Analysis of Nonresponse Bias in the 2010 Post-Election Voting Survey of Uniformed Service Members

Timothy Markham[1], David McGrath[1]

[1]Defense Manpower Data Center, 4800 Mark Center Drive, Suite 04E25-1, Alexandria, VA 22350

Abstract

In an era of declining response rates, survey organizations are increasingly examining the effects of nonresponse bias on their survey estimates. As fewer people respond to surveys, the nonresponse bias increases if survey respondents and non-respondents are systematically different from each other, even after the application of sophisticated weighting methods to account for survey nonresponse. To assess the prevalence of this bias in estimates from the Federal Voting Assistance Program's *2010 Post-Election Voting Survey of Uniformed Service Members*, the Department of Defense Manpower Data Center (DMDC) performed a non-response bias study by surveying a random sub-sample of survey non-respondents. Because the nonresponse sample members were interviewed by telephone and the production survey was a self-administered, Web instrument, DMDC also selected a sub-sample of the production sample for phone interviews to test for mode effects. Results show that modest mode effects exist for some variables, and DMDC uses this result while interpreting the nonresponse study. Analyses of the nonresponse survey reveal that unweighted estimates show large nonresponse bias, but DMDC's weighting methods effectively reduce, but may not eliminate, this bias in production estimates.

From this study, DMDC compared unweighted data from the production and nonresponse surveys with weighted production data; created estimates with the same weighting process as if the survey had closed two weeks earlier or if respondents to the nonresponse study had responded to the production survey to determine the effect of late responders; and created composite estimates comprised of the production data and non-respondent data.

Key Words: Nonresponse bias, voting, mode effects, military

1. Introduction

In accordance with the Government Accountability Office's recommendation, DMDC performed a nonresponse bias study of the *2010 Post-Election Voting Survey of Uniformed Service Members*. To gain participation from non-respondents to the production survey, DMDC changed the mode of survey administration from self-administration via the Web to an interviewer-administered telephone survey. In order to account for potential mode effects, DMDC also performed a mode study by sampling a subset of the original production sample for phone interview to isolate the effect of the telephone mode on the survey estimates.

DMDC assessed the nonresponse bias in the *2010 Post-Election Voting Survey of Uniformed Service Members* using many methods, including sample composition, response rates across subgroups, composite estimates that combine production and nonresponse bias survey data, comparing early and late responders to the production survey (continuum of resistance model), and computing various sets of weights for the nonresponse bias survey data.

DMDC concludes that the sophisticated weighting greatly reduces the nonresponse bias present in unweighted estimates. For questions on both voting rates and voter registration rates, Tables 9 and 10 show that the unweighted estimates from the production survey greatly exceed the unweighted estimates from the nonresponse bias study, but the weighting appropriately adjusts the production estimates toward the data from the survey non-respondents. Because of the large variance of the results from the mode and nonresponse bias studies, most comparisons are not statistically significant, but some general patterns in the data emerge. First, there is some evidence that phone respondents over report 'positive' answers (voting, use of FVAP Websites, etc.), and therefore caution needs to be used when assessing the nonresponse bias study. The nonresponse bias study shows that DMDC's weighting methods reduce the nonresponse bias for both the voting and registration rates, possibly slightly under-adjusting for the voting rate and over-adjusting for the registration rate. Although nonresponse bias likely remains the largest source of error in *the 2010 Post-Election Voting Survey of Uniformed Service Members*, our judgment is that the magnitude is only modest.

2. Nonresponse Bias Considerations

With the continuing decline of response rates, survey organizations have been looking to better understand, quantify, and account for nonresponse bias (NRB). For instance, although low response rates may indicate the presence of NRB, especially when they vary across subgroups, they do not guarantee NRB. As Groves points out, improving response rates may not solve NRB issues, especially because the amount of NRB can be different across items in the same survey (p. 649).

One common method of assessing NRB is to assume a continuum of responses. Respondents who complete the survey late in the field period may be more like non-respondents than respondents that complete the survey when it first opens. Based on this assumption, a regression can be run to determine the predicted values of non-respondents (Armstrong and Overton, p. 399-400). Armstrong and Overton also note that one measure of nonresponse bias can be a comparison with known population estimates. In the case of the 2010 Post-Election Voting Survey of Uniformed Service Members, extensive administrative data is available to characterize the population, make comparisons with survey respondents, and create weights to account for differences.

Fowler, et al., argue that contacting non-respondents through telephone interviews can increase response rates and decrease bias in estimates (p. 191). This procedure was tested for the 2010 Post-Election Voting Survey of Uniformed Service Members, where non-respondents were called and the survey was administered over the phone. While this data was not used in the creation of the production estimates, comparing non-respondent estimates with production estimates allows for analysis on the presence of NRB, the effectiveness of weighting procedures, and the existence of a continuum of responses as described above.

3. Design

3.1 Production Survey

The production survey for the *2010 Post-Election Voting Survey of Uniformed Service Members* had a sample of 76,244 uniformed service members, comprised of both active duty and 'activated' Reserve Component members.[1] In order to ensure that enough members were selected in each domain of interest, the sample was stratified. The stratifying variables consisted of region,[2] service or Reserve component, paygrade, age,[3] gender, and race.[4] Each member received e-mail and postal notifications to complete the Web survey, which fielded from January 10 to February 28, 2011.

3.2 Mode Study

To assess the validity of comparisons between the data collected from the (NRB) study and the production study, DMDC needed to test for presence of mode effects. To determine if responding to the survey by telephone as opposed to the Web had an effect on survey responses, DMDC selected a subsample of 3,000 members from the 76,244 in the production sample. The subsample was composed so that each of the original 295 strata was expected to have at least 2 complete and eligible respondents based on expected response rates (estimated from responses to the *2008 Post-Election Voting Survey of Uniformed Service Members*). The numbers were adjusted so that no stratum had more members in the mode study than in the production survey. Each member was then called and an abbreviated form of the survey was administered over the phone. The mode study was planned to field at the same time as the production survey, but delays caused the field period to slip to April 12 to May 12, 2011. Due to the difference in survey administration periods, 181 members of the mode sample (6%) that had been eligible at the fielding of the production survey became record ineligible, due to factors such as separation from the military.

3.3 Nonresponse Bias Study

First, DMDC needed to identify who was eligible for the NBR study. Of the non-respondents to the production survey, those who were not ineligible when the main survey's field period began and did not refuse the production survey were considered eligible non-respondents. One thousand respondents to the nonresponse bias study were desired to study the possible existence of nonresponse bias (NRB) in the *2010 Post-Election Voting Survey of Uniformed Service Members*. Assuming a response rate of approximately 25 percent to the nonresponse bias study, DMDC sampled 3,999[5] eligible non-respondents. As in the mode study, each member in the sample was then called and the abbreviated survey was administered over the phone. As the nonresponse bias study was fielded from June 2 through June 22, 2011, 169 members (4%) who had been eligible non-respondents to the production survey became record ineligible for the NRB study. Although the NRB study fielded later, and hence DMDC would expect more members to become ineligible, the percent ineligible is lower in the NRB study than the

---

[1] DMDC identified 'activated Reservists' who were likely to be covered by UOCAVA law by selecting AGR (permanently activated guard and Reserve) and Reserve members activated on DMDC's CTS activation file.

[2] The region variable was collapsed into two categories: US and US territories, and overseas.

[3] The age variable was collapsed into 18-24 years old and greater than 24 years old.

[4] The race variable was collapsed into non-minority and minority.

[5] With the expected response rate of 25%, the sample size would have been 4,000. However, due to rounding, the actual sample taken included 3,999 members.

mode study because only eligible non-respondents from the production survey were sampled for the NRB survey, creating a smaller pool of possible participants.

Table 1: Sample Size and Percent by Disposition Code for the Three Surveys

Disposition Code	Production (Web)		Mode (Phone)		NRB (Phone)	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
1 Record ineligible based on sample file[a]	1,432	2%	181	6%	169	4%
2 Ineligible--Self or proxy report[b]	30	0%	11	0%	2	0%
3 Ineligible--Survey self report[c]	1,820	2%	0	0%	50	1%
4 Complete eligible response	8,546	12%	553	18%	472	12%
5 Incomplete eligible response[d]	647	1%	16	1%	12	0%
8 Refused/other	663	1%	287	10%	276	7%
9 Blank	731	1%	0	0%	0	0%
10 Postal non-deliverable (PND)[e]	9,197	13%	116	4%	354	9%
11 Non-respondents	50,178	69%	1,836	61%	2664	67%
Total	**73,244**	**100%**	**3,000**	**100%**	**3,999**	**100%**

[a]In the month preceding survey fielding, member eligibility for the sample is checked against administrative files. Members that have left the service, been promoted above paygrade 06, etc. are rendered record ineligible.

[b]The member or someone on behalf of the member can contact the data collection agency to declare that the member is ineligible for the survey.

[c]There were several demographic questions on the production survey that were removed from the mode and nonresponse bias surveys for brevity. If a member's responses indicated that they were ineligible, such as he or she had left the service at the time the survey was completed, that member was coded as a self-report ineligible case.

[d]If a member responded to less than 50% of the survey questions, he or she was coded as an incomplete respondent.

[e]For the telephone surveys, members with non-working or incorrect telephone numbers were assigned a disposition code of "postal non-deliverable."

Members were more likely, examining unweighted data, to become complete eligible respondents in the mode study (18%) than either the production or the nonresponse bias survey (both 12%). Because non-respondents are different from respondents, DMDC can only compare the survey mode and production study, where this study indicates that the military is more likely to respond to a telephone survey than a Web survey. Refusals occurred far more frequently in both the mode and nonresponse bias studies than in the production survey. This conforms to expectation because phone respondents are forced to make a 'hard refusal' if they don't want to take a survey after answering the phone unaware that it is a survey request. In contrast, respondents to the production survey don't need to make a hard refusal because they can avoid taking the survey by simply ignoring the e-mail and postal contacts. Most of the non-respondents (disposition code '11' in Table 1) to the production survey are likely 'soft refusals,' but DMDC cannot confirm this.

The response rates for each of the three surveys, shown in Table 2, have been calculated in accordance with the American Association for Public Opinion Research (AAPOR) RR3 recommendations. All response rates are weighted using base weights.

Table 2: Location, Completion, and Response Rates for the Three Surveys

Rate	Production	Mode	Nonresponse Bias
Adjusted Location Rate[a]	85%	97%	92%
Adjusted Completion Rate	18%	21%	18%
Adjusted Response Rate	15%	21%	16%

[a]Location Rate is based on PNDs, which are determined from mailing address for the production survey and telephone numbers for the other two surveys.

4. Sample Composition

4.1 Production Survey

In order to make valid comparisons of unweighted estimates from these three different surveys, DMDC needs to analyze the sample composition to make sure that differences in estimates are attributable to mode effects (when comparing the production survey with the mode study) and nonresponse bias when comparing the production survey (or mode study) with the nonresponse bias sample. For example, if the unweighted voting rates differ between the phone mode study and the non-response bias study, DMDC needs to confirm that the reason for the difference was not because a large percentage of the sample respondents from one of the two studies came from a group typically less likely to vote. The goal of the mode sample was to create estimates with low overall variance to increase the power on comparisons between the production and mode studies. However, the primary goal of the production survey was to make a large number of domain estimates with sufficient precision. Therefore, differences in unweighted estimates can be attributable to both 'mode effects' on survey measurement and differences in sample composition. Table 3a and 3b show the population distribution, sample distribution, and respondent distribution by region, a key demographic. These data depict the total number and percent of these groups that fall into each category. Due to space limitations, the table splits across an 'a' and 'b' version.

Table 3a: Distribution of Population and Production Survey Sample and Respondents, by Region[6]

Category	Population		Production Sample		Production Respondents	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
US & US territories	1,192,729	76%	35,644	49%	4,641	54%
Overseas	372,812	24%	37,600	51%	3,905	46%
Total	**1,565,541**	**100%**	**73,244**	**100%**	**8,546**	**100%**

---

[6] The region variable is created based on a Service member's Unit Identification Code (UIC). In some cases, a member may be overseas but their UIC is still located in the United States. In this instance, the member would be categorized as US & US territories, although he or she would be physically overseas.

Table 3b: Distribution of Mode and Nonresponse Bias Samples and Respondents, by Region[7]

Category	Mode Sample Frequency	Mode Sample Percent	Mode Respondents Frequency	Mode Respondents Percent	NRB Sample Frequency	NRB Sample Percent	NRB Respondents Frequency	NRB Respondents Percent
US & US territories	2,224	74%	447	81%	1,919	48%	274	55%
Overseas	776	26%	106	19%	2,080	52%	222	45%
Total	**3,000**	**100%**	**553**	**100%**	**3,999**	**100%**	**496**	**100%**

The composition of the production sample is disproportionately 'overseas' relative to the mode study (51% vs. 26%). To meet the required estimation domains, DMDC sampled overseas members at a high rate for the production sample. The NRB sample shared the skewed composition because the NRB sample was a sub-sample of the production sample. These kind of demographic differences in sample composition are prevalent in several key variables, including Service, paygrade, and race, and disallow (make invalid) the testing for mode effects between the production and mode surveys using comparisons of unweighted estimates.

5. Subgroup Response Rates

Tables 3a and 3b showed population numbers, sample sizes, and number of respondents for each of the three studies. Differences in percentages between the sample and the respondents are driven by response propensities, so the discussion of response rates in this section is closely related to the discussion on sample composition. In other words, subgroups that have higher response propensities will comprise a larger proportion of the respondents than of the sample. For example, Army is 42% of the frame population, but 50% of the sample for the production survey and 52% of the nonresponse bias study sample (data not shown due to space limitations). The distribution of respondents, however, is different than these sample distributions. While Army makes up only 40% of eligible respondents to the production survey, they comprise 51% of the nonresponse bias study eligible respondents, showing that Army responded well to the phone NRB study. Tables 4 and 5 show these effects using response rates weighted with base weights.

Tables 4 and 5 show the following:

1. Response rates for the telephone surveys exceed response rates to the production survey (Web survey with e-mail and postal notification);

2. Using the telephone greatly improves the response rates for typically low responding military subgroups (Army, Marine Corps, and E1-E4).

For example, Table 5 shows that junior enlisted members responded at a higher rate to the nonresponse bias study than to the production survey, whereas senior officers had the opposite trend. This is an important finding because DMDC generally expects non-

---

[7] The region variable is created based on a Service member's Unit Identification Code (UIC). In some cases, a member may be overseas but their UIC is still located in the United States. In this instance, the member would be categorized as US & US territories, although he or she would be physically overseas.

respondents to the production survey to be a 'difficult' to survey group, but the phone calls to E1-E4 non-respondents produced higher response rates than e-mail and postal contacts to the full E1-E4 population. This finding could be attributable to improved survey contact or improved survey cooperation. It is possible that e-mail and postal mail contacts are not reaching E1-E4 at high rates, a survey contact problem, and the phone is simply a better method for contacting this group. It is also possible that both methods are equally effective at contacting this group, but the phone influences their participation decision (perhaps due to more social pressure to respond).

Table 4: Location, Completion, and Response Rates, by Service

Service	Production (Web)	Mode (Phone)	NRB (Phone)
Army	12%	20%	14%
Navy	15%	16%	17%
Marine Corps	12%	22%	21%
Air Force	20%	22%	15%
Coast Guard	28%	41%	18%

Table 5: Location, Completion, and Response Rates, by Paygrade

Paygrade	Production (Web)	Mode (Phone)	NRB (Phone)
E1-E4	6%	17%	16%
E5-E9	17%	24%	17%
W1-W5	16%	13%	18%
O1-O3	23%	21%	16%
O4-O6	36%	26%	11%

6. Mode Study Findings

Due to the different modes used to implement the production survey and NRB study, direct comparison between the two only makes sense if there is no mode effect on survey responses. If there is a mode effect, then nonresponse bias cannot be isolated and a direct comparison between estimates from the production and NRB studies is not meaningful. The previous section indicates that response rates are affected at both the subgroup and overall levels by the implementation of a telephone survey, but does not inform any possible effect on response data (mode effects).

Unweighted responses to the mode study are significantly different from the production study for about 1/3 of the mode-study questions. For example, Table 6 shows that the unweighted proportion of members who responded that they had 'visited the Federal Voting Assistance Program's Website prior to the 2010 election' in the mode study was lower than the Web-administered survey (12% vs. 22%).

Table 6: "In preparation for the 2010 primaries and general election, did you visit the Federal Voting Assistance Program Web site?" (Frequency and Percent Responding Yes)

	Production		Mode	
Frequency	*Percent (unweighted)*		*Frequency*	*Percent (unweighted)*
1,902	22%		69	12%

This difference is affected by both mode effects (measurement error) and sample composition. For instance, the respondents to the mode study are disproportionately young and enlisted relative to the respondents from the production survey. Therefore, their responses may be different simply because they are younger and not because of a mode effect.

Because the mode study sample is independent from the production sample, the mode study respondents can be weighted to represent the full population in the same way that the production sample is weighted. The consistent weighting scheme accounts for the differing distributions of respondents and allows for comparisons between the estimates to investigate mode effects.[8]

Tables 7 and 8 show the weighted estimates for two questions of interest: voter registration and voter participation. For both questions, the weighted production estimates are within the margins of error of the mode study estimates.

Table 7: "In 2010, were you registered to vote in the United States?" (Frequency and Percent Responding Yes)

Production - Weighted			Mode Study - Weighted		
Frequency	*Percent*	*Margin of Error*	*Frequency*	*Percent*	*Margin of Error*
1,201,250	77%	±3	1,129,639	79%	±5

Table 8: "In the election held on November 2, 2010, did you definitely vote…?" (Frequency and Percent Responding Yes)

Production - Weighted			Mode Study - Weighted		
Frequency	*Percent*	*Margin of Error*	*Frequency*	*Percent*	*Margin of Error*
459,155	29%	±2	506,835	34%	±6

The estimates from Tables 7 and 8 are not statistically different from one another due to the margins of error. The differences that do exist, however, are in the direction hypothesized in the literature: the belief underpinning social desirability bias is that respondents will report higher voting rates to a telephone or face-to-face interview for the sake of self-presentation (Holbrook and Krosnick, 37). Although these estimates are not

---

[8] The production weighting generally uses a form of weighting classes. For the mode study, because there are far fewer cases, it is impractical to use the same number of weighting classes, and therefore DMDC collapsed some classes to ensure sufficient numbers of cases per class.

statistically different, the possible existence of social desirability bias could affect comparisons between the production and NRB surveys.

The results for all other questions that were asked in the mode study could not be included due to length requirements. Most questions exhibit little to no difference between the production and mode surveys, and the margins of error for each survey cover the estimate for the other in most cases. For these questions, therefore, the choice of mode does not seem to affect survey responses.

However, two groups of questions seem to show dramatic mode effects: questions asking from where the survey respondent received information about the election in 2010 and questions asking about access to the Internet and other technologies. For example, 34% of respondents from the production survey said that they had received information from local television, radio, or print media in 2010, compared with 58% from the mode study. This drastic difference is echoed in nearly all the questions about receiving information and conforms to the idea of social desirability – namely, when taking a survey with a more personal medium, such as a phone call, respondents are more likely to give answers that they believe the surveyor 'wants' to hear. In this group of questions, it is possible that respondents to the mode study were influenced by the use of the phone interview into over-reporting their sources of knowledge.

This phenomenon also exists in the group of questions concerned with access to the Internet and other technologies, although the differences are not as dramatic as in the first group. For instance, 74% of respondents to the production survey reported having access to a personal e-mail address around the time of the 2010 election, but 89% of mode study respondents claimed the same access. Again, this seems to indicate that the use of the phone call encourages respondents to provide socially desirable answers, in this case affirming they have access to e-mail and other technologies.

It is notable, however, that mode effects seem modest in questions related specifically to voting. Fewer people in the mode study than in the production survey (47% vs. 50%) responded that they had voted in the 2008 election. The margins of error of both surveys indicate that these responses are not statistically different.

7. Nonresponse Bias Study Findings

When examining unweighted data, NRB is present in many of the questions for the *2010 Post-Election Voting Survey of Uniformed Service Members*. However, as in the case of the mode effect, its presence can at least partially be explained by the differing distributions of eligible respondents to the three surveys. DMDC's weighting methodology successfully reduces a large amount of NRB present in unweighted estimates; however it remains difficult to assess the remaining level of NRB.

Table 9, which shows voting rates across the three surveys (shown unweighted and weighted for production survey), shows the effectiveness of the weighting process in reducing NRB for this question.

Table 9:"In the election held on November 2, 2010, did you vote…?"
(Frequency and Percent Responding Yes)

Production (Unweighted)		Mode (Unweighted)		NRB (Unweighted)		Production (Weighted)	
Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
3,720	44%	197	35%	109	22%	459,155	29% (±2)

According to the unweighted data, there appears to be both a mode effect and an NRB effect. Namely, those members who responded by phone were less likely to vote (35%) than those that responded to the Web-administered production survey (44%), and those who responded to the NRB study were even less likely to have voted (22%). However, recall that the phone surveys were successful at gaining response from young, junior enlisted members, who vote at much lower rates. Differences in voting rates are likely to be, at least in part, influenced by sample composition. The weighted voting rate (29%), which adjusts for demographic differences between the respondents to the production survey and the frame population, is 15 percentage points lower than the unweighted voting rate (44%). This indicates that the weighting process 'moved' the unweighted estimate in the direction of data from the NRB study, which indicates the weighting process is reducing NRB. The question remaining is whether the magnitude of movement from 44 percent to 29 percent is correct, which is assessed in the following paragraph.

If the respondents to the NRB survey perfectly represent all non-respondents to the production survey,[9] and DMDC lets the production survey respondents represent only 15 percent of the population (weighted response rate was 15%), then a weighted average of the data from the two surveys should give an accurate estimate of the population value. This is an unlikely assumption to make, but does provide some indication of the effectiveness of the weighting process. In this case, with a 15% response rate to the production survey, a weighted average would be $(44)^{10} *(.15)^{11} + (22)^{12} *(.85)^{13} = 25\%$. Therefore, the weighting process adjusts the unweighted estimate from the production survey downward in the direction of the non-respondents, as shown by the estimates from the NRB survey. The composite estimate of 25% is not within the margin of error (±2) of the weighted estimate of 29%. However, with considerably fewer respondents in the NRB study than in the production survey, DMDC would expect the margin of error to be much wider around the NRB estimate and would need to test whether the difference between the two estimates are statistically significant. Recall that DMDC saw some evidence of potential over-reporting of voting via telephone in the mode study, so it is possible the 22% voting rate from the NRB study should be closer to 18-20 percent, and this may indicate the weighting greatly reduces the NRB, but may fail to correct enough and eliminate the NRB for the voting rate.

---

[9] Note that this is an untestable assumption.
[10] Unweighted voting rate.
[11] Production survey response rate.
[12] Voting rate from NRB study.
[13] Letting respondents only represent the 15% response rate, the non-respondents represented by the NRB study comprise the other 85% of the population.

Table 10: "In 2010, were you registered to vote in the United States?"
(Frequency and Percent Responding Yes*)*

Production (Unweighted)		Mode (Unweighted)		NRB (Unweighted)		Production (Weighted)	
Frequency	*Percent*	*Frequency*	*Percent*	*Frequency*	*Percent*	*Frequency*	*Percent*
7,140	84%	439	81%	384	80%	1,201,250	77% (±3)

The same composite estimate analysis can be applied to other areas of interest, such as voter registration. Following the same formula described above, the composite estimate for the percent of the military that was registered to vote in 2010 would be $(84)*(.15) + (80)*(.85) = 81\%$. Because the production estimate was 77%, this indicates the weighting process may have slightly overcorrected for NRB. While the composite value is slightly outside of the margin of error of the weighted production estimate (77 ± 3), the difference is not statistically significant due to the large variance of the composite estimate due to the small number of respondents in the NRB survey, as was the case with the question on voter participation.

To this point, the NRB analysis has treated all respondents to the production survey as one group and non-respondents as another group. If the Missing at Random (MAR) assumption is true, and that survey respondents in the same weighting classes are MAR, then DMDC should be able to weight the respondents from the NRB study to the full population and produce an unbiased estimate of the voting rate (similar to the production survey). DMDC maintains a uniquely rich set of correlates to use in the weighting process, and this weighting would be a very powerful display of the ability to remove bias from estimates by essentially re-producing the same weighted estimates regardless of the survey respondents. However, the weighted NRB study estimates show voting rates lower than the weighted production estimates, possibly indicating that non-respondents are systematically different than respondents, and DMDC cannot completely control for these differences during weighting through the use of administrative variables. Tables 12 and 13 show the results of the production and NRB surveys each weighted to the full population for the voter registration and participation questions (estimates for all other questions are omitted due to space limitations).

Tables 12 and 13 also show estimates from a continuum of resistance model. Respondents to the production and NRB surveys are not likely to be homogeneous. For example, someone who responds on the last day of the production survey field period may be more similar to NRB respondents than to the first respondents to the production survey. In other words, respondents could be on a continuum where respondents to the NRB study behave similarly to late respondents from the production survey.

To evaluate this, the production estimates are compared to two other estimates. The first recognizes that respondents often come in waves due to continued contact efforts. In the production survey, non-respondents continued to receive e-mail notifications until the closing of the survey field period. If two fewer e-mail notifications had been used, which would cut the field period short by approximately two weeks, this reduced group of respondents could be weighted to the full population as if it had been the intended production survey.

The second group of estimates can be created by adding the NRB respondents to the production respondents, as if the field period had been extended and these NRB

respondents had been additional respondents. This group can also be weighted up to the full population, creating a third estimate for comparison. The first estimate uses a survey field period that is shorter than the production survey, the production survey itself acts as the second estimate, and the third estimate extends the production survey by incorporating the NRB respondents.

Table 11 shows the sample disposition codes for the production survey as well as what the sample disposition codes would have been for a shorter or longer field period. Due to the small number of NRB respondents relative to the production survey, only 496 compared to 8,546, the distribution changes very little between the production and production plus NRB survey periods.

Table 11: Sample Disposition Codes for the Three Survey Field Periods

Disposition Code	Production Minus 2 E-mails		Production		Production Plus NRB	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
1 Record ineligible based on sample file[a]	1,432	2%	1,342	2%	1,432	2%
2 Ineligible--Self or proxy report[b]	29	0%	30	0%	36	0%
3 Ineligible--Survey self report[c]	1,385	2%	1,820	2%	1,871	3%
4 Complete eligible response	7,024	10%	8,546	12%	9,042	12%
5 Incomplete eligible response[d]	500	1%	647	1%	660	1%
8 Refused/other	652	1%	663	1%	947	1%
9 Blank	546	1%	731	1%	693	1%
10 Postal non-deliverable (PND)[e]	9,197	13%	9,197	13%	8,966	12%
11 Non-respondents	52,479	72%	50,178	69%	49,597	68%

Note: Footnotes a-e are identical to the footnotes on Table 1 and can be found there.

With so few respondents added by incorporating the NRB respondents, it is unsurprising that none of the estimates from the production plus NRB time frame are statistically significant from the production survey. The shorter time frame, which eliminates 1,522 respondents from the 8,546 in the production survey, also does not have any estimates that are statistically significant from the production survey. However, several questions do show signs that early responders were more likely to have been active in the voting process.

Table 12: "In the election held on November 2, 2010, did you vote…?"
(Frequency and Percent Responding Yes)

Production (Weighted)		Production Minus 2 E-mails (Weighted)		Production Plus NRB (Weighted)		NRB Only (Weighted to Full Population)	
Percent	Margin of Error	Percent	Margin of Error	Percent	Margin of Error	Percent	Margin of Error
29%	±2	30%	±3	29%	±2	21%	±6

Table 13: "In 2010, were you registered to vote in the United States?"
(Frequency and Percent Responding Yes)

Production (Weighted)		Production Minus 2 E-mails (Weighted)		Production Plus NRB (Weighted)		NRB Only (Weighted to Full Population)	
Percent	Margin of Error	Percent	Margin of Error	Percent	Margin of Error	Percent	Margin of Error
77%	±3	77%	±3	76%	±2	74%	±6

Table 12 shows that voter participation for the military based on the weighted estimate for the shorter time frame is 30%, but the production estimate is 29%. While this difference is very small, in conjunction with several other questions, a pattern begins to emerge. For instance, the shorter time frame produces slightly higher estimates for the usage of voting resources such as unit voting assistance officers, installation voting assistance officers, and the FVAP Website. Although none of these estimates is statistically different from the production estimate, the fact that each is higher than the production estimate by 1-2% indicates that early respondents are possibly more active voters.

The differences are most distinct for the NRB only data. When taken on their own and weighted to the full population, the NRB respondents do have some statistically significant differences from the production survey. For example, the weighted estimate for the percent of the military that usually plan on voting is much lower than the production estimate (42% vs. 54%). The NRB respondents also seem to be less interested in the election and, although they reported higher use of unit voting assistance officers, they are less likely to use installation voting assistance officers, the Voting Assistance Guide, or FVAP's Website. While the use of Unit Voting Assistance Officers, Installation Voting Assistance Officers, and the Voting Assistance Guide are not statistically significant, the general trend seems to point toward the possibility of NRB, as non-respondents to the production survey may have different tendencies than respondents.

8. Conclusions and Future Considerations

Based on the results of the mode and NRB studies, DMDC believes that significant biases exist in the unweighted data. However, the post-survey weighting process does seem to have minimized the effects of NRB on production estimates. Despite the small sample sizes of the two studies and the presence of some questions that indicate persistent biases, such as the questions on sources of election knowledge that seem to have mode effects, nearly all the unweighted production estimates seem to be brought more in line with the NRB data by the weighting process. Therefore, while NRB is likely still present in the production estimates, its effect seems to be greatly diminished.

In an effort to reduce NRB in the *2012 Post-Election Voting Survey of Uniformed Service Members*, DMDC plans to contact 'hard-to-survey' junior enlisted survey members by telephone to encourage survey response. If the phone contacts successfully increase response rates for junior enlisted as they did in 2010, the survey estimates will be less dependent on weighting methods to reduce non-

response bias, and our risk of NRB significantly effecting survey estimates will be diminished. Because there was evidence from this study that mode effects exist for voting questions, DMDC does not plan to collect any survey data by telephone.

References

The American Association for Public Opinion Research. 2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 7th edition. AAPOR.

Armstrong, Scott and Terry Overton. "Estimating Nonresponse Bias in Mail Surveys." *Journal of Marketing Research*. 1977. Vol 14 (3): 396-402.

Fowler, Jr., Floyd, et al. "Using Telephone Interviews to Reduce Nonresponse Bias to Mail Surveys of Health Plan Members." *Medical Care*. 2002. Vol 40 (3): 190-200.

Holbrook, Allyson and Jon Krosnick. "Social Desirability Bias in Voter Turnout Reports." *Public Opinion Quarterly*. 2010. Vol 74 (1): 37-67.

Groves, Robert. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly*. 2006. Vol 70 (5): 646-675.