What's Wrong with the CPS?

Michael P. McDonald

University of Florida

michael.mcdonald@ufl.edu • www.electproject.org •@ElectProject

Abstract

Responses to the Census Bureau's Voting and Registration Supplement have missing data: people who are never administered the voting supplement, people who refuse to answer the voting question, and those who don't know if they - or a person they are reporting for by proxy - voted. The Census Bureau scores these response items as persons who did not vote when computing turnout rates. In contrast, the scholarly community typically applies listwise deletion to compute turnout rates. Recently, Hur and Achen (2013) propose to further reweight the listwise deleted data to state VEP turnout rates to compute more accurate CPS turnout rates. I investigate these missing data mechanisms and show they are related to administration of the CPS, although I do not propose my own correction, yet. I then demonstrate how the various missing data corrections have profound consequences on perceptions of turnout rates. For example, where the Census Bureau reports African-American turnout first exceeded non-Hispanic White turnout in 2012, African-American turnout was higher in 2008 when applying listwise deletion and the Hur and Achen correction. Chief Justice Roberts in the 2013 *Shelby County* case overturning important provisions of the Voting Rights Act cited Census Bureau statistics to argue Mississippi African-American turnout rates are higher than those in Massachusetts, but this is not observed among the other missing data corrections. Even recently published *AJPS* research (Burden, et al. 2013) investigating turnout effects of election laws is sensitive to missing data specification.

Prepared for the 2014 American Political Science Association Conference, Washington, DC August 27-31

Introduction

The Current Population Survey Voting and Registration Supplement (CPS) is among the most utilized election surveys. A simple search yields an impressive 30,807 articles citing the survey.[1] One attractive feature of the CPS is its large sample size, typically at least one thousand respondents for each state, which permits robust comparisons of turnout rates among states and among key demographic sub-groups. Another often-lauded feature of the CPS is its apparent low vote over-report bias. For example, in the 2012 presidential election, the Census Bureau reported from the CPS a citizen voting-age population (CVAP) turnout rate of 61.8% compared to voting-eligible (VEP) turnout rate of 58.7%. In comparison, the American National Election Study reported a 2012 turnout rate of 78%.[2]

A closer look at the CPS reveals troubling patterns. The 2012 CPS reports Mississippi had the highest CVAP turnout rate of 74.5%, while its VEP turnout rate was just above the nation's average, at 59.7%, a difference of 14.8 percentage points. In contrast, the CPS reports Massachusetts as having a turnout rate of 70.9%, 4.6 percentage points higher than its VEP turnout rate of 66.3%. These discrepancies should be alarming to consumers of the CPS, especially those that use the CPS to estimate determinants of voting and registration. They affect politics, too. Chief Justice Roberts, when questioning the relevance of the Voting Rights Act's Section 4 coverage formula in the landmark 2013 case *Shelby County v. Holder*, asked the Solicitor General of the United States why Section 5 voting protections were necessary if African-Americans living in Mississippi had higher turnout rates than those living in Massachusetts.

An unusual data Census Bureau data practice may explain these patterns. In 2012, 12.8% of CPS respondents were missing data to the vote question. These respondents either were never administered the voting supplement, refused to provide a response, or did not know the answer for themselves or a proxy household member. The Census Bureau scores these respondents as did not vote, incredibly even the 9.7% who were never asked the voting question. Those in the scholarly community mostly - but not all, as one study explored herein demonstrates - apply listwise deletion to these missing data. Hur and Achen (2013) propose a further weighting correction to match the CPS listwise deleted turnout rates to McDonald's (2011) VEP turnout rates. These competing corrections produce different turnout estimates, raising troubling concerns not only about published research, but important policy matters based on the Census Bureau statistics, like the Supreme Court's decision regarding the Voting Rights Act.

Missing Data and Turnout Rates

The Current Population Survey is a monthly survey of the domestic non-institutional civilian population that the federal government administers to determine monthly national and state employment rates. The biennial November Voting and Registration Supplement is one among many

[1] JSTOR search terms ((((Current) AND (Population)) AND (Survey)) AND (Voting)), accessed February 14, 2014.
[2] There are some differences between these rates' denominators, but these differences cannot explain the apparently higher vote over-report bias of the ANES compared to the CPS. The CPS sample frame is the non-institutional domestic institutional population of the United States. The American National Election Study has a similar sample frame. The voting-eligible population includes some institutional populations, such as school dormitories and other group quarters, and also includes overseas citizens (McDonald 2005).

supplements to the main questionnaire that the Census Bureau administers on varying schedules, ranging across topics such as disability, fertility, volunteering, food security, and internet usage. Since the government administers the CPS, the voting and registration supplement has only a limited number questions that exclude many potentially politically sensitive behavioral and attitudinal questions found on other election surveys known to be correlated with voter participation.

The Census Bureau produces biennial voting and registration reports with statistics calculated from the CPS voting supplement. There are a number of nuances to the way the Census Bureau calculates these statistics that require a deep dive into how the CPS main survey and Voting and Registration Supplement have evolved since the first supplement was administered in 1964. As an unfortunate accident of history the earliest supplement data have been lost to time; the November, 1972 election marks the first publicly available individual level dataset.

To begin, consider response items which describe respondents' answers and participation in the Voting and Registration Supplement questions.[3] For the vote question, respondents to the main employment questionnaire who are not citizens or are under the age of eighteen are coded *Not in Universe*. Respondents who are never administered the supplement are coded *No Response* (also referred to in CPS documentation as *Noninterview* or *No Answer*). For all surveys except 1972, 1982, and from 1994 onward respondents who refuse to answer a question are also coded *No Response*. For 1972, 1982, and from 1994 onward refusals are coded into a separate category, *Refused*; in other surveys refusals are coded as *No Response*. Respondents who report they voted are coded *Voted*. Those who report they did not vote are coded *Did Not Vote*. Respondents who do not know the answer for themselves or a household member they are providing a proxy response for are coded *Don't Know*.

Figure 1 plots the magnitude of the various response items to the CPS vote question from 1972 to 2012. Respondents *Not in Universe* are omitted.[4] Where *Refused* is not a permitted response item, these responses are grouped with *No Response*. The VEP turnout rate is plotted as a solid line for reference (McDonald and Popkin 2001; McDonald 2011). From 1972 to 1994, the *No Response*, *Refused* and *Don't Know* categories are a relatively consistent share of the response items to the vote question, averaging 5.2 percent during this period (unweighted). After 1994, these response items account for a growing share of the response items to the vote question, reaching a high of 16.0 percent in 2010 and then reducing to 12.8 percent in 2012. The growth is almost entirely due to a growing percentage of respondents in the *No Response* category, people who are never administered the survey, reaching 12.5 percent in 2010 and then reducing to 9.7 percent in 2012.

---

[3] Note, all pre 1994 data files were obtained from the ICPSR archive. All 1994 to current data files were obtained directly from the Census Bureau through the Data Ferrett (sic) service available at http://dataferrett.census.gov/ (first accessed Feb. 25, 2007). Data files on the Data Ferrett service extract concurrently variables and records from the main employment survey and supplements. By design of this data distribution mechanism, no filter is applied to any 1994 thru current data files.

[4] From 1976 to 1980 and again in 1988, *Not in Universe* appears to be conflated with *No Response* as some respondents under age 18 or not a citizen of the United States are included in the *No Response* category. In these years respondents are selected by age and citizenship status to measure the *No Response* category.

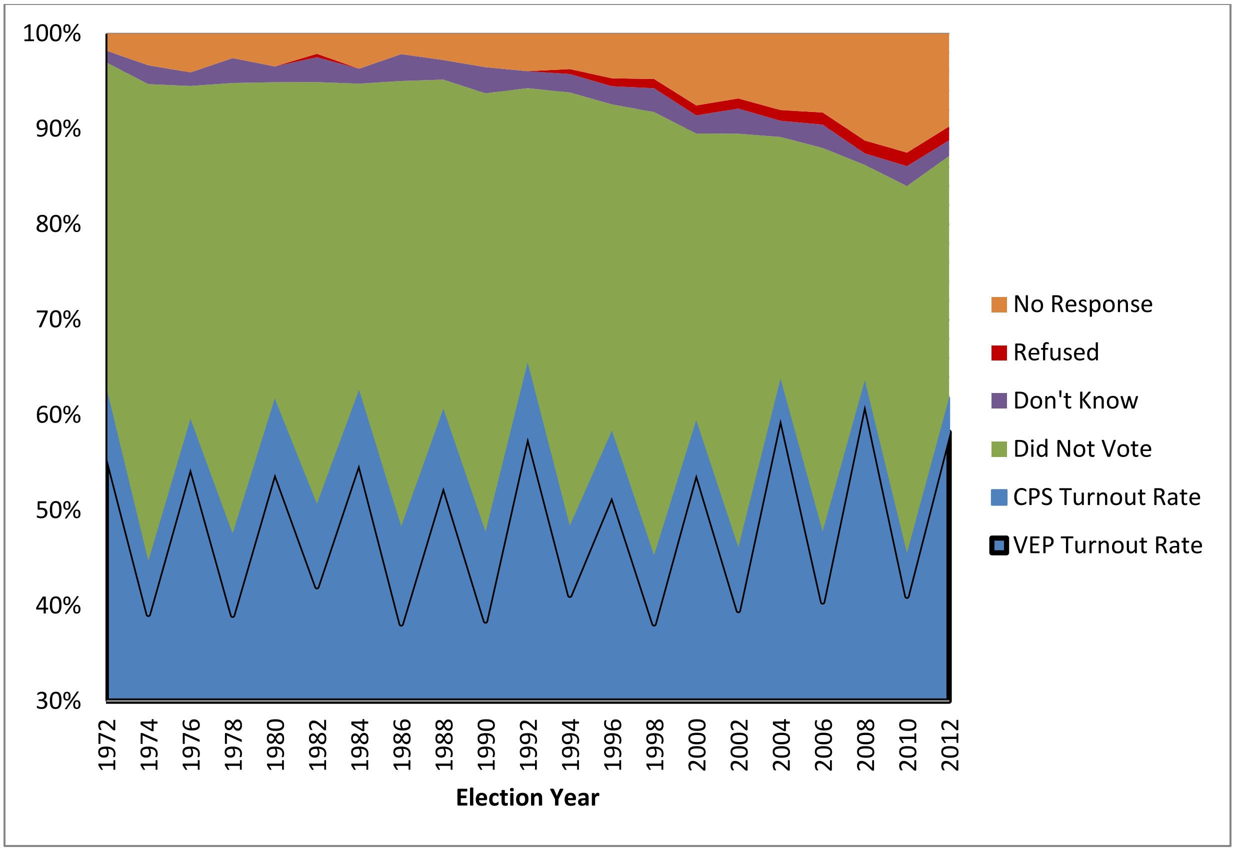

Figure 1. CPS Voting and Registration Vote Question Response Item Components, 1972-2012.

These response items would perhaps be a passing curiosity for many consumers of the CPS, if it were not for an unconventional practice that the Census Bureau engages in. As noted by Hur and Achen (2013), when the Census Bureau reports the citizen voting-age population turnout rate the *No Response*, *Refused* and *Don't Know* categories are classified as *Did Note Vote*. The American National Election Study (2009, 83-85), as well as some scholars who analyze the CPS (e.g., Wolfinger and Rosenstone 1980; Leighley and Nagler 2013) classify these responses as missing data and apply listwise deletion when analyzing voter turnout.

Figure 2 plots the CPS turnout rate from 1972 to 2012, applying listwise deletion to the missing data. When the listwise deletion is applied to the missing data, the CPS has a 2012 CVAP turnout rate of 70.9%, more similar to the ANES turnout rate of 78% than the 61.8% CVAP turnout rate calculated by assuming the missing data are *Did Not Vote* responses (Gara et al. 2010). Furthermore, since the percentage of *No Response* has dramatically increased in recent elections the two missing data treatments illustrated in Figures 1 and 2 produce substantially different turnout rate estimates. The Census Bureau's method of coding missing data results in a reduction in the CPS vote over-report bias over time from 6.9 percentage points in 1972 to 3.6 percentage points in 2012. The scholarly community's method of applying listwise deletion to the missing data indicates that the vote over-report bias is increasing from 8.7 percentage points to 12.7 percentage points. The bloom thus fades from one of the oft-cited virtues of the CPS, its low vote over-report bias, compared to other election surveys.

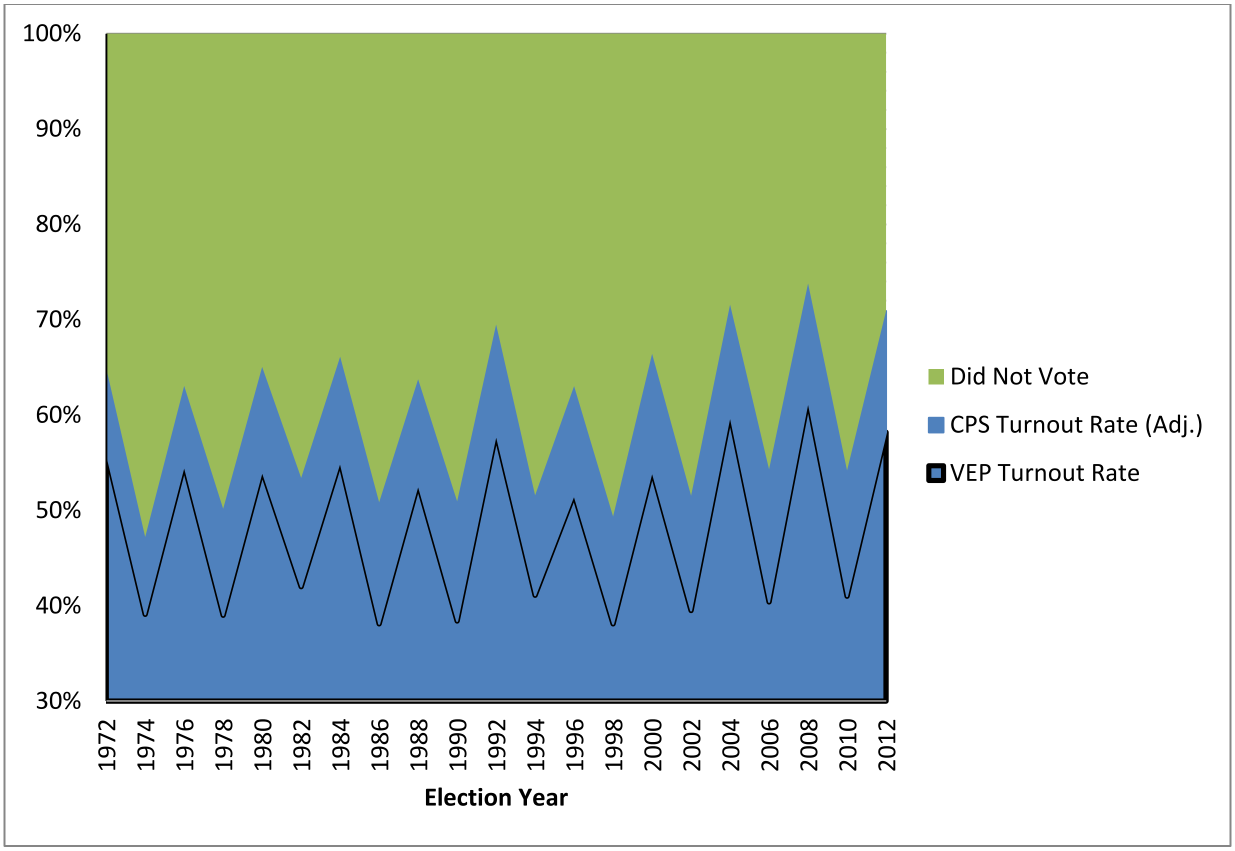

Figure 2. CPS Voting and Registration Vote Question Response Item Components, 1972-2012, Applying Listwise Deletion to Missing Data

Modeling the Missing Data Mechanism

These two competing corrections, in effect, model the missing data mechanism. Define responses to the voting question, Y, where $\{Y = 0 \text{ if respondent } Did\ Not\ Vote\}$ and $\{Y = 1 \text{ if respondent } Voted$. Partition Y into two categories, observed (*O*) and missing (*M*) such that

$$Y = \{Y_O | Y_M\}$$

The Census Bureau assumes all respondents with missing data *Did Not Vote*, or $Y_M = 0$.

The survey research community commonly applies listwise deletion to the missing data, thereby assuming that the distribution of the missing data is the same as the observed data, or $Y_M \sim Y_O$.

For listwise deletion to produce unbiased estimates the data must be missing completely at random (MCAR), that is, any observation has an equal chance of being observed or missing (Little and Rubin 1987, 14). As Hur and Achen (2013, 992) note, this is "unlikely to be [a] perfectly accurate" assumption.

Scholars commonly soften the strong MCAR assumption by assuming that the missing data mechanism is constant across observable subgroups, or that the data are Missing at Random (MAR). Assume that all observations in Y can be exhaustively partitioned into G partitions, then

$$Y_{g_M} \sim Y_{g_O}\ , for\ g = 1 \dots G$$

For example, if upper income respondents have lower (and constant within income strata) probabilities of producing missing data than lower income respondents, then the overall turnout rate can be estimated without bias by applying, within these two subgroups, the turnout rates of the observed to the missing and then aggregating the resulting subgroup estimates. In perhaps the most basic application, assume upper income respondents have higher turnout rates and that half of the respondents are upper income and half are lower income. Upper income respondents have no missing data and half of the lower income respondents have missing data. Apply the turnout rate of lower income respondents with observed data to those with missing data and then average the turnout rates of the two income levels together to estimate the overall turnout rate. Upon reflection one can see why MAR is superior to MCAR in this example: MCAR applies the overall turnout rate to the missing data while MAR applies only the lower income turnout rate to those lower income respondents missing data.

If missing data are MAR, then statistical models that include covariates that identify the partitions where the missing data are MAR will produce unbiased estimates. Thus, scholars who estimate comprehensive multivariate turnout models might, by ingenuity or luck, include the right mix of independent variables to estimate unbiased models, that is if the missing data are MAR.

Little and Rubin (1987, 8) argue, "Knowledge, or absence of knowledge, of the mechanisms that led to certain values being missing is a key element in choosing an appropriate analysis and in interpreting the results." Aspects of the missing data mechanism may be observed by examining aspects of the CPS data collection process such as the interview mode, codified in what are known as paradata. For example, as described in more detail below, the contemporary CPS is a mixed mode survey consisting of phone interviews from centralized call centers and interviews conducted in the field. Respondents are not randomly assigned to an interview mode since, among other reasons, a household must have a phone to be assigned a phone interview. In 2012, 3.8% of interviews conducted by centralized call centers were coded *No Response*, while in-person field interviews resulted in a 11.2% *No Response* rate. If assignment of households to an interview mode is related to the turnout rate and other observable information collected from respondents, the missing data mechanism is not MCAR or MAR, the missing data is what is called non-ignorable. If the missing data are non-ignorable then listwise deletion results in biased estimates; to make unbiased inferences, the missing data mechanism must be modeled.

Exploring the Missing Data Mechanism

Mode effects

Survey researchers describe the method by which interviewers ask questions of respondents as the survey mode. A mode may be face-to-face, phone, mail, or internet. The mode may further describe data collection, such as in the case of face-to-face a confidential paper survey instrument given to respondents to complete versus interviewers asking questions and recording respondents' answers or in the case of phone surveys interviewers questioning respondents person-to-person or automated surveys where responses are collected by respondents pushing buttons. Mode is known to affect the presence of missing data (McDonald and Thornburg 2012).

Beginning in 1987, the Census Bureau began performing CPS interviews through two modes, face-to-face and phone interviews (CPS 2006). Currently, about ten percent of CPS household interviews are computer assisted telephone interviews (CATI) conducted from centralized call centers (CPS 2006, Chpt. 4, p.6). In 1994, the Census Bureau began using computer assisted personal interviewing (CAPI) for its face-to-face interviews. The Census Bureau conducted experiments in 1992 to explore CATI and CAPI mode effects, and at the same time the questionnaire was redesigned. The Census Bureau and found the innovations did not affect the unemployment rate calculation, but did affect other measures, such a duration of unemployment and reported occupation.

The assignment of respondents to CATI and CAPI is not random (CPS 2006, Chpt. 4, p.6). Households in the first and fifth months are generally not eligible for CATI (the fifth month interview is inaptly named since eight months pass between the fourth and fifth month). Respondents are asked in the first month if they are willing to accept phone interviews. Then, the regional office assigns roughly ten percent of households for CATI. The Census Bureau documentation is unclear on exactly how the assignment occurs, but criteria include if a field interviewer requests CATI or if a field interviewer is ill. If an interview cannot be completed by CATI, it is assigned to field interviewers for additional contact attempts.

The CATI and CAPI modes are an important part of the CPS data collection. However, telephone interviews are conducted in the field. CAPI field interviewers may, at their discretion, conduct a phone interview if a respondent has given approval in the first month interview (CPS 2006 Chpt. 7, p.4). In the fifth month, the field interviewer must first attempt a face-to-face interview before conducting a phone interview to re-establish rapport with the household.[5]

The rate of *No Response* to the 2012 vote question by interview mode is reported in Table 1. The numbers are unweighted to reveal the incidence of *No Response* for individual respondents. The CATI interviews conducted from the centralized call centers have the lowest *No Response* rate. This may be because call center interviewers can be closely monitored by supervisors listening in on interviews (Tarnai 2007). Phone interviews conducted at the discretion of field interviewers have the next lowest *No Response* rate of 8.5%, and in-person CAPI interviews had the highest *No Response* rate of 11.2%.

[5] The fifth month interview is thus also the point where replacement households may enter the sample (CPS 2006 Chpt. 7, p.4).

Mode	Percent No Response to Vote Question	Percent Response to Vote Question	Respondents (Number)	Respondents (Percent)
Unknown	92.0%	8.0%	399	0.4%
CATI (Phone)	3.8%	96.3%	10,638	11.3%
CAPI (Phone)	8.5%	91.5%	49,413	52.4%
CAPI (In-person)	11.2%	88.8%	33,861	35.9%
All	9.3%	90.7%	94,311	100.0%

Table 1. 2012 Vote Question (PES1) Non-Response by Interview Mode (Unweighted)

Household Size and Proxy Reporting

The Current Population Survey's sample frame is households, not individuals. When interviewers obtain responses from individuals in a household, one or more persons present at the interview may provide responses for themselves, as well as a proxy report for other household members not present. Proxy reports are known to affect reporting to the vote question. Highton (2005, 113), applying listwise deletion, finds that on average from 1992 to 2000 "proxy-reported turnout is about 4 percentage points lower than self-reported turnout." However, the potentially higher vote over-report bias of self-reports does not appear to confound "individual level correlates of turnout and interstate differences [, which] are mostly similar for these two measures."

The effect of proxy reporting on *No Response* cannot be directly observed since no persons with a response to the vote question, self or proxy reported, are coded as *No Response*. The potential for proxy reporting might be indirectly measured through household size since a proxy report is recorded when a household member provides responses for other household members. While more than one person may be present for an interview, household size might be an indicator of the necessity of providing proxy reports. Larger households have more persons for whom a proxy report might be provided. Attrition on supplement questions might occur when respondents (or the interviewer), having just completed the exhaustive main employment survey, find themselves faced with answering additional questions, with pressures to satisfice (Krosnick 1991) by skipping the supplement increasing with household size.

The rate of *No Response* to the 2012 vote question by household size is reported in Table 2. The size of the household is generated by counting the number of citizens of voting age in each household, using a household identifier provided in the CPS data. Household size appears to be related to *No Response*. Households with one or two citizens of voting age have a *No Response* rate of 8.7%. The failure to administer the vote question then increases with household size, reaching 19.8% for a household size of six. The No Response rate then mostly decreases as household size increases, excepting the largest *No Response* rate of 33.9% for the 56 persons with a household size of eight.

Household Size	No Response to Vote Question	Response to Vote Question	Respondents (Number)	Respondents (Percent)
1	8.7%	91.3%	18,775	19.9%
2	8.7%	91.3%	50,968	54.0%
3	9.4%	90.7%	15,762	16.7%
4	12.7%	87.3%	6,548	6.9%
5	16.0%	84.0%	1,610	1.7%
6	19.8%	80.2%	414	0.4%
7	12.1%	87.9%	140	0.1%
8	33.9%	66.1%	56	0.1%
9	5.6%	94.4%	18	0.0%
10	5.0%	95.0%	20	0.0%
All	9.3%	90.7%	94,311	100.0%

Table 2. 2012 Vote Question (PES1) Non-Response by Household Size (Unweighted)

What is evidently curious about these largest household sizes is that some household members must provide a response to the vote question since the size of the household is not evenly divided by the number of *No Responses*. Although proxy status cannot be observed for those who do not participate in the vote question, proxy status on the main employment questionnaire is recorded separately by interviewers. The proxy status of those who provide *No Response* to the vote question by proxy status to the main employment questionnaire are presented in Table 3. Those who provide a self report to the main employment questionnaire less often have a *No Response* to the vote question (7.9%) than those for whom a proxy report is provided (9.3%). The small number of persons are reported as having both a self and proxy report - presumably the respondent is present for only some of the lengthy employment questionnaire - have the lowest *No Response* rate to the vote question (6.3%), but if vote response has a random sampling component, this lower level is not statistically significant. It is the case that persons in the largest households who report for themselves on the main employment questionnaire are those who provide a response to the vote question, thus suggesting that satisficing behavior is present since the respondent or interviewer does not seek further responses for non-present household members, which could have been recorded as *Don't Know* if the questions were asked.

Caution is warranted in interpreting Table 2 since the main employment questionnaire may not be a perfect indicator of proxy status on the Voting and Registration Supplement. A simple cross-tabulation of the two proxy status indicators are presented in Table 4. Somehow during the interview process, 9.8% of respondents for whom an employment proxy report is given report their own voting status. On the other side of the coin, 2.2% of respondents who provided a self report to the employment questionnaire have a proxy report for their vote status. Combined, 5,344 of the 85,344 (6.3%) respondents who participated in the Voting and Registration Supplement have an inconsistent proxy status across the main employment and voting supplement.

Self or Proxy Report	No Response to Vote Question	Response to Vote Question	Respondents (Number)	Respondents (Percent)
Unknown	99.1%	0.9%	755	0.8%
Self Report	7.9%	92.1%	49,889	52.9%
Proxy Report	9.3%	90.7%	43,460	46.1%
Self and Proxy Report	6.3%	93.7%	207	0.2%
Overall	9.3%	90.7%	94,311	100.0%

Table 3. 2012 Vote Question (PES1) Non-Response by Proxy Status to Main Employment Questionnaire (Unweighted)

There may be innocuous reasons for this inconsistency since the survey is lengthy and respondent attrition is expected, as well respondent accretion (to coin a new term for a respondent with a proxy report at the start of a survey converting to a self report at a later stage). However, The inconsistencies raise suspicions that interviewers are not following protocols. If this were the case, we might expect respondents assigned to call centers for interviewing to have lower inconsistent proxy report status since they are being monitored while interviews are taking place. Indeed, they do, with only 363 of 10,217 (3.6%) respondents who participated in the supplement with an inconsistent reporting status, compared to 4.3% of CAPI phone interviews. (One wonders from the outside how this varied proxy reporting process works - do respondents pass a phone around or are these indicators of interviewer coding errors?) The CAPI in-person interviews are the most inconsistent, with 10.1%. More of the CAP in-person inconsistency is due to respondent accretion, which accounts for 86.7% of all CAPI in-person inconsistency, compared to 59.5% in the CATI phone interviews and 72.6% in the CAPI phone interviews.

	Vote Proxy Status				
Employment Proxy Status	**No Response**	**Self**	**Proxy**	**Respondents (Number)**	**Respondents (Percent)**
Unknown	99.2%	0.1%	0.7%	755	0.8%
Self	8.2%	89.6%	2.2%	49,889	52.9%
Proxy	9.5%	9.8%	80.7%	43,460	46.1%
Self and Proxy	7.3%	63.8%	29.0%	207	0.2%
All	9.5%	52.1%	38.4%	94,311	1.0%

Table 4. Proxy Status in the Main Employment Questionnaire and Vote Supplement

Month in Sample

The CPS is a panel study consisting of eight months of a household participating in the sample. Efforts are made to complete the first and (inaptly named) fifth months by CAPI in-person interviews. Since Table 1 demonstrates that CAPI in-person interviews have the highest *No Response* rate across modes, presumably the first and fifth months have the highest *No Response* rate. Table 5, which reports *No Response* by month-in-sample demonstrates they do not. The first month has the lowest *No Response* rate and the fifth month has the third highest.

Month-in-Sample	No Response to Vote Question	Response to Vote Question	Respondents (Number)	Respondents (Percent)
1	7.9%	92.1%	11,764	12.5%
2	8.8%	91.2%	11,987	12.7%
3	9.5%	90.5%	11,717	12.4%
4	10.4%	89.6%	11,872	12.6%
5	9.6%	90.4%	11,576	12.3%
6	9.1%	90.9%	11,809	12.5%
7	9.9%	90.1%	11,967	12.7%
8	9.1%	90.9%	11,619	12.3%
All	9.3%	90.7%	94,311	100.0%

Table 5. 2012 Vote Question (PES1) Non-Response by Month-in-Sample (Unweighted), All

Part of the explanation lies in that only 74.0% of the first and fifth month interviews are conducted CAPI in-person. However, as reported in Table 6, not only are most respondents interviewed CAPI in-person in these months, but the *No Response* rate is the lowest for the first and fifth month. For the CAPI telephone interviews reported in Table 7, the first and fifth months have the two highest *No Response* rates. As Table 8 reports, no CATI interviews are conducted in the first or fifth months and there is no readily evident pattern to *No Response* rates by month.

Month-in-Sample	Percent No Response to Vote Question	Percent Response to Vote Question	Respondents (Number)	Respondents (Percent)
1	7.2%	92.8%	9,844	29.1%
2	14.3%	85.7%	3,043	9.0%
3	14.6%	0.9%	2,710	8.0%
4	15.3%	84.7%	2,636	7.8%
5	9.2%	90.8%	7,424	21.9%
6	14.1%	85.9%	2,880	8.5%
7	14.6%	85.4%	2,877	8.5%
8	13.7%	86.3%	2,447	7.2%
All	11.2%	88.8%	33,861	100.0%

Table 6. 2012 Vote Question (PES1) Non-Response by Month-in-Sample (Unweighted), CAPI in-person only

Month-in-Sample	Percent No Response to Vote Question	Percent Response to Vote Question	Respondents (Number)	Respondents (Percent)
1	9.9%	90.1%	1,888	3.8%
2	7.4%	92.6%	7,232	14.6%
3	8.6%	91.4%	7,066	14.3%
4	9.3%	90.8%	7,017	14.2%
5	9.2%	90.8%	4,098	8.3%
6	7.9%	92.1%	7,465	15.1%
7	8.8%	91.3%	7,508	15.2%
8	8.2%	91.8%	7,139	14.4%
All	8.5%	91.5%	49,413	100.0%

Table 7. 2012 Vote Question (PES1) Non-Response by Month-in-Sample (Unweighted), CAPI phone only

Month-in-Sample	Percent No Response to Vote Question	Percent Response to Vote Question	Respondents (Number)	Respondents (Percent)
1			0	0.0%
2	3.4%	96.7%	1,671	15.7%
3	3.7%	96.3%	1,904	17.9%
4	4.5%	95.5%	2,130	20.0%
5			0	0.0%
6	3.3%	96.7%	1,430	13.4%
7	3.9%	96.1%	1,533	14.4%
8	3.6%	96.4%	1,970	18.5%
All	3.8%	96.3%	10,638	100.0%

Table 8. 2012 Vote Question (PES1) Non-Response by Month-in-Sample (Unweighted), CATI phone only

Presence of Phone in Household

These proceeding statistics suggest that a respondents' willingness to be interviewed by phone is related to their participation in the Voting and Registration Supplement. Field interviewers ask respondents if a phone is present in the household. Only 3.0% of respondents report not having a phone in their household (there is double counting here as this is not a household measure). Among these, 23.6% had *No Response* compared with 8.8% for those respondents in a household with a phone.

Other Demographics

Household income may be relevant, as some respondents may not be able to afford a phone in their household. Indeed, income is related to *No Response* rates, with the lowest two aggregate household income categories have the two lowest *No Response* rates, 13.2% for respondents with a household income less than $5,000 and 11.1% for respondents with a household income from $5,000 to $7,499. There is no discernible pattern for higher income levels, ranging between 8.3% for household income between $100,000 to $149,999 to 10.2% for household income between $35,000 to $39,999. However, caution is warranted since household income has a high degree of missing values, with 18.7% of respondents missing income.

Similarly, persons with at least some college education are less likely to have a No Response (8.6%) than those without (10.3%). Persons age 18-29, who might be poorer and more often unable to afford a phone, do have the highest *No Response* rate of 10.7%, but this is again not much higher than the 8.2% reported for persons age 60 and older.

Predicting *No Response*

Key demographic variables thus appear only modestly correlated with failure to administrate the Voting and Registration supplement. The missing data mechanism does not appear to be MAR for these variables. As Little and Rubin (1987) caution, it is thus unlikely that inclusion of similar demographic measures in multivariate models results in unbiased estimates. A model of the missing data mechanism is needed.

The dependent variable is an indicator variable coded as '1' if a respondent provided an answer to the vote question and '0' if the respondent had *No Response*. A logit model is employed to predict if a respondent provides a response.

The independent variables include the interview mode: whether or not an interview was conducted by *CATI*, which originate from the centralized call centers, or field interviewers conduct a *CAPI* phone interview. The excluded interview model is where field interviews conduct an CAPI in-person interview.[6] Indicator variables identify if the interview was conducted in the 1st Month in Sample and 5th Month in Sample, with all other six months the excluded categories. Respondents with a *Phone in Household* are identified. The *Size of Household* for the respondent varies from one to ten citizens of voting age. Two further demographic controls are included: if the respondent is *Non-Hispanic White* , has at least *Some College or More* education, and the *Age* of the respondent, incrementally coded into four ascending categories (18-29, 30-44, 45-59, and 60+).[7]

[6] The omitted category of interview mode also includes a small number (0.4% of all interviews) with an unknown interview mode.

[7] Some variables that were tested but not incorporated into the model include household income, which become insignificant in explanatory power when education is entered into the equation. Gender was also tested with no correlation found.

Variable	Coefficient (Standard Error)	*p*
CATI (Phone)	1.345	0.00
	(.056)	
CAPI (Phone)	0.458	0.00
	(.026)	
1st Month in Sample	0.549	0.00
	(.039)	
5th Month in Sample	0.264	0.00
	(.036)	
Phone in Household	0.918	0.00
	(.047)	
Household Size	-0.138	0.00
	(.011)	
Non-Hispanic White	0.133	0.00
	(.026)	
Age	0.038	0.00
	(.011)	
Some College or More	0.125	0.00
	(.023)	
Constant	1.050	0.00
	(.057)	
Number of Observations	94,311	
Log Likelihood	-28,250.845	

Table 9. Modeling Predicting Response to Vote Question

All the independent variables are highly statistically significant in the expected directions from the preceding section, often with coefficients estimated ten or more standard errors from zero. The *CATI* phone interviews more often result in a response to the vote question relative to the CAPI in-person interviews, as do the *CAPI* phone interviews. The first and fifth month-in-sample interviews more often result in a response. The higher response rates in these months do not simply appear to be a function of persons who might have been interviewed by phone in another month, who happen to have higher socio-economic status. Demographic controls such as presence of a *Phone in Household* and respondents' race, age, and education levels should wipe out the month-in-sample effects if this was simply caused by a greater pool of people who are predisposed to not participate in the vote question.

This is a first cut exploratory analysis; further investigation into other demographic controls will follow to confirm. Also needed are corrections for other missing data categories of *Refused* and *Don't Know*. Still, the relationships found in the model, especially among paradata variables like interview mode and month-in-sample, strongly suggest that these missing data mechanisms should be incorporated into corrections for the missing data.

Applying Proposed Corrections

Four candidate corrections to the Current Population Survey Voting and Registration Supplement emerge.

Method One: The Census Method

The Census Method assumes that CPS supplement missing data are equal to zero for turnout and registration statistics.

Method Two: The Listwise Deletion Method

The Listwise Deletion Method favored by the scholarly community applies listwise deletion to records with missing data for turnout and registration statistics. A drawback of listwise deletion is that it restricts the sample to fewer observations, thereby increasing statistical measures of uncertainty. Listwise deletion should not affect greatly most CPS voting and registration statistics since the CPS has a large sample size, but it may affect small subsamples.

Method Three: The Hur and Achen Method

Hur and Achen (2013, 988) note that after applying listwise deletion, the CPS citizen voting-age population turnout rate is larger than turnout rates calculated from administrative records (see also Figure 2, above). Furthermore, the over-report bias is systematically higher and lower for some states. The average, minimum and maximum CPS CVAP turnout rate in the 2000 to 2012 presidential elections, calculated using listwise deletion for missing data, is plotted for all states plus DC in Figure 3. States are ordered on their average over-report bias for clarity.

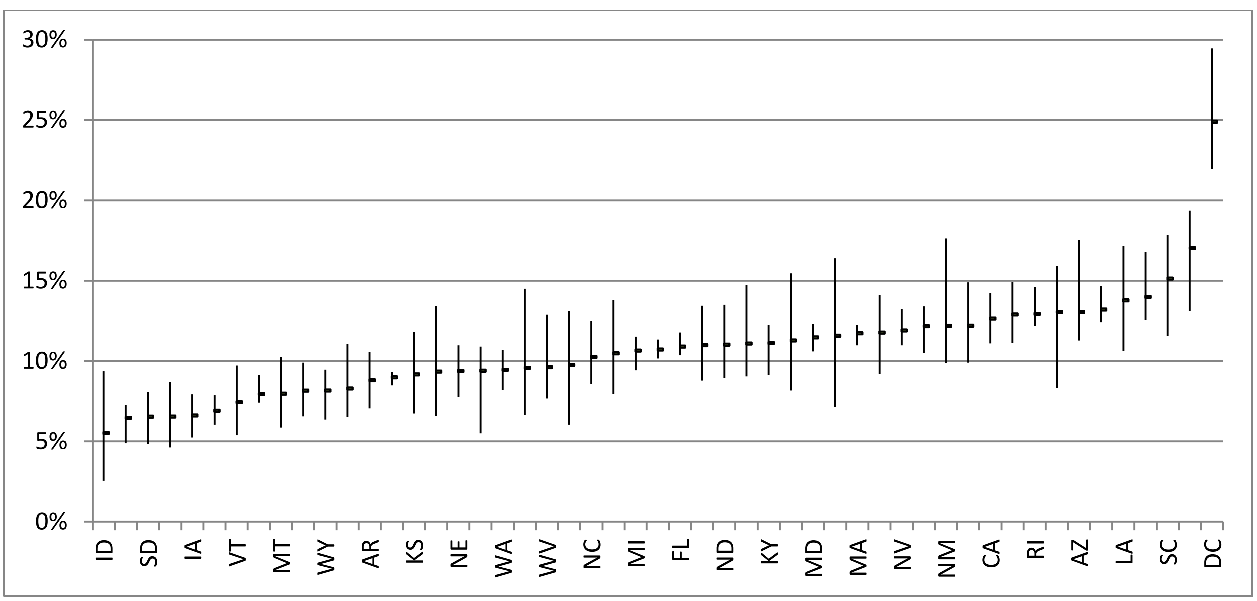

Figure 3. Average, Minimum, and Maximum CPS CVAP Over-Report Bias, 2000 to 2012 Presidential Elections (Listwise Deletion Applied)

The District of Columbia, Mississippi, South Carolina, New York, Louisiana, Alabama, Arizona, and Georgia all averaged thirteen percentage points or more difference between the CPS CVAP turnout rate (with listwise deletion applied) and the VEP turnout rate. The District of Columbia particularly stands out, with an average over-report bias of nearly twenty-five percentage points. Conversely, states like Idaho, Minnesota, South Dakota, Maine, Iowa, and New Hampshire average over-report bias of less than seven percentage points. These latter states have higher turnout rates, so perhaps they are limited in how much over-report bias they can exhibit. However, these patterns persist when computed as percent deviation from baseline turnout rates.

To account for varied over-report bias across states, which then affects estimates for the nation as a whole, Hur and Achen (2013, 991) recommend reweighting the CPS by state level VEP turnout rates. They recommend adjusting the CPS individual level weight, PWSSWGHT, in the following manner. (1) Compute the CVAP listwise deleted turnout rate by state. (2) Divide the percent voting by the VEP turnout rate, and multiply this by PWSSWGHT for respondent who report voting. (3) Divide the percent not voting by one minus the VEP turnout rate, and multiply this by PWSSWGHT for respondent who report not voting.[8] This procedure produces CPS voter turnout rates that closely match the VEP turnout rates.

Method Four: Multiple Imputation Method

The Scholarly Method and the Hur and Achen improvement upon it assume that the missing data are MAR. The Multiple Imputation Method models the missing data mechanism using the model presented in Table 9, creates multiple datasets with randomly imputed missing values using the model, and then produces estimates from averages of statistics calculated from statistical procedures applied to generate each imputed dataset.

Regrettably, for now, I will not proceed with the Multiple Imputation Method since more work is needed on modeling the missing data mechanism.

Comparing Estimates from the Missing Data Correction Methods

These four missing data correction methods theoretically affect substantive estimates. However, if the missing data are sufficiently small relative to the overall data or the correction method is orthogonal to estimation, substantive results will be robust to applied correction method.

To explore the effect of correction method on substantive estimates, I apply the correction methods to the voter turnout estimates in two contexts. First, I compute CVAP turnout rates for

[8] Although Hur and Achen imply this correction should be applied using the VEP total ballots counted turnout rate, CPS consumers will have to estimate the total ballots counted in states that do not report this statistic. I choose, for now, to apply the correction using the VEP vote for highest office turnout rate, or VEP turnout rate for president, since this rate does not require estimating total ballots counted for some states.

demographic subgroups, such as race and ethnicity and age. Second, I replicate a recent *AJPS* article by Burden, et al. (2013) that explores the correlations between various election laws and turnout.

Demographics

CPS CVAP turnout rates, calculated employing the three of the correction methods, for the demographic characteristics of race and ethnicity, sex, age, and education are presented in Table 10.

Demographic		Census	Scholarly	Hur & Achen
Race/Ethnicity	*White*	66.1%	75.5%	65.2%
	Black	65.2%	79.4%	69.1%
	Hispanic	49.9%	59.5%	46.5%
Sex	*Male*	61.5%	71.4%	60.1%
	Female	65.7%	76.0%	65.7%
Age	*18-29*	51.1%	60.9%	48.4%
	30-44	61.8%	71.9%	60.7%
	45-59	68.5%	79.1%	69.5%
	60+	70.8%	80.3%	71.0%
Education	*Less than H.S.*	39.4%	46.3%	34.1%
	High School	54.9%	65.1%	53.0%
	Some College	68.0%	78.1%	68.3%
	Bachelor's Degree +	78.9%	89.7%	83.8%
Income	*<$25K*	54.1%	59.3%	46.9%
	$25-$50K	63.3%	68.9%	57.3%
	$50-$100K	72.8%	78.9%	69.3%
	$100K+	61.4%	80.3%	70.8%

Table 10. Selected 2008 CPS CVAP Turnout Rates by Correction Method

Among the widely reported statistics from the 2012 Voting and Registration Supplement Report (CPS 2013, 3) is that Blacks voted at a higher rate than non-Hispanic Whites for the first time "first time since the Census Bureau started publishing voting rates by the eligible citizenship population in 1996." The Census Bureau reported in 2008 that the turnout rate for Non-Hispanic Whites was 66.1% to 65.2% for Blacks. However, applying listwise deletion to the 2008 data, Blacks first voted at higher rates than non-Hispanic Whites in 2008, 79.4% to 75.5%.[9] Further applying Hur and Achen's weighting correction, Blacks voted at higher rates than non-Hispanic Whites, 69.1% to 65.2%. The primary reason for this substantive changes is that 14.3% of Blacks had *No Response*, compared to 9.3% for non-Hispanic Whites (both have a similar percentage of *No Response* and *Don't Know* of a little more than 2%). The Census Bureau claim about voting patterns by race thus hinges on their choice to count respondents who were never administered the voting supplement as having not voted.

Chief Justice Roberts in the 2013 *Shelby County v. Holder* decision overturning a key provision of the Voting Rights Act cited Census Bureau reports that African-Americans' 2008 turnout rates in Mississippi were higher than in Massachusetts to argue that the Section 4 coverage formula may no longer accurately define where voting discrimination persists. His representation of the Census Bureau data is correct; according to the Census Bureau 72.9% of Mississippi Blacks voted compared to 65.8% in Massachusetts (weighted by PWSSWGHT). However, applying listwise deletion the differences narrow to a 81.4% turnout rate in Mississippi compared to 81.0% in Massachusetts. Recognizing that the over-report bias in Mississippi is among the highest in the country (Figure 3), further applying Hur and Achen's correction indicates that Massachusetts African Americans have a higher turnout rate than those in Mississippi, 69.5% to 65.2%. Chief Justice Robert's justification for overturning the Section 4 coverage formula thus rests partially on shaky methodological assumptions that he is blithely unaware of.

Among the demographic characteristics investigated, the turnout rates by race and ethnicity stand out in that relative rankings are altered. For other categories such as age and education, turnout rate differences between the highest and lowest categories tend to increase from the Census Bureau's calculations when applying listwise deletion and further applying Hur and Achen's correction. For age, the Census Bureau calculates a 19.7 percentage point difference between the turnout rates of persons age 60+ and age 18-29. Applying listwise deletion, the difference is 19.4 points, and further applying Hur an Achen's weighting correction the difference is 22.6 percentage points, and the estimated turnout rate is now below fifty percent for youth. Similar patterns are observed with education; the Census Bureau calculates a 39.5 percentage point difference between persons with at least a Bachelor's degree and those with less than a high school education; applying listwise deletion the difference is 43.4 percentage points, and further applying Hur and Achen's weighting correction the difference is 49.8 percentage points

The corrected turnout rate calculations reveal that on age and education demographics, participation inequalities may be larger than previously recognized. The reason for the increased differential turnout rates is that even though the lowest category of age and education has a higher *No Response* rate than the highest category (11.3% for age 18-29 vs. 9.0% for age 60+, 11.0% for

[9] See also, Michael P. McDonald. "2012 Turnout: Race, Ethnicity, and the Youth Vote." May 5, 2013. Available at: http://www.huffingtonpost.com/michael-p-mcdonald/2012-turnout-race-ethnict_b_3240179.html, accessed March 29, 2014.

persons less than a high school diploma vs. 9.2% for persons with at least a Bachelors' degree), the turnout rate for the highest category is much higher than the lowest category, therefore the correction has a greater effect on the highest category.

On household income there may be less participation inequality. The Census Bureau calculates a 25.3 percentage point difference between individuals in households making more than $100,000 and those living in households making less than $25,000; applying listwise deletion the difference is 21.0 percentage points, and is 23.9 percentage points further applying Hur and Achen's weighting correction. However, these patterns may be an artifact of reporting issues since 18.7% of respondents are missing income data. Whereas 10.2% of all respondents have *No Response* to the vote question, only 5.8% of those without missing family income have *No Response*. As a consequence, among those with household income data the percentages of *No Response* across family income levels are relatively smaller and are not much different (6.3% for persons in households with less than $25K vs. 5.8% in households with more than $100K).[10]

The Unintended Consequences of Election Laws

In a provocative 2013 *AJPS* article Burden, Canon, Mayer and Maynihan (2013) argue that election laws may have unintended consequences, most notably that early voting reduces voter turnout. Among the key results are an analysis of the 2008 and 2004 CPS turnout rates (Burden, et al. 2013, 101). These scholars' dependent variable is constructed in a similar manner as the Census Bureau, counting the *No Response*, *Don't Know*, and *Refused* response items as *Did Not Vote* (Burden, et al. 2013, 101, footnote 11). The article thus makes a good candidate for replication to determine how modeling missing data may affect statistical estimates.[11] If the missing data are MAR and a good mix of covariates are entered into the model, then it may be that missing data will not affect the published estimates.

Results from the replications are presented in Table 11, focusing on the key substantive variables of interest and excluding numerous control variables. Model 1 presents the faithful replication, which defines the vote variable in the same way as the Census Bureau. Model 2 presents the results applying listwise deletion to the missing data. Model 3 presents the results applying listwise deletion and weighting using the Hur and Achen weight correction. Model 4 presents results using the same procedures in Model 3 and excluding the family income variable, which as noted above has a high degree of missingness, and adding DC, which was inadvertently dropped from the original model. All *p*-values are computed for a two-tailed test.

The coefficients and their statistical significance are relatively stable when following the Census Bureau's practice in Model 1 or using listwise deletion in Model 2. Election Day Registration's statistical significance is on a knife's edge and drops below conventional levels of

[10] This raises an interesting observation that missing data may be systematic across interviews.
[11] The authors generously shared the data and Stata code, from which I was able to faithfully replicate their 2008 logit model from canonical CPS data downloaded from the Census Bureau's Data Ferrett service.

statistical significance in Model 2, but one might reasonably apply a one-tailed test since theory predicts a positive correlation. The coefficient for states with Early Voting, Election Day Registration, and Same Day Registration changes sign across the two models, but the coefficient is practically not different from zero in either model, so this sign change appears due to small random perturbations of the data.

	Model 1: Original Replication		**Model 2: Listwise Deletion**		**Model 3: Hur & Achen**		**Model 4: Hur & Achen + add DC, remove income**	
	Coef (S.E.)	***p***	**Coef (S.E.)**	***p***	**Coef (S.E.)**	***p***	**Coef (S.E.)**	***p***
EDR	0.191	0.04	0.195	0.10	0.312	0.06	0.332	0.02
	(.092)		(.12)		(.167)		(.141)	
Early Voting	-0.180	0.00	-0.179	0.01	-0.152	0.00	-0.100	0.10
	(.058)		(.068)		(.053)		(.06)	
Early Voting + EDR	-0.069	0.59	-0.051	0.83	0.141	0.10	0.190	0.06
	(.128)		(.24)		(.087)		(.102)	
Early Voting + SDR	0.008	0.87	-0.007	0.91	-0.014	0.75	0.028	0.55
	(.048)		(.064)		(.042)		(.046)	
Early Voting + EDR + SDR	0.134	0.10	0.132	0.08	0.310	0.00	0.293	0.00
	(.082)		(.075)		(.064)		(.068)	
Number of Observations	73,333		67,780		67,780		80,667	

Table 11. Replication of Burden, et al. (2013), With Different Missing Data Treatments

Difference emerge when Hur and Achen's weighting correction is applied. There are some features of the model that can be interpreted more cleanly. Theoretically, there is not a strong expectation why Election Day Registration and Early Voting should interact with one another, but they appear to do so in Model 1 since adding the EDR coefficient and the Early Vote coefficients yields a value of 0.011, while the estimated coefficient for Early Vote + EDR equals -0.069. In Model 3, adding the two coefficients together yields 0.160, while the estimated coefficient for Early Vote + EDR equals a more similar 0.141. Also note that the coefficient for EDR has increased from 0.191 in Model 1 to 0.312 in Model 2. Thus, if EDR has a stronger effect than predicted in the original model, it may logically follow that in Model 1 Early Voting and EDR has a non-statistically significant negative coefficient ($p < .59$) whereas in Model 3 it has a positive coefficient near statistical significance ($p < 0.10$).

The most controversial finding in the original research, that early voting has a depressing effect on turnout, still stands in Model 3. In the course of replication, I discovered that DC had inadvertently been removed from the analysis, due to the fact that a state level competitiveness variable was missing data for DC.[12] To reincorporate DC in the analysis, I set the value for DC to the least competitive state, Wyoming.[13] Furthermore, I noticed the model included household income, which noted above, has a high degree of missingness (McDonald and Thornburg 2012). In Model 4, the Model 3 analysis is conducted, with further adding DC back into the model and excluding household income. The results of Model 4 appear similar to Model 3, except that the coefficient on Early Voting is negative but no longer statistically significant at conventional levels (p <.10). Still, if one believes early voting increases turnout, their model does not support it (although others do so).

Conclusion

The Census Bureau's practice of scoring missing data on the CPS Voting and Registration Supplement, primarily composed of respondents who were never administered the vote question, as *Did Not Vote* affects perceptions of the magnitude of vote over-report bias, an indicator used to gauge the reliability of the CPS. Furthermore, because the percentage of missing data is increasing over time, the Census Bureau's practice masks trends in vote over-report bias. Whereas the Census Bureau reports decreasing vote over-report bias, applying listwise deletion to the missing data indicates increasing vote over-report bias.

Modeling assumptions about the missing data mechanism produce substantively different results that affect perceptions of the state of participation in the United States. The Census Bureau reported that African-American turnout exceeded non-Hispanic White turnout for the first time in 2012, when applying listwise deletion to the missing data, because more African-Americans are not administered the voting supplement, African-American turnout exceeded non-Hispanic White turnout in 2008. Vote over-report bias varies among states, too, in a persistent manner. When a further weighting correction recommended by Hur and Achen (2013) is applied, revised turnout rates change perceptions about the continued need for special provisions of the Voting Rights Act struck down recently by the Supreme Court. These corrections further affect estimates of published research, such as a recent controversial AJPS article by Burden, et al. (2013) that finds early voting depresses turnout.

The corrections applied in this paper assume the data are missing at random. However, the CPS Voting and Registration Supplement data are not missing at random. The missing data mechanism is related to administration of the Current Population Survey. In future research, I intend to model the missing data correction and employ multiple imputation methods to turnout estimates and models.

[12] I confirmed the inadvertent omission DC with the authors by personal communication.

[13] DC was dropped because it was given a missing value for state competitiveness, which was constructed from the final Pollster poll averages prior to the election. DC had no polling, so it had no value. I set DC's value equal to the least competitive state, Wyoming.

References

Burden, Barry C., David T. Canon, Kenneth R. Meyer, and Donald P. Moynihan. 2013. "Election Laws, Mobilization, and Turnout: The Unanticipated Consequences of Election Reform." *American Journal of Political Science* 58(1): 95-109.

Current Population Survey, Technical Paper 66. 2006. "Design and Methodology: Current Population Survey." Washington, DC: U.S. Department of Commerce, U.S Census Bureau.

Current Population Survey. 2013. "The Diversifying Electorate - Voting Rates by Hispanic Origin in 2012 (and Other Recent Elections)." Washington, DC: U.S. Department of Commerce, U.S Census Bureau.

Highton, Benjamin. 2005. "Self-Reported Versus Proxy-Reported Voter Turnout in the Current Population Survey." *Public Opinion Quarterly* 69(1): 113-23.

Hur, Aram and Christopher Achen. 2013. "Coding Voter Turnout Responses in the Current Population Survey." *Public Opinion Quarterly* 77(4): 985–93.

Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5: 213-36.

Little, Roderick J.A. and Donald B. Rubin. 1987. *Statistical Analysis with Missing Data*. New York, New York: Wiley and Sons.

Michael P. McDonald. 2005. "Reporting Bias" in *Polling in America: An Encyclopedia of Public Opinion*. Benjamin Radcliff and Samuel Best, eds. Westport, CT: Greenwood Press.

Michael P. McDonald. 2011. "Voter Turnout: Eligibility Has Its Benefits" in *Controversies in Voting Behavior*, 2nd Edition, Richard G. Niemi, Herbert F. Weisberg, and David Kimball, eds. Washington, DC: CQ Press

Michael P. McDonald and Samuel Popkin. 2001. "The Myth of the Vanishing Voter." *American Political Science Review* 95(4): 963-974.

Michael P. McDonald and Matthew Thornburg. 2012. "Evidence for Mode of Interview Effects: The Case of Supplementing Exit Polls with Early Voter Phone Surveys." *Public Opinion Quarterly* 76(2): 326-49.

Tarnai, John. 2007. "Monitoring CATI Interviewers." Paper prepared for the 2007 American Association of Public Opinion Researchers Conference.

Wolfinger and Rosenstone. 1980. *Who Votes?* New Haven: Yale University Press.