 years old with people ages 25-34 years old to create a category of people 18-34 years old for the weighting procedure. For education, we combined no high school diploma and high school diploma categories to produce a high school diploma or less category. Finally, separated and divorced were combined into a single divorced or separated marital status category. This

[19] Separated people also made up only 2.3% of the population.

modification ensured that none of the age, education, and marital status demographic categories used to compute weights contained less than 5% of the respondents.

Addressing potential problems caused by the small number of Hispanics in the sample required a different solution. Given that the ethnicity demographic variable had only two categories (Hispanic and non-Hispanic), combining those categories would have eliminated all variability in ethnicity. To ensure that the weights assigned to the small number of Hispanics in the sample did not increase differences between the sample and population percentages for other demographic categories, ethnicity was not included in the weighting procedure.

The distributions of sex, age, race, education, income, and marital status in the total population of adult U.S. citizens living in a household in the 50 states that was served by a landline served as target values for our weighting procedure. We also included base weights provided by ANES (i.e., “wgtbase”) to account for unequal selection probabilities. The values were entered in the raking-ratio estimation routine of the R 2.13.1 statistical program (i.e., “anesrake”) to assign weights to respondents who provided substantive answers to the registration and turnout questions. This routine created weights so that the weighted respondents matched population percentages of sex, age, race, education, income, and marital status as closely as possible.

Distributions of weighted demographics were similar to the distributions of those variables among the adult population of U.S. citizens living in a U.S. household served by a landline (see Table OA2.1). The only demographic variable for which the percentage of weighted panel respondents differed from the population percentage by more than 5 percentage points was ethnicity. Due to the small percentage of unweighted panel respondents in the Hispanic ethnicity category (4.7%), ethnicity was excluded from the weighting procedure. As a result, the weighted percentage of panel respondents in the Hispanic category (6.4%) was 6.4 percentage points lower than the percentage of adult U.S. citizens living in a household served by a landline in one of the 50 states in that category (12.8%). Of the demographic variables that were included in the weighting procedure, only one weighted survey percentage deviated from the population percentage by more than one percentage point. This occurred in the case of race, for which the weighted sample percentage of White respondents (78.5%) was 1.3 percentage points higher than the target population percentage for that demographic category (77.2%). Thus, the weighted sample respondents closely resembled the population in terms of known demographics included in the weighting procedure.

## WEIGHTING RESPONDENTS TO REFLECT STATEWIDE DISTRIBUTIONS

We also used the raking-ratio estimation routine to weight respondents in each state to match the 2008 ACS's assessments of the states' adult population in terms of race, ethnicity, education, income, and marital status. We used the NHIS data to adjust the 2008 ACS's assessments of sex and age distributions in the state to determine the distributions of those demographics in the adult population living in a household served by a landline in the target states. With the exceptions of California and Florida, Hispanics made up less than 5% of respondents in the target states. Thus, we only included ethnicity in the weighting procedure for California and Florida. For other demographic variables, any category that included fewer than 5% of respondents in a state or fewer than 5% of a state's adult population residing in a household served by a landline were combined with other categories for the routine. The following summarizes how low frequency demographic categories were combined for the within-state weighting procedures:

- The 18-24 and 25-34 years age categories were combined into an 18-35 years age category for Florida, New York, North Carolina, and Pennsylvania.
- Black and other race categories were combined into a non-White category for North Carolina, Ohio, and Pennsylvania.
- The no high school diploma and high school diploma education categories were combined into a high school diploma or less category for California, Florida, New York, and North Carolina.
- The $85,000-$174,999 and $175,000 or more income categories were combined into an $85,000 or more income category for North Carolina and Ohio.
- Divorced" and separated marital status categories were combined into a divorced or separated category for California, Florida, New York, North Carolina, and Ohio.
- Divorced, separated, and widowed marital status categories were combined into a widowed, divorced, or separated category for Pennsylvania.

Distributions of demographic categories after weighting resembled population distributions of most demographic variables in most states (see Table OA2.2). The weighted percentages in demographic categories differed from the percentages in a state's adult population of U.S. citizens living in a household served by a landline by less than 5 percentage points for 90% of the demographic categories examined. Whereas 84% of the demographic categories yielded differences between weighted survey data and the state adult populations of U.S. citizens living in a household served by a landline exceeding 2.5 points, all of these involved demographic categories with small numbers of panel respondents in a state. The small numbers of panel respondents in a demographic category in a state precluded inclusion of the category in the weighting procedure. As a result, the weighted percentages of panel respondents in the affected categories sometimes deviated from the population percentages by more than 2.5 percentage points.

Of the demographic categories with sufficient numbers of panel respondents to be included in the weighting procedure for a state, only 2 differed by more than one percentage point after weighting. Both of these discrepancies occurred in Ohio. The weighted sample percentage of males among respondents living in Ohio in a household served by a landline (48.5%) was 1.3 percentage points higher than the percentage of males among the adult population of U.S citizens living in Ohio in a household served by a landline (47.2%). Also, the weighted sample percentage of whites among respondents living in Ohio in a household served by a landline (88.4%) was 2.5 percentage points higher than the percentage of whites among the adult population of U.S citizens living in Ohio in a household served by a landline (86.0%). Overall, weighted respondents resembled state populations along key demographic variables included in the weighting procedure, although this was somewhat less true for Ohio than the other target states.

**Online Appendix 4 - Estimating State-Level Population Sizes, Registration and Turnout Rates**

The ANES 2008-2009 Panel Study sample was drawn from the population of U.S. citizens, who were 18 years or older on Election Day 2008, and who lived in one of the 50 states in a household served by a landline. Government records do not provide information about this specific subpopulation's size, demographics, registration rate, or turnout rate. This appendix explains how we calculated those values.

ESTIMATING THE SIZE OF THE RELEVANT POPULATION IN EACH STATE

Our target populations are the numbers of U.S. citizens who were 18 years or older on Election Day 2008, and living in a U.S. household served by a landline in each of the six target states. We used data from multiple sources to estimate these target populations. We started with McDonald's (2009) estimate of the number of residents in a state who were 18 years or older on Election Day 2008, a number known as the state's Voting Age Population ($VAP_s$).

Given that people in prison were not part of the population from which the ANES sample was drawn, we subtracted the number of incarcerated persons ($INC_s$) from each *VAP* estimate, as measured by McDonald (2009).

We then accounted for the number of non-US citizens in each state, as measured by McDonald (2009). Thus, the formula ($VAP_s$ - $INC_s$) * (1-$\%NC_s$) estimated the number of non-incarcerated U.S. citizens living in each state.[20]

[20] This procedure assumes that the proportion of non-U.S. citizens living in a state is equivalent to the proportion of non-U.S. citizens in the state's incarcerated population. This assumption allowed us to avoid eliminating incarcerated non-citizens from the VAP twice (once with the

We then estimated the percentage of people living in households served by a landline in each state using data from the National Health Interview Survey (NHIS, Center for Disease Control and Prevention, 2009) and the American Community Survey (ACS, U.S. Census Bureau, 2009). The ACS data allowed estimation of the number of adults living in each target state by sex and age categories. The NHIS data indicated that the percentage of people living in a household served by a landline varied by sex and age. For example, according to NHIS data, 70.0% of 18 to 24 year old males lived in a landline household. This compares to 95.8% of males 65 year or older who lived in such households and 68.0% of 18-24 year old females who lived in such households. Hence, we used the NHIS estimates to adjust the ACS estimates of males and females in each age category so that the resulting estimates better represent the percentage of males and females in the different age categories living in landline households in each state.

For example, the ACS data estimated the number of males 18 to 25 years old living in California was 1,987,094. The NHIS data indicated that the percentage of 18 to 25 years old males living in landline households was 69.0%. Multiplying the ACS number by the NHIS percentage produced an estimate of 1,370,697 18-25 year old male living in Californians in landline households (1,987,094 × .6898 = 1,370,697). We computed this product for every sex × adult age demographic group in a state and summed those products to produce the total number of adults living in a household served by a landline for each state. Dividing this sum by the ACS

---

adjustment for incarcerated residents and then again with the non-citizen adjustment). We explored an alternative method that uses the numbers of incarcerated non-citizens in each target state published by the U.S. Department of Justice's Bureau of Justice Statistics (2014). This alternative produced a target population estimate across the six target states that differed from the estimate we report by less than .03%, and did not alter our conclusions.

estimate of a state's adult population produced our estimated proportion of adults living in a household served by a landline in that state ($e\%LL_s$). Our equation for computing $e\%LL_s$ was:

$$e\%LL_s = \sum_{g=1}^{12} N_g \times \%LL_g$$

where:

- $N_g$ = a state's population for one of the 12 sex × age demographic groups, and
- $\%LL_g$ = the percentage of the national population of the sex × age demographic group living in a household served by a landline.[21]

The state populations in the demographic groups ($N_g$) are shown in Table OA3.1, and the national numbers used to estimate each $\%LL_g$ are shown in Table OA3.2.

In conclusion, our estimate of a state's target population ($N_s$) is:

$$N_s = (VAP_s - INC_s) \times (1 - \%NC_s) \times e\%LL_s$$

Our estimated target population for each state ($N_s$) and the values used to estimate each estimate are shown in Table OA3.3.

## ESTIMATING THE NUMBERS OF PEOPLE WHO WERE REGISTERED IN THE STATE WITH WHICH THEY HAD A RESIDENTIAL ADDRESS

The numbers of people's registration records we obtained from each state were not equal to the numbers of people registered in those states according to the states' official published statistics. With the exception of Florida, we received fewer people's registration records from each state than the number of people registered according to published statistics. For example,

[21] The 12 age × sex combinations result from crossing sex (male or female) by six age ranges (18-24 years, 25-34 years, 35-44 years, 45-54 years, 55-64 years, and 65 years or more.

we received registration records for 17,094,209 people from California while California's published statistics claimed 17,304,091 people were registered to vote in that state.  Ohio showed the smallest discrepancy between the number of people's records we obtained and the number of people registered according to published state statistics (people's records=8,246,881; published number=8,287,665; difference=40,784) while Florida showed the largest discrepancy (people's records=12,570,869; published number=11,247,634; difference=1,323,235).  This means that estimates of the number of people with a residential address in a target state who were registered to vote in that same state would vary depending on the which source of registration numbers we used.

We used both the number of people's records obtained from a target state and each state's published registration numbers as starting points for estimating the number of people with a residential address in a target state who were registered to vote in that state.[22]  We then adjusted these numbers were adjusted to remove people who did not live in a household served by a landline and people who lived overseas. We now explain each adjustment.[23]

---

[22] The survey sample results we report include only respondents who had a residential address in the state in which they were registered, which required us to make a comparable adjustment to the state registration numbers.

[23] We did not attempt to account for duplicates in the state records.  Doing so would require estimates of the record duplication rates for the states.  The only estimates of state-level duplication rates we could locate were reported by Ansolabehere and Hersh (2010).  However, adjusting our estimates by their reported duplication rates yielded implausible numbers.  For example, Ansolabehere and Hersh (2010) reported a duplication rate in Florida records of 37%. Applying that rate to the people's records we received from Florida would have resulted in more

*Estimating the numbers of people registered to vote in demographic subgroups in the target states*. Some records in all target states included information that would allow us to determine the age of the person (e.g., date of birth), and every state other than Ohio provided the sex of at least some registrants. Some people's records had inaccurate ages or dates of birth (e.g., 1% of records from California listed implausible ages of 107 to 246 years old). Ages of younger than 18 years old or older than 106 years old on Election Day 2008 were treated as missing.

We assumed that the distributions of age and sex in people's records accurately represented those variables' distributions among all people in each state's records. State-level summaries of the age and sex distributions from records with those data and our estimated distributions for all state records are shown in Table OA3.4.

*Accounting for registrants in a household not served by a landline*. To estimate the proportion of people in government records without a landline, we used the NHIS to gauge the proportions of demographic subgroups living in households with landlines. Multiplying the estimated number of people's records for each demographic subgroup by the corresponding $\%LL_g$ yielded an estimate of the number of people in the demographic subgroup with a landline. Summing these products for all demographic subgroups in a state produced an estimate of the total number of people with a landline. Dividing this sum by the number of people's registration records in a state produced an estimate of the state-wide percentage of records that were for people with a landline ($e\%RRLL_s$).

---

people turning out in the 2008 election than there were people with registration records. We therefore did not make adjustments for duplicate records.

*Accounting for overseas eligible registrants*. Some people listed in the government records lived outside the United States. For example, military personnel stationed abroad may be registered to vote in a U.S. state. Because these people were not eligible to participate in the ANES survey, we used McDonald's (2009) estimates of the number of such persons in each state to adjust our registered voter population estimates downward. Specifically, we divided the number of overseas eligible residents of a state ($OE_s$) by $VEP_s$ to obtain the percentage of all people eligible to vote in a state who lived outside the U.S. ($\%OE_s$). We multiplied one minus this percentage by (1) the number of people's records, and (2) the registration numbers published by states in order to remove people living outside the United States.

*Accounting for registrants in a target state who lived in a different state.* The 2008-2009 ANES Panel Study respondents were asked not only whether they were registered to vote but in which state they were registered, and Knowledge Networks knew the residential address of each respondent. Therefore, we used these data to estimate the number of people with a registration record in each target state who lived in a different state (see Online Appendix 4 for the wordings of the registration address questions). The proportions of respondents reporting registration in a target state other than the state in which they resided were used to account for out-of-state registrants within each state.

The full equation for estimating the number of registered voters in each state who lived in that state with a landline ($R_s$) was:

$$R_s = (RR_s \times e\%RRLL_s) \times (1 - \%OE_s) \times (1 - e\%OSR_s)$$

where:

- $RR_s$ = either the number of people's registration records from a state or the number of people who the state said in published statistics were registered to vote in the 2008 general election in the state,
- $e\%RRLL_s$ = the estimated percentage of a people with a registration record living in a household served by a landline,
- $\%OE_s$ = a state's proportion of registrants living outside the U. S., and
- $e\%OSR_s$ =the estimated proportion of people with a registration record who resided in a different state.

The values used to compute $R_s$ are shown in Table OA3.5. The equation used for estimating each state's $e\%RRLL_s$ was:

$$e\%RRLL_s = \sum_{g=1}^{k} eRR_g \times \%LL_g$$

where:

- $eRR_g$ = the estimated number of a people registered in a state for one of $k$ demographic subgroups, and
- $\%LL_g$ = the national percentage of a demographic subgroup living in a household served by a landline.[24]

The values for $LL_g$ are shown in Table OA3.2, and $eRR_g$ numbers are shown in Table OA3.4.

ESTIMATING THE NUMBERS OF PEOPLE WHO TURNED OUT TO VOTE IN THE SIX TARGET STATES

[24] Given that Ohio's records did not include registrant's sex, $k$=6 age groups for that state. For all other states, $k=12$ age × sex combinations.

Each state reported the number of people who turned out to vote in the 2008 general election in that state ($NV_s$). These totals include people in our target population who turned out, but also people with a residential address outside the state, people living outside the United States, and voters who do not live in a household served by a landline.

We removed people who voted but did not have a landline using $e\%LL_s$, which estimates the percentage of a state's adult population that lived in a household served by a landline. We removed people who voted but lived outside the U.S. using the percentage of records for Overseas Eligible voters ($\%OE_s$), assuming this is also the ratio of the number of people living abroad who voted in a state to the total number of people who voted in that state.

The equation used to adjust $NV_s$ to estimate the number of people in our target population who turned out ($TO_s$) was:

$$TO_s = NV_s \times e\%LL_s \times (1 - \%OE_s)$$

where:

- $NV_s$ = the number of people who voted in a state according to statistics published by the state[25],
- $e\%LL_s$ = the estimated percentage of a state's adult population living in a household served by a landline.
- $OE_s$ = a state's overseas eligible population.

[25] We could have used the number of people who voted in the records we obtained from a state as an alternative to the number of people who voted according to statistics published by a state. However, turnout histories in some state's registration records are not updated for several months following an election (McDonald, 2007). This means that turnout histories in the registration records we obtained could underestimate the actual number of people who voted in the election.

- $VEP_s$ = a state's voting eligible population.

We report NVs and TOs in Table OA3.6.  The $e\%LL_s$ values are shown in Table OA3.1, and OEs and VEPs are shown in Table OA3.4.

**Online Appendix 5 - Commentary on Deadwood Adjustments**

Deadwood refers to registration records for people who no longer resided at the address on file for them, had died, had committed a felony, or who had become ineligible to vote for some other reason (McDonald, 2007). The registration records we obtained from the target states include such records, and deadwood influences the official registration statistics published by the states by inflating the number of people supposedly living in the state and eligible to vote. The amount of deadwood in registration records can vary substantially across states (Ansolabehere & Hersh, 2010), so accounting for deadwood requires different estimates for different states.

Ansolabehere and Hersh (2010) published state-specific deadwood rates, which could be used to adjust estimated registration statistics. However, two issues dissuaded us from doing so. First, we are unable to evaluate the accuracy of the published deadwood rates. Our inability stems from the fact that the commercial firm that generated the estimates for Ansolabehere and Hersh do not describe their methodology in sufficient detail to permit independent replication or review. Second, recent research reveals important problems with the accuracy of analogous data supplied by such firms (Pasek, 2012). For these reasons, we did not use the published deadwood rates in our article's state-level registration estimates. This appendix shows how using such data would have affected our estimates.

INFORMATION ABOUT THE ANSOLABEHERE AND HERSH DEADWOOD ESTIMATES

We asked the commercial firm that generated Ansolabehere and Hersh's (2010) deadwood estimates for information about how they were generated. Our communications with the firm revealed the type of data they used to generate their deadwood estimates, but in many cases left us unable to determine how they used the data. For example, in cases where multiple registration records exist for a person, the firm determined that one record was more accurate than another. However, the information the firm shared with us does not explain the criteria by which they reached such conclusions. What we describe below is a summary of what we learned from the commercial firm.

The commercial firm uses data from several sources to categorize individual registration records as "Likely Deadwood", "Probably Deadwood", "Possibly Deadwood", or "Not Deadwood". Sources of data included registration records, turnout histories, the Social Security Death Index file (SSDI), the National Change of Address file (NCOA), and information from another commercial vendor that maintained a consumer database.[26] They used this data to determine a record's deadwood status:

1. *Deceased Categorizations*. A person's age and the SSDI file record influenced whether an individual was categorized as deceased. The additional commercial vendor also flagged individual records as belonging to a deceased individual.[27] People listed in the SSDI file or who were flagged as deceased by the additional commercial vendor, and who were older than 65 years old were categorized as *likely deceased*. People listed in the SSDI file or flagged as deceased and younger than 65 years old were categorized as *possibly deceased*.

---

[26] The additional commercial vendor created records for consumers based on data provided by businesses. Those data come from product registrations and consumer account information.

[27] The additional commercial vendor flags people as deceased when the companies from whom they obtain information report that correspondence mailed to people are returned as undeliverable.

2. *Inactive Categorizations.* According to HAVA, registrants may be classified as inactive if they have not voted in two consecutive election cycles. Registration records from some states listed some people as inactive; other states only released information for active registrants. However, states differed in the strategies used to identify inactive registrants. Some states applied strict criteria, whereas others were more relaxed in identifying registrants as inactive. Records identified as inactive from states that followed strict procedures were categorized as *Strong Inactive*, whereas people from states with more relaxed procedures were categorized as *Weak Inactive*.
3. *The state has dropped a registrant's record*. The commercial vendor obtained government records from the states at different points in time. The vendor compared recently obtained records to those in their existing database to identify new records and records that were changed or dropped. The firm categorized as dropped people in their database who do appear in recently obtained records..
4. *The record is for an individual who has moved*. A registrant may have been categorized as having moved for two different reasons. First, a record existed for a registrant in one state that was "better" or more recent than the registrant's record in a different state, so the person was assumed to have moved out of the latter state. Second, the NCOA indicated a registrant had moved, and the registrant had a mailing address in a different county than the county listed in the vendor's database.
5. *The turnout history indicated that a registrant had not voted recently.* If records indicated that a registrant has not voted since 2000, the registrant was categorized as *not having voted recently*.

These five factors determined if a record was categorized as Likely, Probably, Possibly, or Not Deadwood. The combinations of factors that defined each deadwood level are as follows:

1. *Likely Deadwood*. A record was categorized as Likely Deadwood if: (1) the registrant was possibly or likely deceased, and had been dropped or classified as inactive by the state; or (2) the registrant had moved or had a "better" record in a different state, and had been dropped by a state or been categorized as Strong Inactive.
2. *Probably Deadwood*. A record was categorized as Probably Deadwood if: (1) the registrant had moved or had a "better" record in a different state and was categorized as Weak Inactive; (2) the registrant had not voted recently and had moved or had a "better" record in a different state; or (3) the registrant had not voted recently and had been dropped or classified as inactive by the state.
3. *Possibly Deadwood*. A record was categorized as Possibly Deadwood if: (1) the registrant was likely deceased; (2) the registrant had moved or had a "better" record in a different state; or (3) the registrant had not voted recently.
4. *Not Deadwood*. All records that were not categorized as Likely, Probably, or Possibly Deadwood were categorized as Not Deadwood.

Ansolabehere and Hersh (2010) used the number of records labeled "Likely Deadwood" or "Probably Deadwood" by the vendor to compute estimated deadwood rates for every state ($\%DW_s$).

## ESTIMATING THE NUMBER OF REGISTRATION RECORDS IN EACH TARGET STATE USING ANSOLABEHERE AND HERSH'S DEADWOOD RATES

Table OA3.5 in Online Appendix 3 reports estimated numbers of people registered based on the records obtained from the six target states and the statistics published by those states ($eRRLL_s$). Those numbers have been adjusted to account for people living in non-landline households, overseas, or in a different state. We multiplied each of those estimates by one minus Ansolabehere and Hersh (2010)'s deadwood estimates ($1\text{-}\%DW_s$). The full equation for estimating the number of non-deadwood registration records for people residing in a U.S. landline household in a target state ($eRRNDLL_s$) was:

$$eRRNDLL_s = eRRLL_s \times (1 - \%DW_s)$$

where:

- $eRRLL_s$ = the estimated total number of registration records for the target population of a state presented in Table OA3.5,
- $\%DW_s$ = Ansolabehere and Hersh's estimated deadwood rate for the state.

The state-level variables used to compute $eRRNDLL_s$ and *eRRNDLL* for the target states are in Table OA5.1. That table displays two $eRRNDLL_s$ estimates for each state: one based on the records we received from the states, and the other based on each state's published statistics.

## THE EFFECTS OF DEADWOOD ADJUSTMENTS ON REGISTRATION AND TURNOUT RATES

Deadwood adjustments reduce the estimated number of valid registration records and thereby reduce the percent of the nation or a state that was registered below the rates reported in Table 4 of our paper. As a result, the adjusted registration rate in the population of U.S citizens over 18 years of age who lived in a landline household in target states as of November, 2008 is lower than the estimated registration rate we report in that population (see Table OA5.2). Across the six target states, deadwood adjustments lowered the estimated registration rate by 3.29 percentage points (estimated registration rate=83.7% without deadwood adjustment; estimated rate=80.4% with deadwood adjustment).

Adjusting registration records in this way produced a slight shift in which matching method's estimated registration rates most closely resembled population rates. Without adjustment, the difference between the estimated population rate (83.7%) and the rate based on self-reports (87.4%; difference=3.7 percentage points) is smaller than the difference between the population rate and the rate based on LEAST (78.3%; difference=5.3 percentage points). With adjustment using Ansolabehere and Hersh's deadwood estimate, the population estimate (80.4%) is closer to the rate based on LEAST (difference=2.1 percentage points) than to the rate based on self-reports (difference=7.0 percentage points) Hence, the LEAST algorithm with deadwood adjusted registration numbers performed slightly better than any other method in estimating the population registration rate.

However, LEAST did not perform uniformly better across the six states in this domain. LEAST produced an estimated California registration rate (73.4%) that was more similar to the deadwood adjusted California population rate (74.4%) than any other method. LEAST and self-reports produced estimates that were similarly close to deadwood adjusted estimates for New York and North Carolina (New York: deadwood adjusted=82.7%, self-reports=82.7%, LEAST=83.3%; North Carolina: deadwood adjusted=87.9%, self-reports=95.7%, LEAST=81.1%). For Florida, Ohio, and Pennsylvania, self-reports produced turnout rates closest to deadwood adjusted rates (Florida: deadwood adjusted=76.3%, self-reports=81.0%;

Ohio: deadwood adjusted=90.6%, self-reports=94.0%; Pennsylvania: deadwood adjusted=82.2%, self-reports=86.7%). Thus, no matching algorithm consistently outperformed self-reports in resembling the population even after using Ansolabehere and Hersh's deadwood estimates.

These adjustments do not alter the conclusion that registered respondents turned out at a higher rate than did the population of registered people. These adjustments increased the population's estimated turnout rate across the six target states by about three percentage points (from 72.1% to 75.0%, see Table OA5.3). This adjusted turnout rate was still more than 13 percentage points below registration rates among respondents matched to a government record using LEAST, MOD, or STRICT. The adjusted turnout rate was also more than 19 percentage points lower than the self-reported turnout rate among respondents who said they were registered. Thus, the data continue to indicate that the survey respondents voted at a higher rate than the population from which they were drawn.

These adjustments also did not change the conclusion that no matching algorithm was uniformly superior to the others in estimating turnout in the state populations of people registered to vote. LEAST produced turnout rates most similar to the adjusted rates for California and Ohio (California: deadwood adjusted turnout rate among people registered=83.0%; rate among respondents matched to a government record using LEAST=91.8%; Ohio: deadwood adjusted=72.7%, LEAST=89.1%). MOD worked best for North Carolina and Pennsylvania (North Carolina: deadwood adjusted=74.5%, MOD=70.5%; Pennsylvania: deadwood adjusted=75.5%, MOD=97.3%), and STRICT outperformed the others in Florida and New York (Florida: deadwood adjusted=85.1%, STRICT=87.4%; New York: deadwood adjusted=70.7%, STRICT=69.5%).

In sum, compared to the rates we report in the paper, deadwood adjustment yielded (1) higher estimated population registration estimates and (2) lower estimated turnout rates among people registered to vote. However, the deadwood adjusted rates did not alter the following conclusions:

1. No matching algorithm was consistently most effective in producing registration rates across the six target states according to government records.
2. No matching algorithm was consistently effective in producing turnout rates among people registered to vote across the six target states according to the government records.
3. All methods of estimating turnout rates among registered survey respondents indicate that those respondents voted at a higher rate than did the populations of people registered to vote.

**References**

Ansolabehere, Stephen, and Eitan Hersh. 2008. Vote Validation in the 2006 CCES. Unpublished Manuscript. Available at http://web.mit.edu/polisci/portl/cces/material/website%20paper%20changes%2010_3_08.pdf

Ansolabehere, Stephen, and Eitan Hersh. 2010. The quality of voter registration records: A state-by-state analysis. Available at http://www.vote.caltech.edu/drupal/files/report/quality_of_voter_report_pdf_4c45d05624.pdf.

California Secretary of State. 2009. Voter participation statistics by county. Available at http://www.sos.ca.gov/elections/sov/2008_general/.

Center for Disease Control and Prevention. 2009. National Health Interview Survey [2008 Data Release]. Available at http://www.cdc.gov/nchs/.

DeBell, Matthew, and Krosnick, Jon A. 2010. Computing Weights for American National Election Study Survey Data. Available at http://www.electionstudies.org/resources/papers/nes012427.pdf

Florida Division of Elections. 2009. Voter Registration Statistics - By Election. Available at http://election.dos.state.fl.us/voter-registration/statistics/elections.shtml.

McDonald, Michael P. 2007. "The true electorate: A cross-validation of voter registration files and election survey demographics." *Public Opinion Quarterly* 71:588-602.

McDonald, Michael P. 2009. United States Elections Project [2008 general election turnout rates]. Available at http://elections.gmu.edu/.

McDonald, Michael P and Justin Levitt. 2008 . "Seeing Double Voting: An Extension of the Birthday Problem." *Election Law Journal* 7, 111-22.

New York State Board of Elections. 2009. 2008 election results. Available at http://www.elections.state.ny.us/.

North Carolina State Board of Elections. 2009. 2008 general election. Available at http://results.enr.clarityelections.com/.

Ohio Secretary of State. 2009. Voter turnout: November 4, 2008. Available at http://www.sos.state.oh.us/SOS/elections/electResultsMain/.

Pasek, Josh. November, 2012. Can microtargeting improve survey sampling? An assessment of accuracy and bias in consumer file marketing data. Paper presented at The Future of Survey Research: A Pair of Conferences at the National Science Foundation. Washington DC. Available at https://iriss.stanford.edu/content/future-survey-research-nsf.

Pennsylvania Department of State. 2009. Voter registration statistics: November 4, 2008. Available at http://www.dos.state.pa.us/portal/server.pt/community/voter_registration_statistics/12725.

United States Census Bureau. 2009. American community survey. Available at http://www.census.gov/acs/www/data_documentation/data_via_ftp/.

United States Department of Justice, Bureau of Justice Statistics. 2014. Reported number of non-U.S. citizens held in state or federal prisons, by gender, region, and jurisdiction, June 30, 2007-2008. Available at http://www.bjs.gov/.

**Table OA1.1**. Government Files Matching Variables by State

| State | Last Name | First Name | Address | Age | Birth Month | Birth Day | Birth Year |
|---|---|---|---|---|---|---|---|
| California | X | X | X | | X | X | X |
| Florida | X | X | X | | X | X | X |
| New York | X | X | X | | X | X | X |
| North Carolina | X | X | X | X | | | |
| Ohio | X | X | X | | | | X |
| Pennsylvania | X | X | X | | X | X | X |

**Table OA2.1** Match Rates of Different Criteria and Previous Vote Validation Studies

| | 2008 ANES Panel Study | | | ANES Time Series | | | | | | | 2006 CCES study |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample group | STRICT | MOD | LEAST | 1976[a] | 1978[b] | 1980[a] | 1984[b] | 1986[b] | 1988[b] | 1990[b] | |
| Respondents living in a location for which government records were available | 45.6% | 65.1% | 77.4% | 65.2% | 63.6% | 64.3% | 72.7% | 63.4% | 68.5% | 60.8% | 62.5% |
| *N* | *972* | *972* | *972* | *2,908* | *2,304* | *1,522* | *2,257* | *2,140* | *2,005* | *1,980* | *22,930* |
| Respondents who said (or implied) they were registered in their state of residence or an area for which government records were available | 50.4% | 72.3% | 85.3% | 77.6% | 83.6% | 81.0% | 90.8% | 86.3% | 88.5% | 87.1% | 64.9% |
| *N* | *815* | *815* | *815* | *2,256* | *1,650* | *1,095* | *1,622* | *1,572* | *1,383* | *1,383* | *21,806* |

Notes:

[a]Self-reported registration is based on responses to a single registration status question asked during pre-election interviews (V763032 in 1976, and V800134 in 1980).

[b]Self-reported registration is based on responses to turnout and registration questions (the latter of which was only asked to respondents who reported not turning out in a general election) asked during post-election interviews. Respondents were coded as reporting registration if they reported turning out in the general election or being registered (V780470 and V780484 in 1978, V840783 and V840802 in 1984, V860261 and V860276 in 1986, V880756 and V880757 in 1988, and V900279 and V900280 in 1990).

**Table OA2.2** Unstandardized Regression Coefficients for State Dummy Codes Predicting Difference Scores between Official State Turnout Rate and Dummy Coded Turnout Based on for Different Matching Criteria.

| | Unstandardized regression coefficients for different matching criteria | | | | | |
|---|---|---|---|---|---|---|
| | STRICT (n=804) | | MOD (n=804) | | LEAST (n=804) | |
| Target state | Unstandardized coefficient | *SE* | Unstandardized coefficient | *SE* | Unstandardized coefficient | *SE* |
| California | -.19** | .05 | -.04 | .05 | .06 | .05 |
| Florida | -.24** | .07 | -.06 | .08 | .08 | .08 |
| New York | -.31** | .07 | -.04 | .08 | .10 | .08 |
| North Carolina | -.28** | .11 | -.15 | .11 | -.05 | .10 |
| Ohio | -.15* | .06 | -.01 | .06 | .09# | .05 |
| Pennsylvania | -.07 | .07 | .05 | .07 | .09 | .07 |
| $R^2$ | .16*** | | .02 | | .03 | |

Notes: # p<.10; * p<.05; ** p<.01 two-tailed.

**Table OA3.1**. Distributions of Demographic Variables among Weighted Panel Respondents and Among the Population of Adult U.S. Citizens Living In a Household Served By a Landline in One of the 50 States

| Demographic | Category | All ANES Panel Study respondents with registration and turnout self-reports | The landline population of U.S. adult citizens living in one of the 50 states | Difference |
|---|---|---|---|---|
| Sex | Male | 47.8% | 47.8% | .0% |
| | Female | 52.2% | 52.2% | .0% |
| Age | 18-34 | 24.8% | 24.8% | .0% |
| | 35-44 | 18.8% | 18.8% | .0% |
| | 45-54 | 20.6% | 20.6% | .0% |
| | 55-64 | 17.6% | 17.6% | .0% |
| | 65+ | 18.2% | 18.2% | .0% |
| Ethnicity | *Non-Hispanic* | 93.6% | 87.2% | 6.4% |
| | *Hispanic* | 6.4% | 12.8% | -6.4% |
| Race | White | 78.5% | 77.2% | 1.3% |
| | Black | 10.7% | 11.3% | -.6% |
| | Other | 10.8% | 11.5% | -.7% |
| Education | High school diploma or less | 44.5% | 44.5% | .0% |
| | Some college, no bachelor's degree | 29.7% | 29.7% | .0% |
| | Bachelor's degree | 16.3% | 16.3% | .0% |
| | Graduate degree | 9.5% | 9.5% | .0% |
| Income | Less than $25,000 | 17.5% | 17.5% | .0% |
| | $25,000-$39,999 | 12.3% | 12.3% | .0% |
| | $40,000-$84,999 | 34.2% | 34.2% | .0% |
| | $85,000-$174,999 | 28.2% | 28.2% | .0% |
| | $175,000 or more | 7.9% | 7.9% | .0% |
| Marital status | Married | 56.1% | 56.1% | .0% |
| | Widowed | 7.4% | 7.4% | .0% |
| | *Divorced* | 10.6% | 10.4% | .2% |
| | *Separated* | 1.8% | 2.0% | -.2% |
| | Never married | 24.1% | 24.1% | .0% |

Italicized categories within a demographic variable were combined when weighting panel respondents.

**Table OA3.2**. Distributions of Demographic Variables among Weighted Panel Respondents With a Residential Address in One of Six Target States and Among the Populations of Adult U.S. Citizens Living In a Household Served By a Landline In Those States

| | | | Weighted panel respondents | Landline population | Difference |
|---|---|---|---|---|---|
| California | Sex | Male | 48.7% | 48.7% | .0% |
| | | Female | 51.3% | 51.3% | .0% |
| | Age | 18-24 | 11.8% | 11.8% | .0% |
| | | 25-34 | 15.4% | 15.4% | .0% |
| | | 35-44 | 20.5% | 20.5% | .0% |
| | | 45-54 | 20.2% | 20.2% | .0% |
| | | 55-64 | 14.7% | 14.7% | .0% |
| | | 65+ | 17.4% | 17.4% | .0% |
| | Ethnicity | Hispanic | 30.6% | 30.6% | .0% |
| | | Non-Hispanic | 69.4% | 69.4% | .0% |
| | Race | White | 62.8% | 62.7% | .0% |
| | | Black | 5.8% | 5.8% | .0% |
| | | Other | 31.4% | 31.4% | .0% |
| | Education | *No high school diploma* | 6.3% | 19.8% | -13.5% |
| | | *High school diploma* | 37.3% | 23.9% | 13.4% |
| | | Some college, no bachelor's degree | 29.4% | 29.4% | .0% |
| | | Bachelor's degree | 17.1% | 17.1% | .0% |
| | | Graduate degree | 9.8% | 9.8% | .0% |
| | Income | Less than $25,000 | 15.1% | 15.1% | .0% |
| | | $25,000-$39,999 | 11.0% | 11.0% | .0% |
| | | $40,000-$84,999 | 32.0% | 32.0% | .0% |
| | | $85,000-$174,999 | 31.6% | 31.6% | .0% |
| | | $175,000 or more | 10.3% | 10.3% | .0% |
| | Marital status | Married | 54.8% | 54.8% | .0% |
| | | Widowed | 6.5% | 6.5% | .0% |

| | | | | | |
|---|---|---|---|---|---|
| | | *Divorced* | 8.6% | 9.5% | -.9% |
| | | *Separated* | 3.1% | 2.2% | .9% |
| | | Never married | 27.0% | 27.0% | .0% |
| Florida | Sex | Male | 47.6% | 47.6% | .0% |
| | | Female | 52.4% | 52.4% | .0% |
| | Age | *18-24* | 13.1% | 9.2% | 3.9% |
| | | *25-34* | 8.6% | 12.5% | -3.9% |
| | | 35-44 | 17.9% | 17.9% | .0% |
| | | 45-54 | 19.1% | 19.1% | .0% |
| | | 55-64 | 16.2% | 16.2% | .0% |
| | | 65+ | 25.1% | 25.1% | .0% |
| | Ethnicity | Hispanic | 18.6% | 18.6% | .0% |
| | | Non-Hispanic | 81.4% | 81.4% | .0% |
| | Race | White | 79.6% | 79.6% | .0% |
| | | Black | 13.3% | 13.3% | .0% |
| | | Other | 7.2% | 7.2% | .0% |
| | Education | *No high school diploma* | 7.4% | 15.8% | -8.4% |
| | | *High school diploma* | 39.9% | 31.5% | 8.4% |
| | | Some college, no bachelor's degree | 28.6% | 28.6% | .0% |
| | | Bachelor's degree | 15.5% | 15.5% | .0% |
| | | Graduate degree | 8.6% | 8.6% | .0% |
| | Income | Less than $25,000 | 18.6% | 18.6% | .0% |
| | | $25,000-$39,999 | 14.6% | 14.6% | .0% |
| | | $40,000-$84,999 | 36.7% | 36.7% | .0% |
| | | $85,000-$174,999 | 24.0% | 24.0% | .0% |
| | | $175,000 or more | 6.1% | 6.1% | .0% |
| | Marital status | Married | 55.8% | 55.8% | .0% |
| | | Widowed | 8.9% | 8.9% | .0% |
| | | *Divorced* | 14.0% | 11.9% | 2.1% |
| | | *Separated* | .1% | 2.1% | -2.1% |
| | | Never married | 21.3% | 21.3% | .0% |

| | | | | | |
|---|---|---|---|---|---|
| New York | Sex | Male | 46.8% | 46.8% | .0% |
| | | Female | 53.2% | 53.2% | .0% |
| | Age | *18-24* | 14.9% | 11.0% | 3.9% |
| | | *25-34* | 9.4% | 13.3% | -3.9% |
| | | 35-44 | 19.3% | 19.3% | .0% |
| | | 45-54 | 20.6% | 20.6% | .0% |
| | | 55-64 | 16.0% | 16.0% | .0% |
| | | 65+ | 19.9% | 19.9% | .0% |
| | Ethnicity | *Hispanic* | 13.5% | 14.5% | -1.0% |
| | | *Non-Hispanic* | 86.6% | 85.5% | 1.0% |
| | Race | White | 68.8% | 68.8% | .0% |
| | | Black | 14.2% | 14.2% | .0% |
| | | Other | 17.0% | 17.0% | .0% |
| | Education | *No high school diploma* | 7.0% | 16.1% | -9.1% |
| | | *High school diploma* | 38.1% | 29.0% | 9.1% |
| | | Some college, no bachelor's degree | 25.1% | 25.1% | .0% |
| | | Bachelor's degree | 17.1% | 17.1% | .0% |
| | | Graduate degree | 12.7% | 12.7% | .0% |
| | Income | Less than $25,000 | 17.8% | 17.8% | .0% |
| | | $25,000-$39,999 | 11.0% | 11.0% | .0% |
| | | $40,000-$84,999 | 31.6% | 31.6% | .0% |
| | | $85,000-$174,999 | 29.5% | 29.5% | .0% |
| | | $175,000 or more | 10.0% | 10.0% | .0% |
| | Marital status | Married | 52.1% | 52.1% | .0% |
| | | Widowed | 8.0% | 8.0% | .0% |
| | | *Divorced* | 10.3% | 8.3% | 2.0% |
| | | *Separated* | .7% | 2.7% | -2.0% |
| | | Never married | 29.0% | 29.0% | .0% |
| North Carolina | Sex | Male | 47.3% | 47.3% | .0% |
| | | Female | 52.7% | 52.7% | .0% |
| | Age | *18-24* | .0% | 10.7% | -10.7% |

| | | | | | |
|---|---|---|---|---|---|
| | | *25-34* | 24.7% | 14.1% | 10.7% |
| | | 35-44 | 19.8% | 19.8% | .0% |
| | | 45-54 | 20.1% | 20.1% | .0% |
| | | 55-64 | 16.5% | 16.5% | .0% |
| | | 65+ | 18.9% | 18.9% | .0% |
| | Ethnicity | *Hispanic* | .0% | 5.4% | -5.4% |
| | | *Non-Hispanic* | 100.0% | 94.6% | 5.4% |
| | Race | White | 73.0% | 73.0% | .0% |
| | | *Black* | 26.1% | 19.6% | 6.5% |
| | | *Other* | .8% | 7.3% | -6.5% |
| | Education | *No high school diploma* | 2.6% | 17.4% | -14.8% |
| | | *High school diploma* | 44.5% | 29.7% | 14.8% |
| | | Some college, no bachelor's degree | 29.0% | 29.0% | .0% |
| | | Bachelor's degree | 15.9% | 15.9% | .0% |
| | | Graduate degree | 8.0% | 8.0% | .0% |
| | Income | Less than $25,000 | 21.6% | 21.6% | .0% |
| | | $25,000-$39,999 | 14.7% | 14.7% | .0% |
| | | $40,000-$84,999 | 36.5% | 36.5% | .0% |
| | | *$85,000-$174,999* | 26.3% | 22.0% | 4.3% |
| | | *$175,000 or more* | .9% | 5.2% | -4.3% |
| | Marital status | Married | 57.9% | 58.0% | .0% |
| | | Widowed | 7.6% | 7.6% | .0% |
| | | *Divorced* | 3.6% | 9.8% | -6.2% |
| | | *Separated* | 9.1% | 2.9% | 6.2% |
| | | Never married | 21.7% | 21.7% | .0% |
| Ohio | Sex | Male | 48.5% | 47.2% | 1.3% |
| | | Female | 51.5% | 52.9% | -1.3% |
| | Age | 18-24 | 10.2% | 10.2% | .0% |
| | | 25-34 | 13.1% | 13.1% | .0% |
| | | 35-44 | 18.5% | 18.5% | .0% |
| | | 45-54 | 21.2% | 21.2% | .0% |

| | | | | | |
|---|---|---|---|---|---|
| | | 55-64 | 16.4% | 16.4% | .0% |
| | | 65+ | 20.6% | 20.6% | .0% |
| | Ethnicity | *Hispanic* | .7% | 2.0% | -1.3% |
| | | *Non-Hispanic* | 99.3% | 98.0% | 1.3% |
| | Race | White | 88.4% | 86.0% | 2.5% |
| | | *Black* | 8.6% | 10.5% | -1.8% |
| | | *Other* | 2.9% | 3.6% | -.6% |
| | Education | No high school diploma | 13.5% | 13.5% | .0% |
| | | High school diploma | 36.2% | 36.2% | .0% |
| | | Some college, no bachelor's degree | 27.9% | 27.9% | .0% |
| | | Bachelor's degree | 14.2% | 14.2% | .0% |
| | | Graduate degree | 8.2% | 8.2% | .0% |
| | Income | Less than $25,000 | 19.5% | 19.5% | .0% |
| | | $25,000-$39,999 | 14.0% | 14.0% | .0% |
| | | $40,000-$84,999 | 37.1% | 37.1% | .0% |
| | | *$85,000-$174,999* | 27.3% | 24.5% | 2.9% |
| | | *$175,000 or more* | 2.1% | 5.0% | -2.9% |
| | Marital status | Married | 56.6% | 56.6% | .0% |
| | | Widowed | 8.1% | 8.1% | .0% |
| | | *Divorced* | 12.7% | 11.4% | 1.3% |
| | | *Separated* | .4% | 1.7% | -1.3% |
| | | Never married | 22.3% | 22.3% | .0% |
| Pennsylvania | Sex | Male | 47.0% | 47.0% | .0% |
| | | Female | 53.0% | 53.0% | .0% |
| | Age | *18-24* | 6.0% | 10.2% | -4.1% |
| | | *25-34* | 16.0% | 11.8% | 4.1% |
| | | 35-44 | 18.1% | 18.1% | .0% |
| | | 45-54 | 20.9% | 20.9% | .0% |
| | | 55-64 | 16.4% | 16.4% | .0% |
| | | 65+ | 22.6% | 22.6% | .0% |
| | Ethnicity | *Hispanic* | 2.0% | 3.5% | -1.5% |

| | | | | |
|---|---|---|---|---|
| | *Non-Hispanic* | 98.0% | 96.5% | 1.5% |
| Race | White | 85.9% | 85.9% | .0% |
| | *Black* | 10.0% | 9.1% | .9% |
| | *Other* | 4.2% | 5.0% | -.9% |
| Education | No high school diploma | 13.5% | 13.5% | .0% |
| | High school diploma | 37.7% | 37.7% | .0% |
| | Some college, no bachelor's degree | 24.4% | 24.4% | .0% |
| | Bachelor's degree | 15.2% | 15.2% | .0% |
| | Graduate degree | 9.3% | 9.3% | .0% |
| Income | Less than $25,000 | 19.6% | 19.6% | .0% |
| | $25,000-$39,999 | 13.1% | 13.1% | .0% |
| | $40,000-$84,999 | 35.4% | 35.4% | .0% |
| | $85,000-$174,999 | 25.9% | 25.9% | .0% |
| | $175,000 or more | 6.1% | 6.1% | .0% |
| Marital status | Married | 19.6% | 19.6% | .0% |
| | *Widowed* | 6.4% | 9.1% | -2.7% |
| | *Divorced* | 9.9% | 8.9% | 1.1% |
| | *Separated* | 3.6% | 2.0% | 1.6% |
| | Never married | 23.7% | 23.7% | .0% |

Italicized categories within a demographic variable and state were combined when weighting panel respondents.

**Table OA4.1**. State Adult Populations by Age and Sex

| Sex | Age group | California | Florida | New York | North Carolina | Ohio | Pennsylvania |
|---|---|---|---|---|---|---|---|
| | | State | | | | | |
| Male | 18-24 | 1,987,094 | 824,854 | 1,007,173 | 462,325 | 553,430 | 607,766 |
| | | *(7.4%)* | *(5.8%)* | *(6.7%)* | *(6.8%)* | *(6.4%)* | *(6.3%)* |
| | 25-34 | 2,708,685 | 1,155,412 | 1,249,413 | 605,771 | 720,666 | 731,149 |
| | | *(10.0%)* | *(8.2%)* | *(8.3%)* | *(8.9%)* | *(8.3%)* | *(7.6%)* |
| | 35-44 | 2,768,549 | 1,276,221 | 1,410,778 | 663,408 | 795,272 | 861,716 |
| | | *(10.2%)* | *(9.0%)* | *(9.4%)* | *(9.7%)* | *(9.1%)* | *(9.0%)* |
| | 45-54 | 2,523,223 | 1,255,683 | 1,412,507 | 626,602 | 857,673 | 940,375 |
| | | *(9.3%)* | *(8.9%)* | *(9.4%)* | *(9.2%)* | *(9.8%)* | *(9.8%)* |
| | 55-64 | 1,712,141 | 997,599 | 1,015,932 | 482,382 | 626,807 | 693,024 |
| | | *(6.3%)* | *(7.0%)* | *(6.8%)* | *(7.1%)* | *(7.2%)* | *(7.2%)* |
| | 65+ | 1,711,099 | 1,360,670 | 1,050,880 | 456,073 | 637,996 | 774,983 |
| | | *(6.3%)* | *(9.6%)* | *(7.0%)* | *(6.7%)* | *(7.3%)* | *(8.1%)* |
| Female | 18-24 | 1,826,278 | 772,194 | 978,582 | 415,671 | 528,751 | 585,274 |
| | | *(6.8%)* | *(5.5%)* | *(6.5%)* | *(6.1%)* | *(6.1%)* | *(6.1%)* |
| | 25-34 | 2,477,634 | 1,092,561 | 1,259,394 | 602,073 | 722,659 | 714,788 |
| | | *(9.2%)* | *(7.7%)* | *(8.4%)* | *(8.8%)* | *(8.3%)* | *(7.4%)* |
| | 35-44 | 2,655,779 | 1,265,688 | 1,452,235 | 675,343 | 804,370 | 876,008 |
| | | *(9.8%)* | *(8.9%)* | *(9.7%)* | *(9.9%)* | *(9.2%)* | *(9.1%)* |
| | 45-54 | 2,543,917 | 1,318,274 | 1,481,827 | 662,614 | 888,431 | 969,781 |
| | | *(9.4%)* | *(9.3%)* | *(9.9%)* | *(9.7%)* | *(10.2%)* | *(10.1%)* |
| | 55-64 | 1,829,343 | 1,109,002 | 1,143,469 | 534,159 | 672,690 | 750,347 |
| | | *(6.8%)* | *(7.8%)* | *(7.6%)* | *(7.8%)* | *(7.7%)* | *(7.8%)* |
| | 65+ | 2,292,146 | 1,743,396 | 1,510,818 | 652,280 | 912,460 | 1,118,192 |
| | | *(8.5%)* | *(12.3%)* | *(10.1%)* | *(9.5%)* | *(10.5%)* | *(11.6%)* |
| | TOTALS | 27,035,888 | 27,035,888 | 14,171,554 | 14,973,008 | 6,838,701 | 8,721,205 |
| | | *(100.0%)* | *(100.0%)* | *(100.0%)* | *(100.0%)* | *(100.0%)* | *(100.0%)* |

**Table OA4.2**. People Living In a U.S. Household Served By a Landline by Age and Sex

| Sex | Age group | U.S population ($N_g$) | Number of people living in a household served by a landline ($LL_g$) | Percent of the demographic group living in a household served by a landline ($\%LL_g$) |
|---|---|---|---|---|
| Male | 18-24 | 13,657,368 | 9,421,471 | 69.0% |
| | 25-34 | 18,717,219 | 11,750,659 | 62.8% |
| | 34-44 | 19,825,211 | 16,231,637 | 81.9% |
| | 45-54 | 20,951,391 | 18,351,873 | 87.6% |
| | 55-64 | 15,726,208 | 14,358,256 | 91.3% |
| | 65+ | 15,648,465 | 14,983,276 | 95.8% |
| Female | 18-24 | 13,474,712 | 9,159,234 | 68.0% |
| | 25-34 | 19,070,683 | 13,247,159 | 69.5% |
| | 34-44 | 20,489,648 | 17,510,962 | 85.5% |
| | 45-54 | 21,899,609 | 19,510,205 | 89.1% |
| | 55-64 | 17,006,092 | 15,842,650 | 93.2% |
| | 65+ | 20,914,629 | 20,385,718 | 97.5% |
| Total | | 217,381,235 | 180,753,100 | 83.2% |

**Table OA4.3**. State Population Estimation Numbers

| State | Population estimation values | | | | |
|---|---|---|---|---|---|
| | Voting age population of the state ($VAP_s$) | Imprisoned population in the state ($INC_s$) | Percent of the state population that was non-U. S. citizen ($\%NC_s$) | Estimated percent of the population in landline households ($e\%LL_s$) | Estimated Target Population ($N_s$) |
| California | 27,279,556 | 173,670 | 17.9% | 82.1% | 18,449,926 |
| Florida | 14,395,399 | 102,388 | 11.2% | 84.2% | 10,763,124 |
| New York | 15,048,837 | 60,347 | 12.4% | 83.1% | 11,038,294 |
| North Carolina | 7,029,536 | 39,482 | 5.9% | 82.9% | 5,475,561 |
| Ohio | 8,802,396 | 51,686 | 2.2% | 83.4% | 7,111,379 |
| Pennsylvania | 9,790,263 | 49,215 | 3.1% | 83.9% | 8,003,081 |

**Table OA4.4**. Original and Estimated Sex and Age Distributions in State Records

| | | California | | Florida | | New York | | North Carolina | | Ohio | | Pennsylvania | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sex | Age | Records with data | Estimated total ($eRR_{gs}$) | Original | Estimated total ($eRR_{gs}$) | Original | Estimated total ($eRR_{gs}$) | Original | Estimated total ($eRR_{gs}$) | Original | Estimated total ($eRR_{gs}$) | Original | Estimated total ($eRR_{gs}$) |
| Male | 18-24 | 631,316 | 938,390 | 579,847 | 615,318 | 565,353 | 567,779 | 306,336 | 313,803 | | | 157,469 | 445,740 |
| | 25-34 | 945,879 | 1,304,193 | 873,724 | 900,810 | 973,714 | 977,894 | 474,256 | 483,207 | | | 387,415 | 662,853 |
| | 35-44 | 1,019,312 | 1,408,503 | 978,626 | 1,000,002 | 1,077,911 | 1,082,386 | 547,725 | 555,807 | | | 428,475 | 726,550 |
| | 45-54 | 1,130,131 | 1,653,301 | 1,076,231 | 1,096,806 | 1,148,498 | 1,153,187 | 557,219 | 564,207 | | | 488,525 | 810,693 |
| | 55-64 | 834,581 | 1,316,189 | 897,024 | 912,139 | 905,636 | 909,380 | 449,898 | 454,797 | | | 389,761 | 630,592 |
| | 65+ | 761,168 | 1,300,918 | 1,276,358 | 1,296,033 | 1,075,521 | 1,080,384 | 461,501 | 466,264 | | | 424,549 | 699,158 |
| | Missing | *24,725* | | *24,592* | | *24,343* | | *12,567* | | | | *4,986* | |
| Female | 18-24 | 782,584 | 1,139,522 | 647,440 | 688,535 | 693,649 | 697,090 | 351,831 | 361,129 | | | 190,950 | 514,982 |
| | 25-34 | 1,135,535 | 1,552,316 | 1,008,357 | 1,039,712 | 1,195,560 | 1,201,488 | 585,191 | 596,486 | | | 415,678 | 725,288 |
| | 35-44 | 1,144,299 | 1,596,932 | 1,087,902 | 1,112,631 | 1,245,990 | 1,252,337 | 636,433 | 646,757 | | | 441,767 | 776,822 |
| | 45-54 | 1,222,221 | 1,830,309 | 1,205,942 | 1,229,737 | 1,286,490 | 1,293,141 | 638,122 | 647,143 | | | 519,171 | 881,308 |
| | 55-64 | 910,353 | 1,469,856 | 1,042,944 | 1,060,418 | 1,040,004 | 1,045,315 | 519,121 | 525,518 | | | 422,884 | 693,593 |
| | 65+ | 956,887 | 1,583,781 | 1,595,991 | 1,618,727 | 1,451,493 | 1,458,390 | 609,339 | 615,629 | | | 568,059 | 876,737 |
| | Missing | *39,824* | | *28,178* | | *34,534* | | *18,457* | | - | | *5,635* | |
| Missing | 18-24 | *632,573* | | *69,501* | | *11* | | *12,663* | | 1,003,929 | 1,017,238 | *606,898* | |
| | 25-34 | *731,876* | | *47,927* | | *17* | | *13,684* | | 1,464,559 | 1,483,974 | *577,238* | |
| | 35-44 | *796,353* | | *34,659* | | *12* | | *11,096* | | 1,451,299 | 1,470,538 | *624,672* | |
| | 45-54 | *1,078,552* | | *31,765* | | *13* | | *8,647* | | 1,602,658 | 1,623,904 | *674,785* | |
| | 55-64 | *998,959* | | *21,902* | | *4* | | *5,338* | | 1,235,160 | 1,251,534 | *504,090* | |
| | 65+ | *1,122,999* | | *26,619* | | *5* | | *4,478* | | 1,381,382 | 1,399,694 | *574,421* | |
| | Missing | *194,082* | | *15,340* | | *13* | | *6,847* | | *107,894* | | *36,889* | |

Italicized values are the number of records with missing age, sex, or age and sex data.

**Table OA4.5**. State Registration Population Estimation Numbers

| State | $RR_s$ source | Population estimation values: Number of registration records ($RR_s$) | Estimated percent of the population in landline households ($e\%LL_s$) | Voting eligible population ($VEP_s$) | Sub-population of VEP living overseas ($OE_s$) | Percent of VEP living overseas ($\%OE_s$) | Estimated percent of records for registrants residing outside the state ($e\%OSR_s$) | Estimated number of people living in a landline household who were registered ($eRRLL_s$) |
|---|---|---|---|---|---|---|---|---|
| California | Registration records | 17,094,209 | 83.5% | 27,279,556 | 486,207 | 1.8% | 4.0% | 13,458,719 |
| | Published statistics | 17,304,091 | 83.5% | 27,279,556 | 486,207 | 1.8% | 4.0% | 13,623,965 |
| Florida | Registration records | 12,570,869 | 84.7% | 14,395,399 | 451,907 | 3.1% | 4.1% | 9,885,271 |
| | Published statistics | 11,247,634 | 84.7% | 14,395,399 | 451,907 | 3.1% | 4.1% | 8,844,727 |
| New York | Registration records | 11,660,114 | 84.1% | 15,048,837 | 263,787 | 1.8% | 4.3% | 9,219,284 |
| | Published statistics | 12,031,312 | 84.1% | 15,048,837 | 263,787 | 1.8% | 4.3% | 9,512,779 |
| North Carolina | Registration records | 6,154,773 | 83.6% | 7,029,536 | 133,483 | 1.9% | 1.2% | 4,987,424 |
| | Published statistics | 6,226,204 | 83.6% | 7,029,536 | 133,483 | 1.9% | 1.2% | 5,045,307 |
| Ohio | Registration records | 8,246,881 | 83.1% | 8,802,396 | 174,703 | 2.0% | 3.4% | 6,490,918 |
| | Published statistics | 8,287,665 | 83.1% | 8,802,396 | 174,703 | 2.0% | 3.4% | 6,523,018 |
| Pennsylvania | Registration records | 8,444,317 | 83.8% | 9,790,263 | 203,791 | 2.1% | 3.1% | 6,715,175 |
| | Published statistics | 8,755,588 | 83.8% | 9,790,263 | 203,791 | 2.1% | 3.1% | 6,962,707 |

**Table OA4.6**. State-Level Turnout

| State | Number of people who voted: Number according to statistics published by the state ($NV_s$) | Number of people who voted: Estimated number living in the U.S. in a household served by a landline ($TO_s$) |
|---|---|---|
| California | 13,743,177 | 11,191,383 |
| Florida | 8,456,329 | 6,945,853 |
| New York | 7,722,019 | 6,378,107 |
| North Carolina | 4,353,739 | 3,555,663 |
| Ohio | 5,773,777 | 4,702,216 |
| Pennsylvania | 5,995,137 | 4,977,465 |

**Table OA6.1**. Estimates of the Number of People Who Were Registered In a State Using Deadwood Rates from Ansolabehere and Hersh (2010)

| State | Source of the number of a state's registration records | Estimated number of people living in a landline household who were registered (*eRRLL*$_s$) | Ansolabehere and Hersh (2010) deadwood rates (*%DW*$_s$) | Estimated non-deadwood records for people living in a landline household (*eRRNDLL*$_s$) |
|---|---|---|---|---|
| California | People's records | 13,458,719 | .3% | 13,414,791 |
| | Published statistics | 13,623,965 | .3% | 13,579,498 |
| Florida | People's records | 9,885,271 | 7.9% | 9,105,621 |
| | Published statistics | 8,844,727 | 7.9% | 8,147,145 |
| New York | People's records | 9,219,284 | 5.2% | 8,743,091 |
| | Published statistics | 9,512,779 | 5.2% | 9,021,426 |
| North Carolina | People's records | 4,987,424 | 4.9% | 4,740,899 |
| | Published statistics | 5,045,307 | 4.9% | 4,795,921 |
| Ohio | People's records | 6,490,918 | .9% | 6,435,831 |
| | Published statistics | 6,523,018 | .9% | 6,467,659 |
| Pennsylvania | People's records | 6,715,175 | 6.4% | 6,283,982 |
| | Published statistics | 6,962,707 | 6.4% | 6,515,620 |

**Table OA6.2**. Estimated Rates of Registration According To the States' Published Numbers of People Registered With and Without Deadwood Adjustment

| State | Estimated registration rate among people living in a household with a landline | | |
|---|---|---|---|
| | Without deadwood adjustment | With deadwood adjustment | Difference |
| All target states | 83.7% | 80.4% | 3.3% |
| California | 74.6% | 74.4% | .2% |
| Florida | 82.8% | 76.3% | 6.5% |
| New York | 87.2% | 82.7% | 4.5% |
| North Carolina | 92.5% | 87.9% | 4.6% |
| Ohio | 91.4% | 90.6% | .8% |
| Pennsylvania | 87.9% | 82.2% | 5.6% |

**Table OA6.3**. Estimated Turnout Rates among People Registered To Vote According To the States' Published Numbers of People Registered With and Without Deadwood Adjustment

| State | Estimated turnout rate among people living in a household with a landline and registered to vote: Without deadwood adjustment | With deadwood adjustment | Difference |
|---|---|---|---|
| All target states | 75.0% | 72.1% | 3.0% |
| California | 79.7% | 79.4% | .3% |
| Florida | 81.6% | 75.2% | 6.4% |
| New York | 67.7% | 64.2% | 3.5% |
| North Carolina | 73.6% | 69.9% | 3.6% |
| Ohio | 70.3% | 69.7% | .6% |
| Pennsylvania | 73.2% | 68.5% | 4.7% |