

PERILS OF POLLING IN THE 2008 ELECTION

Scott Keeter, Jocelyn Kiley, Leah Christian and Michael Dimock
The Pew Research Center

Presented at the Annual Meeting of the American Association for Public Opinion Research
Hollywood, Florida, May 14-17, 2009

Abstract

The validity of public opinion polling in the presidential election of 2008 was thought to be seriously imperiled by a wide range of potential survey errors, including coverage error due to the growth of the wireless-only population, nonresponse error potentially caused by differential nonresponse among Republicans and racially conservative voters, and measurement error from racially-related underreporting of support for the Republican candidate and greater-than-usual difficulties in forecasting turnout and identifying likely voters. Despite these obstacles, polls at both the national and state level performed well, matching or exceeding the accuracy achieved in 2004. Using a range of data sources including several dual frame surveys with cell phones, a survey of respondent households from previous surveys, and a race-of-interviewer analysis, we examine the possible sources of bias in the election polls and demonstrate that these were neither large enough nor systematic enough to create serious errors in most polls' election projections.

The analysis of total survey error has evolved over many decades to consider a wide variety of potential threats, including concerns about the contribution of both bias and variance, and an attention to errors of both observation and non-observation (Groves 1989). The validity of public opinion polling in the presidential election of 2008 was thought to be seriously imperiled by a wide range of these potential errors. Among these were coverage error due to the growth of the wireless-only population, nonresponse error potentially caused by differential nonresponse among Republicans and racially conservative voters, and measurement error potentially resulting from racially-related understatement of support for the Republican candidate and greater-than-usual difficulties in forecasting turnout and identifying likely voters.

Despite these obstacles, polls performed very well, with 8 of 17 national polls predicting the final margin in the presidential election within one percentage point and most of the others coming within three points. Both at the national and state levels, the accuracy of the polls matched or exceeded that of 2004, which was itself a good year for the polls. The performance of election polls is no mere trophy for the polling community, for the credibility of the entire survey research profession depends to a great degree on how election polls match the objective standard of election outcomes. The consequences of a poor performance were dramatically demonstrated in the reaction to the primary polls' inaccurate prediction that

Barack Obama would win in New Hampshire, portrayed as one of polling's great failures in the modern political era (AAPOR 2009).

We examine the challenges of potential coverage bias from excluding cell phones and potential measurement and non-response bias due to race in detail using data from a wide range of sources, including a summary analysis of state and national pre-election polls, six telephone surveys conducted among both landline and cell phone samples, and a comparison of a survey conducted by landline with reluctant and elusive respondents with a survey conducted at the same time with a fresh sample using standard methodology. Our conclusion is that some of the threats were very real but overcome by the techniques normally employed in surveys to address potential bias from various sources of error, while other threats turned out to be less serious than some anticipated.

1. POLLING ACCURACY

Pre-election polls conducted by telephone did very well in forecasting the outcome of the election in 2008. This was true for polls using live interviewers and those conducted with recorded voices. It was true for those based only on landline interviews and those that included cell phones. The basic methodology of the telephone survey remains robust in the face of the many challenges now facing this mode of data collection.

Our assessment uses data and estimates compiled by the National Council on Public Polls (NCPP), which evaluated 17 national presidential polls and 236 state polls conducted in the final week of the campaign, covering the presidential vote and votes for U.S. Senate and governor. Its measure of accuracy was the average candidate estimate error, defined as half of the difference between the actual election margin minus the poll's margin.

For the 17 national telephone polls evaluated, the mean candidate estimate error less than 1 percentage point error on each presidential candidate (0.8%). Among the 11 national landline-only polls, four underestimated Obama's support, five overestimated it, and two had the margin exactly right. The absolute average candidate error for these landline-only surveys was 0.8%. Among the six dual frame surveys, one underestimated Obama's margin and four overestimated it; one had the margin exactly right. The average candidate estimate error for the dual frame surveys was also 0.8%.

Errors in polling at the state level were larger but still relatively small. The NCPP collected statewide polling data on the presidential race from 146 polls conducted from October 27, 2008 through Election Day, with an average candidate error of 1.6 percentage points. Including additional statewide races for senate and governor for a total of 237 races, the average candidate error for these races was 1.9 percentage points, about the same as in 2004 (1.7 percentage points). Of all state races polled by landline telephone and tracked by NCPP with most interviews conducted October 27 or later (237), more had errors favoring the Republican candidate (125) than the Democrat (86). But the mean error in each direction was about the same (approximately 2.0% for each). The mean error among IVR polls (1.7%) was slightly lower than among those with live interviewers (2.1%).

While the polling errors were greater at the state level than at the national level, the fact that they were little changed from 2004 was notable, given the sharp increase in the percentage of Americans with no landline phone and our presumption that all or nearly all of the state polling was conducted among landline samples. Of course, the landline non-coverage

Table 1
Error Estimates for Telephone Polls
In the 2008 Election

National Polls	Mean <u>Error</u> %	Number <u>of polls</u>
All polls	0.8	17
Landline only	0.8	11
Landline and cell samples	0.8	6
Among errors favoring...		
Obama	1.1	9
McCain	0.8	5
No error	-	3
State Polls	Mean <u>Error</u> %	Number <u>of polls</u>
<i>Presidential Race</i>	1.6	146
Battleground states	1.4	76
Phone with interviewer	1.6	77
Phone with IVR	1.5	69
Among errors favoring...		
Obama	1.9	52
McCain	1.7	74
No error	-	20
<i>All Races*</i>	1.9	237
Battleground states	1.5	77
Phone with interviewer	2.1	121
Phone with IVR	1.7	116
Among errors favoring...		
Democratic candidate	2.1	86
Republican candidate	2.0	125
No error	-	26

Source: Analysis by the authors of rounded poll results from telephone polls completed in the week prior to the election compiled by NCPP <http://ncpp.org/?q=node/114> from public sources. Entries show the mean candidate estimate error, which is the margin of victory in the election minus the poll margin divided by two.

* Note: the number of polls for "all races" includes multiple races on single polls (e.g., Presidential, Senatorial, Gubernatorial)

rate is not uniform across all states. Estimates of the prevalence of wireless-only adults for 2007 by the National Health Interview Survey (NHIS) and State Health Access Data Assistance Center (SHDAC) at the University of Minnesota ranged from 4.0% in Delaware to 25.1% in Oklahoma and 25.4% in the District of Columbia (Blumberg et al., 2009). Thus the potential for bias is greater in some places than others.

2. THE NON-COVERAGE THREAT: A SMALL BUT REAL BIAS IN LANDLINE SAMPLES

The cell phone problem in telephone survey research is well documented. As many as one-in-five voting age adults live in wireless-only households, and there is widespread evidence that they are not only demographically distinct but also differ in certain behaviors – particularly those related to health. (Blumberg and Luke 2009). In addition to the wireless-only coverage problem, evidence that some adults are “wireless mostly” and are difficult to access over landline telephones suggests that coverage problems may be even more widespread. When it comes to political attitudes and voting patterns, however, evidence that adults in wireless-only households differ substantially from their counterparts with landline phones is less definitive, especially when demographic characteristics are held constant (Pew Research Center 2008). As a result, while there is a clear coverage problem in pre-election landline-only surveys, the question of whether effective demographic weighting of landline-only surveys can effectively reduce or eliminate any resulting bias remains an open one.

An analysis of six Pew Research surveys conducted from September through the weekend before the election shows that estimates based only on landline interviews weighted to basic demographic parameters were likely to have a small pro-McCain bias compared with estimates based on both landline and cell phone interviews weighted similarly. Other survey organizations reported a similar result.

But the difference, while statistically significant, was small in absolute terms – smaller than the margin of sampling error in most polls. Obama’s average lead across the six surveys was 9.9 points among registered voters when cell phone and

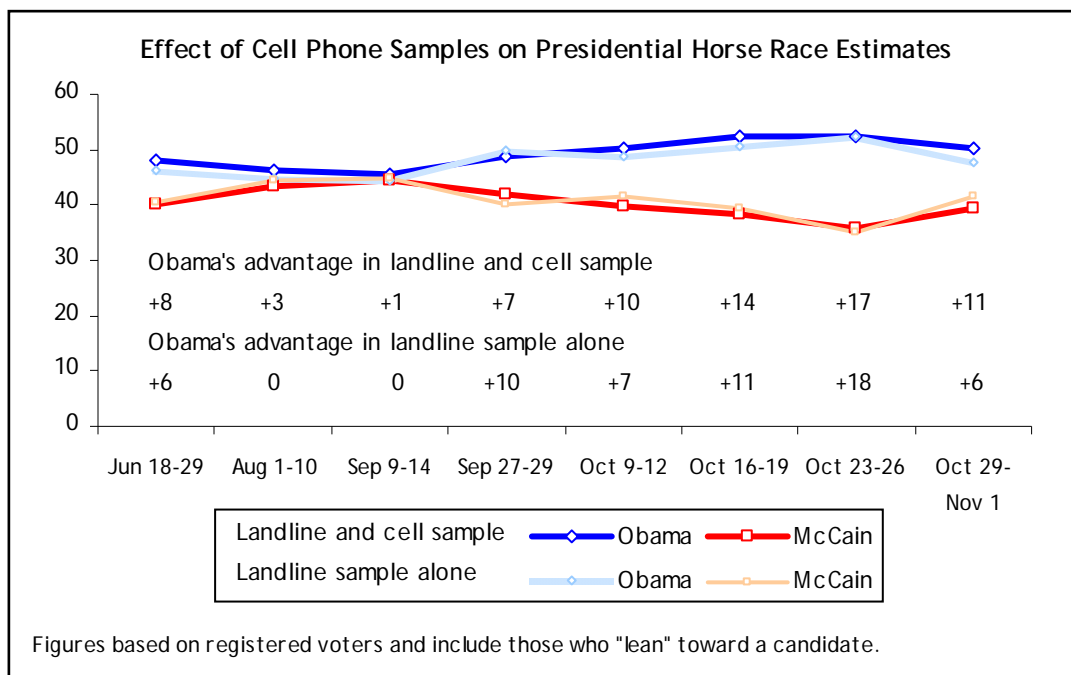
Table 2
Samples With Cell Phones Showed a Larger Obama Advantage

	Landline/ cell phone <i>Registered voters</i>	Landline/ cell phone <i>only</i>	Landline sample <i>sample</i>
	%	%	%
Obama	49.9	49.4	48.5
McCain	40.0	40.3	40.9
Other/DK	<u>10.1</u>	<u>10.3</u>	<u>10.6</u>
	100	100	100
<i>Obama Adv.</i>	+9.9	+9.1	+7.6
Sample size	(11,964)	(10,039)	(9,228)
<i>Likely voters</i>			
Obama	49.8	49.2	48.5
McCain	41.6	42.1	42.7
Other/DK	<u>8.6</u>	<u>8.7</u>	<u>8.8</u>
	100	100	100
<i>Obama Adv.</i>	+8.2	+7.1	+5.8
Sample size	(10,819)	(8,767)	(8,143)

Figures based on weighted data from Pew Research Center pre-election polls conducted from September through November 2008 and include those who “lean” toward a candidate.

landline interviews were combined and weighted. If estimates had been based only on the weighted landline samples, Obama’s average lead would have been 7.6 points, an average bias of 2.3 percentage points on the margin, or about 1.2 points expressed as candidate error. Limiting the analysis to likely voters rather than all voters produced similar results. Obama’s average lead among likely voters was 8.2 points across all six dual frame surveys versus 5.8 points (or 1.2 points as candidate error) when the landline samples are analyzed alone. (See the appendix for a detailed description of the sampling and weighting employed in this analysis.)

While estimates based only on landline interviews typically exhibited a pro-McCain bias, the pattern was not uniform. Four of the six surveys conducted after the August conventions fit the pattern; the largest difference was in the final election weekend survey where Obama led McCain by 11 points in the dual frame sample, but by six points if only landline interviews were considered. Yet in two of the six surveys this pattern did not hold. In late September and late October, Obama’s lead was slightly narrower in the combined landline and cell survey than in the landline survey alone. This indicates that the overall pattern, while important, was not large enough to overcome normal sampling fluctuation.



The fact that the bias related to phone status was relatively small, despite the large demographic differences between the cell-only and landline-accessible populations, is a function both of the proportion of all voters who are cell-only (i.e., the relative size of the cell-only population) and the effects of demographic weighting. Weighting will help minimize this bias as long as the weighting variables correlated with phone status are also related to the

political measures of interest for both cell-only and landline-accessible voters. Put differently, voters reachable by landline who share certain demographic characteristics with cell-only voters are more similar politically to cell-only voters than to other landline voters. Not all of the variables that are strongly associated with phone status and political behavior are currently being used in typical weighting protocols; among these are marital status, presence of children in the household, family income and home ownership. This suggests that there is untapped opportunity for further reduction in cell-only bias with the use of additional weighting variables, assuming these can be measured reliably and that adequate parameters are available. One way to assess the potential effectiveness of weighting is to estimate the impact of cell-only status on the vote with and without these controls.

Logistic regression was used to estimate the probability of voting for Obama among landline voters and cell-only voters. As would be expected, the difference is sizeable; the predicted probability of voting for Obama is 16 points higher for cell-only voters than for landline voters. Adding most of the standard demographic variables used in weighting (e.g., age, sex, race, Hispanicity education, and region) to the model (labeled the “standard model” in Table 3) reduces this difference to 11 points, a result consistent with the notion that weighting helps

Table 3
Effects of Cell Phone Only and Other Demographics on Support for Obama

	<i>Basic model</i>		<i>Standard model</i>		<i>Expanded model</i>	
	<u>Coef</u>	<u>SE</u>	<u>Coef</u>	<u>SE</u>	<u>Coef</u>	<u>SE</u>
Cell phone only	.631 ***	(.12)	.466 ***	(.14)	.181	(.15)
Male	--	--	-.367 ***	(.09)	-.381 ***	(.10)
Black non-Hispanic	--	--	2.673 ***	(.22)	2.50 ***	(.23)
Hispanic	--	--	-.762 ***	(.16)	-.607 ***	(.16)
Other non-white	--	--	.181	(.20)	.227	(.22)
Age	--	--	-.006 *	(.00)	-.004	(.00)
Non college grad.	--	--	-.278 **	(.10)	-.423 ***	(.11)
South	--	--	-.707 ***	(.10)	-.682 ***	(.10)
Income < \$75k	--	--	--	--	.208	(.12)
Married	--	--	--	--	-.430 ***	(.11)
Own home	--	--	--	--	-.303 **	(.13)
<i>Predicted probability of Obama vote among...</i>						
Cell phone only	.59		.55		.50	
Landline accessible	.43		.44		.45	
N	2,500		2,432		2,137	

Unstandardized logistic regression coefficients.
Based on registered voters in the Pew Research Center's 2008 Mid-September pre-election poll.
*p<.05, **p<.01, ***p<.001

reduce but not eliminate the potential for non-coverage bias. Including income, marital status and home ownership in the model reduces the difference even further to 5 points. When these additional demographics are included in the model, being cell phone only is no longer a significant predictor of candidate support, as it was in the first two models.

Although the evidence from the 2008 election indicates that cell-only respondents may pose a relatively minor threat of bias to most telephone surveys, a related threat also attracted attention: respondents who rely mostly on their cell phones and thus might be difficult to reach by landline even if they have one. The issue is whether the wireless mostly group is adequately represented by landline respondents who have both a cell phone and a landline, but rely mostly on their cell phone. Data collected during the 2008 election campaign suggests that while the wireless mostly reached by cell phone are somewhat different from those reached by landline, combined samples of wireless mostly voters from both sampling frames differ only slightly from the wireless mostly who are reached by landline after standard demographic weighting. On the issue of candidate preference, 55% of all wireless-mostly voters interviews in the six Pew Research pre-election surveys supported Obama for president compared with 51% of wireless mostly from the landline sample; differences in party, ideology, and political engagement were smaller. The validity of this generalization depends upon an unknown quantity, namely what proportion of interviews of the cell-mostly group *should* come from each frame to produce the most valid representation of the group. In our surveys approximately 40% of the cell-mostly group comes from the wireless frame. But whatever the best mix, the potential for bias on the total survey estimate is modest, given the fact that wireless-mostly respondents constitute only about 15% of all adults (Blumberg and Luke 2009) and thus far most research suggests that they are reachable by landline surveys.

Table 4
Attitudes and Behavior of the
“Wireless Mostly”

	All wireless mostly	Wireless mostly from landline
	%	%
<i>Party Affiliation</i>		
Republican	24	27
Democrat	37	36
Independent	34	32
Republican/Lean Rep.	37	40
Democrat/Lean Dem.	54	50
<i>Ideology</i>		
Conservative	34	36
Moderate	38	38
Liberal	24	22
Sample size	(1,825)	(1,196)
Registered to vote	77	80
Sample size	(2,349)	(1,524)
<i>Among registered voters:</i>		
<i>Presidential vote</i>		
Obama/lean Obama	55	51
McCain/lean McCain	38	41
A lot of thought given to the election	81	83
Closely following campaign news	88	90
Ever voted in precinct*	76	78
“Always” vote*	50	50
Sample size	(1,908)	(1,289)
Figures based on weighted data from Pew Research Center pre-election polls conducted from September through November 2008.		

3. MEASUREMENT ERROR AND NON-RESPONSE: THE “BRADLEY EFFECT”

Problems with pre-election polls in biracial elections in the 1980s and early 1990s raised the question of whether covert racism remained an impediment to black candidates (Keeter and Samaranayake 2007; Hopkins 2008; Hugick 1990)). White candidates in many of these races generally did better on Election Day than they were doing in the polls, while their black opponents tended to end up with about the same level of support as the polls indicated they might. This phenomenon, often called “the Bradley effect,” was first noticed in the 1982 race for governor of California, where Los Angeles Mayor Tom Bradley, a black Democrat, narrowly lost to Republican George Deukmejian, despite polls showing him with a lead ranging from 9 to 22 points.

The accuracy of the polls in the general election and – with the notable exception of the New Hampshire primary – the long series of Democratic primaries provides more than adequate refutation of a Bradley Effect in the 2008 presidential election, at least at a magnitude that could seriously undermine the accuracy of pre-election polls. Indeed, evidence from five statewide elections in 2006 involving black and white candidates, in which polling was quite accurate, strongly suggested that the Bradley Effect was no longer potent (Keeter and Samaranayake 2007). Still, whether the Bradley Effect would play a different role in a contest for the presidency than in a gubernatorial or Senate race was unknown, and the possibility of seriously biased polls in 2008 was a frequent subject of political discussion.

**Table 5
Exaggerated Margins of Support for Black Candidates in Past Biracial Elections***

<i>Illinois Senate, 1992</i>				
	<u>Braun</u>	<u>Williamson</u>	<u>Diff.</u>	
	%	%		
Actual Vote (11/3)	53	43	+10	
Daily Herald (10/29)	52	32	+20	
Chicago Tribune (10/28)	51	34	+17	
Copley (10/28)	46	28	+18	
<i>Illinois Senate Democratic Primary, 1992</i>				
	<u>Braun</u>	<u>Dixon</u>	<u>Hofeld</u>	<u>Diff.</u>
	%	%	%	
Actual Vote (3/17)	38	35	27	+3
Gallup (3/12)	29	41	21	-12
Political Rsch (3/12)	18	37	35	-19
<i>North Carolina Senate, 1990</i>				
	<u>Gantt</u>	<u>Helms</u>	<u>Diff.</u>	
	%	%		
Actual Vote (11/6)	47	53	-6	
Charlotte Observer (11/1)	47	41	+6	
News & Observer (10/29)	40	44	-4	
Mason-Dixon (10/31)	48	44	+4	
<i>Virginia Governor, 1989</i>				
	<u>Wilder</u>	<u>Coleman</u>	<u>Diff.</u>	
	%	%		
Actual Vote (11/7)	50.2	49.8	+0.4	
Washington Post (11/2)	52	41	+11	
Rich. Times-Dispatch (11/1)	45	36	+9	
Mason-Dixon (11/1)	48	44	+4	
<i>New York City Mayor, 1989</i>				
	<u>Dinkins</u>	<u>Giuliani</u>	<u>Diff.</u>	
	%	%		
Actual Vote (11/7)	50	48	+2	
Daily News/ABC (11/3)	53	39	+14	
New York Observer (10/29)	49	31	+18	
<i>Chicago Mayor, 1983</i>				
	<u>Washington</u>	<u>Epton</u>	<u>Diff.</u>	
	%	%		
Actual Vote (4/12)	51.5	48.1	+3.4	
WLS-TV (4/9)	52	38	+14	
Gallup/Sun-Times (3/30)	51	37	+14	
WLS-TV (3/27)	48	34	+14	

*For each election the African American candidate is listed first.

Despite the accuracy of the 2008 primary polls in hindsight, we concluded that it was prudent to dissect the possible mechanisms by which the Bradley Effect could operate and evaluate the potential for a bias so that precautions could be taken. The Bradley Effect could be the result of two different phenomena: reluctance by racially conservative poll respondents to say that they intended to vote against the black candidate, or a greater resistance among racially conservative voters to be interviewed. The first of these – measurement error due to a “social desirability bias” that manifests itself on many sensitive topics in surveys – can be studied indirectly through the use of such techniques as the “list experiment” and a comparison of interviews conducted by white and black interviewers. To test this, we analyzed differences in responses by race of interviewer to assess the degree of racial sensitivity in questions about Obama’s candidacy and other questions measuring racial attitudes. The second source of potential bias is from non-response error related to the salience or nature of the survey topic or the presumed sponsor (the “mainstream media”). This might be detected by comparing poll respondents reached in a normal survey with those who initially refused to participate or were very difficult to reach for an interview. Non-response bias affected the accuracy of the exit polls in both 2004 and in the 2008 primaries and general election. To test for this second source of error, we made an effort to reach reluctant respondents and compare them with samples reached using our normal interviewing protocol.

Race of Interviewer Analysis

We found little evidence of racial sensitivity in the patterns of responses based on the race of respondent and the race of the interviewer. Unlike previous elections involving white and black candidates (Guterbock, Finkel and Borg 1991), there is little to suggest that voters' responses were significantly affected by the race of the person interviewing them over the phone. Among white non-Hispanic registered voters in the six pre-election Pew Research Center polls beginning in mid-September, there were no systematic differences in candidate support by race of interviewer, either among all white non-Hispanic voters, or among white Democratic voters (Democrats and Democratic leaning independents).

There also were no systematic differences among black voters (not shown), who overwhelmingly supported Barack Obama.

Over these six polls, a significant race of interviewer effect was found only once. In the mid-September poll, counter to the expectation of a social desirability effect, white Democratic voters who spoke with black interviewers were 8 percentage points less likely to express support for Obama. In later surveys, differences by race of interviewer were neither consistent in either direction nor significant.

Multivariate analysis confirms this finding; logistic regressions on candidate support found no significant effect of race of interviewer on support for either Obama or McCain, either among all white non-Hispanic voters or among white non-Hispanic Democratic voters. The results in Table 7 are for the election weekend poll; the effect of race of interviewer was similarly non-significant impact on the two other large pre-election polls (mid-September and Mid-October).

Table 6
No Consistent Race of Interviewer Difference in Candidate Support among White Non-Hispanic Voters

	All White Voters			
	<i>White Non-Hispanic Interviewer</i>		<i>Black Non-Hispanic Interviewer</i>	
	<u>Obama</u>	<u>McCain</u>	<u>Obama</u>	<u>McCain</u>
	%	%	%	%
Mid-September	38	53	36	52
Late September	40	49	41	48
Early October	40	52	44	47
Mid-October	42	48	47	43
Late October	41	42	44	48
Election weekend	44	46	42	46

	White Democrats and Dem Leaning Independents			
	<i>White Non-Hispanic Interviewer</i>		<i>Black Non-Hispanic Interviewer</i>	
	<u>Obama</u>	<u>McCain</u>	<u>Obama</u>	<u>McCain</u>
	%	%	%	%
Mid-September	87	8	79	10
Late September	84	10	91	6
Early October	91	4	87	8
Mid-October	88	6	91	4
Late October	82	6	89	7
Election weekend	88	6	85	7

Based on white non-Hispanic registered voters. Figures based on

While there is little evidence to suggest that respondents were more reluctant to voice opposition to Obama when interviewed by African American interviewers than when interviewed by white interviewers, there was a small difference in the composition of the samples interviewed by white and black interviewers; this difference is consistent with the theory that reluctant whites may have self-selected out of interviews with black

Table 7
Support for Obama among White Non-Hispanic Voters

	<i>Total</i>		<i>Dem/Dem Leaners</i>	
	<u>Coef.</u>	<u>SE</u>	<u>Coef.</u>	<u>SE</u>
<i>Interviewer demographics</i>				
Black non-Hispanic	-.125	(.15)	-.380	(.25)
Hispanic	-.149	(.22)	-.378	(.38)
Other race	.011	(.33)	-.573	(.52)
Male	-.183	(.13)	.031	(.23)
<i>Respondent demographics</i>				
Male	-.052	(.13)	.288	(.24)
Education	.082	(.05)	.107	(.08)
Income	-.014	(.03)	.073	(.06)
Age	-.015 ***	(.00)	-.014 *	(.01)
Republican	-4.553 ***	(.21)	--	--
Independent/Other	-2.384 ***	(.16)	-.172	(.25)
Live in south	-.387 **	(.14)	-.682 **	(.23)
Constant	2.592 ***	(.35)	2.206 ***	(.55)
N	1939		849	

Unstandardized logisitic regression coefficients. Based on white non-Hispanic voters in Pew Research Center's election weekend survey. Excluded categories are: white non-Hispanic interviewers, female interviewers, female, Democratic party ID, and non-southerners.

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

interviewers. Black interviewers were less likely than their white counterparts to interview white respondents (and white Democratic respondents) on most of the six Pew Research Center election polls, and these differences were significant on the penultimate and final surveys before the election. For example, on election weekend, among Democratic respondents interviewed, 66% of those conducted by white non-Hispanic interviewers were with white non-Hispanic respondents, compared with only 59% of interviews conducted by black non-Hispanic interviewers. The previous week, this gap was even larger (68% compared to 51%). A similar pattern holds for the overall white sample on these surveys.

That African American interviewers were less likely to conduct interviews with white respondents could provide support for the hypothesis that racially conservative whites are more reluctant to respond to polls conducted by non-white interviewers, and thus contribute to a possible bias in the results. However, this finding may also be attributable to other differences between white and black interviewers that may be confounded with race. For instance, there was a somewhat uneven gender distribution (the percentage male among black interviewers was slightly higher than among white interviewers) and some differences in the schedules of white and black interviewers that may have affected the mix of respondents they interviewed (e.g., black interviewers were more likely to work on weekends). The fact that black interviewers were more likely to interview black respondents also may be the result of black respondents' greater receptivity to requests for interviews when called by a black

interviewer rather than white respondents' greater resistance to being interviewed by black interviewers.

Are Reluctant Respondents More Racially Conservative?

Evidence that reluctant respondents are more racially conservative is mixed. The Pew Research Center's 1997 non-response study found that the most difficult to interview respondents were slightly more racially conservative than those easier to interview (Pew Research Center 1998). But a follow up study conducted in 2003 found no such pattern. To evaluate this notion in the context of the 2008 campaign, we conducted a recontact survey of hard-to-reach households from earlier survey samples. To do so, we constructed a sample of landline telephone numbers based on households that had either refused to be interviewed or where at least five call attempts had been made with no completion in polls conducted by Pew Research between January and May 2008. The recontact interviews were conducted July 31-August 10, 2008, with 1,000 respondents. Results from these interviews were compared with a new national survey conducted at the same time among a landline sample of 2,254 respondents.¹

In the general election matchup, there were no significant differences in vote choice or strength of support between hard-to-reach voters and the comparable late August sample. McCain and Obama were tied at 44% among the hard to reach; McCain held a narrow 46% to 44% lead in the August sample. In both the August poll and the concurrent hard-to-reach sample, Obama received more strong support than McCain, and these proportions were nearly identical in the two samples. Hard-to-reach voters may have been slightly more likely to be swing voters, but the difference was not statistically significant (35% vs. 32% in the comparable August sample).

<i>All registered voters</i>	Unweighted	
	Hard-to-Reach %	Comparison Survey*
McCain	44	46
<i>Strongly</i>	18	19
Obama	44	44
<i>Strongly</i>	25	26
Other/DK	<u>12</u>	<u>10</u>
	100	100
<i>Democrats & Dem. leaning reg. voters</i>		
<i>Primary</i>		
Obama	43	51
Clinton	48	41
Other/DK	<u>10</u>	<u>8</u>
	100	100
<i>General</i>		
Obama	81	81
McCain	10	10
Other/DK	<u>9</u>	<u>9</u>
	100	100

*The mid-August comparison data were weighted to match the demographic composition of the obtained sample of hard-to-reach cases.

¹ It is important to note that respondents in this survey were not necessarily the same ones who refused or avoided the earlier effort to be interviewed, since data about refusals and non-contacts is necessarily collected at the household and not the individual level.

One area of clear difference between the hard-to-reach sample and the concurrent survey was in primary candidate support among Democratic and Democratic-leaning voters: in the hard-to-reach sample, Democratic voters were considerably more likely to have supported Hillary Clinton in their party's nominating contest. Clinton had a 48% to 43% lead among the hard-to-reach sample, while Obama had a 51% to 41% lead among the comparable August sample. If the analysis is limited to white Democrats and leaners, the magnitude of the difference is similar. These bivariate results were supported by a multivariate analysis that controlled for sex, age, education, region and, where appropriate, race and party (not shown). A logistic regression predicting the nomination preferences of white, non-Hispanic Democrats and Democratic-leaners found a strong and significant effect of being in the hard-to-reach sample on support for Hillary Clinton rather than Barack Obama. A similar regression analysis found no significant difference in general election preferences, either for all registered voters or for white Democrats and Democratic leaners. That differences are more apparent in the primary contest may suggest a greater willingness of racially conservative Democratic voters to report opposition to a black candidate without having to overcome party identification; put differently, a vote by a Democrat for a white candidate against a black candidate in an intra-party contest should be less stigmatizing or dissonant than a general election vote where the Democratic voter is presented with the choice of a white Republican candidate over a black Democratic candidate.

**Table 9
Racial Attitudes**

	Weighted Hard-to Reach	Weighted June '08
<i>We have gone too far in pushing equal rights</i>	%	%
Agree	43	34
Disagree	53	61
DK	<u>4</u>	<u>5</u>
	100	100
	Weighted Hard-to Reach	Weighted June '08
<i>It's all right for blacks and whites to date</i>	%	%
Agree	79	81
Disagree	15	14
DK	<u>6</u>	<u>5</u>
	100	100
	Weighted Hard-to Reach	Weighted Sept '06
<i>Immigrants...</i>	%	%
Strengthen our country	41	41
Are a burden on our country	42	41
Neither/Both/DK	<u>17</u>	<u>18</u>
	100	100

The hard to reach sample and the comparison surveys were weighted to the usual post-stratification demographic parameters from the Current Population Survey.

As with candidate preferences, we also found somewhat mixed results on racial attitudes. Hard-to-reach respondents were as likely as landline respondents in a June 2008 survey to say that it's all right for blacks and whites to date each other (79% in the weighted hard-to-reach survey vs. 81% in June). And like the landline respondents in a September 2006 Pew Research survey, hard-to-reach respondents were divided on whether immigrants strengthen the U.S. or are a burden on the country.

But hard-to-reach respondents were more likely than a June 2008 landline sample to agree with the statement “We have gone too far in pushing equal rights in this country.” About one-third (34%) of those in the June poll agreed with the statement; 43% in the weighted hard-to-reach sample agreed. The patterns among white Democrats and Democratic-leaning respondents were similar to the patterns among all respondents.

One finding consistent with previous research is that hard-to-reach respondents display less interpersonal trust (Keeter et al. 2000). Among the hard-to-reach, nearly six-in-ten (57%) said “you can’t be too careful” in dealing with people; 39% said most people can be trusted. In an October 2006 Pew Research survey, 50% said you can’t be too careful and 45% said most people can be trusted.

But on numerous other comparisons, we found the hard-to-reach sample and standard samples indistinguishable. The hard-to-reach differed little on satisfaction with national conditions, happiness with their personal lives, or political interest and engagement.

While the survey of reluctant households offers evidence of the potential for a bias, the magnitude of such a bias is likely to be quite small. Differences in nomination preferences of Democratic voters between the standard sample and the reluctant respondent sample were sizeable (a 15 percentage point difference in the margin). But there may be less here than meets the eye. It was not at all clear that all of these voters would fail to vote for Obama in the general election; indeed, reluctant respondents indicated that they *would* vote for him at rates comparable to Democratic voters in the standard comparison survey. Further, any potential bias from all of these possibly racially conservative voters abstaining or voting Republican would have been quite modest considering the relatively small size of this group.

4. DISCUSSION

Despite concerns about the growing problems facing polls and the special challenges of an historic election, most pre-election polling in 2008 performed quite well in forecasting the outcome of both the presidential election and statewide races for governor and senator. Sometimes polls yield the right results for the wrong reasons, but the fact that many kinds of polls in various races and places performed well strongly suggests that the underlying methodology of election polling is still robust.

In the general election, serious bias from the so-called Bradley Effect did not materialize. White voters’ support for Obama did not significantly vary with race of interview, and while our

survey of reluctant households offers evidence of the potential for a bias, the magnitude of such a bias is likely to be quite small. Though not a focus of the present study, presidential primary polls, though less accurate than general election polls, also showed no signs of a systematic bias, despite the additional challenges inherent in primary polling. Pro-Obama biases tended to be relatively modest in size and most of the errors that occurred were underestimates of Obama's performance.

Non-coverage bias due to cell phones is a growing problem and might affect the accuracy of polls in the future as the percentage of voters reachable only by cell phone climbs. Even at approximately 20%, the cell-only population was not sufficiently different from other voters to create a large bias in overall survey estimates once normal demographic weighting was applied. But a small bias was apparent and may grow as the size of cell-only population expands. A majority of cell-only voters are ages 30 and older, and demographically they differ more from their landline-accessible age cohorts than do the cell-only voters under age 30. Less clear is whether a similar bias exists with respect to the portion of the population that has both a landline and cell phone but depends mostly on the cell phone.

Finally we should take note of the fact that the 2008 election presented special challenges in identifying likely voters, one of the common problems facing election pollsters. Levels of voter engagement appeared to be extremely high throughout the campaign, and for much of the year Democrats were equally or more engaged than Republicans, an unusual circumstance. Moreover, Barack Obama, as an American of mixed racial background and one parent who was a Muslim, had no precedent among candidates for the nation's highest office. He was especially popular among young voters and African Americans, two groups with historically lower rates of voter turnout compared with older voters and whites. And adding to the novelty of 2008, it was forecast – correctly – that far more voters would vote by absentee ballot or early voting than had ever done so before. Despite these circumstances, pollsters' methods for identifying likely voters (Perry 1960; Perry 1979) were evidently adequate to the task, despite wide variations in approaches and methods used to do so (AAPOR 2009).

REFERENCES

- The American Association for Public Opinion Research (AAPOR). 2009. "An Evaluation of the Methodology of the 2008 Pre-Election Primary Polls." Retrieved May 12, 2009, from http://www.aapor.org/uploads/AAPOR_Press_Releases/AAPOR_Rept_of_the_ad_hoc_committee.pdf
- Blumberg, Stephen J. and Julian V. Luke. 2009. "Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, July-December 2008." National Center for Health Statistics. Retrieved May 12, 2009, from <http://cdc.gov/nchs/data/nhis/earlyrelease/wireless200905.htm>
- Blumberg, Stephen J., Julian V. Luke, Gestur Davidson, Michael E. Davern, Tzy-Chyi Yu, and Karen Soderberg. 2009. "Wireless Substitution: State-level Estimates From the National Health Interview Survey, January-December 2007." National Center for Health Statistics. Retrieved May 28, 2009, from <http://www.cdc.gov/nchs/data/nhsr/nhsr014.htm>
- Groves, Robert M. 1989. *Survey Errors and Survey Costs*. New York: Wiley.
- Guterbock, Thomas M., Steven E. Finkel and Marian J. Borg. 1991. "Race-of-Interviewer Effects in a Pre-Election Poll." *Public Opinion Quarterly* 55:313-330.
- Hopkins, Daniel J. 2008. "No More Wilder Effect, Never a Whitman Effect: When and Why Polls Mislead about Black and Female Candidates." *Journal of Politics*, in press. Manuscript available at <http://people.ig.harvard.edu/~dhopkins/wilder13.pdf>. Retrieved May 27, 2009.
- Hugick, Larry. 1990. "Polls during the Past Decade in Biracial Election Contests." Paper presented at the annual meeting of the American Association for Public Opinion Research, Lancaster, PA
- Keeter, Scott, Carolyn Miller, Andrew Kohut, Robert M. Groves, Stanley Presser. 2000. "Consequences of Reducing Nonresponse in a National Telephone Survey." *Public Opinion Quarterly* 64: 125-148
- Keeter, Scott and Nilanthi Samaranayake. 2006. "Can You Trust What Polls Say about Obama's Electoral Prospects?" Pew Research Center for the People & the Press. Retrieved May 28, 2009, from <http://pewresearch.org/pubs/408/can-you-trust-what-polls-say-about-obamas-electoral-prospects>
- Perry, Paul. 1960. "Election Survey Procedures of the Gallup Poll." *Public Opinion Quarterly* 24: 531-542.

Perry, Paul. 1979. "Certain Problems in Election Survey Methodology." *Public Opinion Quarterly*, 43: 312-325.

Pew Research Center for the People & the Press. 1998. "Possible Consequences of Non-Response for Pre-Election Surveys: Race and Reluctant Respondents." Retrieved May 9, 2009, from <http://people-press.org/report/89/possible-consequences-of-non-response-for-pre-election-surveys>

Pew Research Center for the People & the Press. 2008. "Calling Cell Phones in '08 Pre-Election Polls." Retrieved May 9, 2009, from <http://pewresearch.org/pubs/1061/cell-phones-election-polling>

APPENDIX

Results for the analysis of the impact of cell phone non-coverage are based on combined data from six surveys conducted from September 9 through November 1, 2008. The data are from telephone interviews conducted under the direction of Princeton Survey Research Associates International and Abt SRBI, Inc. among a nationwide sample of adults, 18 years of age or older. Across all six surveys, 10,430 were interviewed on a landline telephone and 3,460 on a cell phone, including 1,160 who had no landline telephone. This sample composition is based on a ratio of approximately three landline interviews to each cell phone interview. This ratio is based on an analysis that attempted to balance cost and fieldwork considerations as well as the resulting demographic composition of the sample and the need for a minimum number of cell only respondents in each survey. The landline and cell phone samples were provided by Survey Sampling International.

The full dual frame sample that includes all interviews from the landline and cell samples is first weighted to account for the fact that respondents with both landline and cell phones have a greater probability of being included in the sample. Respondents with only a landline or only a cell phone are given a weight of 2, respondents with both a landline and cell phone are given a weight of 1 and respondents whose phone status is unknown are given an average weight for the sample based on which phone we reached them on.

The full dual frame sample is then weighted using demographic parameters from the U.S. Census Bureau's Current Population Survey (March 2007 Annual Social and Economic Supplement). This analysis produces population parameters for age, education, race/ethnicity, region, and population density of households with adults 18 years of age or older, which are then compared with the sample characteristics to construct the weights. In addition to the demographic parameters, the sample is also weighted to match current patterns of telephone status and relative usage of landline and cell phones (for those with both), based on extrapolations from the July-December 2007 National Health Interview Survey (NHIS). The final weights are derived using an iterative technique that simultaneously balances the distributions of all weighting parameters.

For comparisons of the full dual frame design with other designs, the landline plus cell-only sample and the landline sample are each weighted using an iterative technique that matches gender, age, education, race/ethnicity, region, and population density to parameters from the March 2007 Annual Social and Economic Supplement to the CPS. The landline plus-cell only sample is also weighted to match current patterns of telephone status and usage, based on extrapolations from the 2007 National Health Interview Survey (NHIS). There is no first stage weight

	%
Cell phone only	17.8
Dual	66.6
All or almost all calls on cell	14.7
All/almost all/some calls on landline	51.9
Landline only	<u>15.6</u>
	100
Based on Pew Research Center extrapolations from NHIS data for July-Dec. 2007.	

	<u>Landline</u>	<u>Cell</u>
	%	%
June	22	23
Mid-September	21	21
Late-September	21	17
Early October	22	21
Mid October	20	19
Late October	24	24
Election Weekend	24	23

applied to either of these samples because all respondents are assumed to have only one chance of being included in the sample.

The significance tests comparing the estimates based on landline respondents with those based on the combined landline and cell respondents and the combined landline and cell phone-only respondents account for the overlap in the two samples due to the landline sample cases being included in the combined estimates. The tests were conducted using software that accommodates complex survey samples.

Response rates for the political surveys described here ranged from 17% to 24%.