LIKELY (AND UNLIKELY) VOTERS AND THE ASSESSMENT OF CAMPAIGN DYNAMICS

ROBERT S. ERIKSON
COSTAS PANAGOPOULOS
CHRISTOPHER WLEZIEN

Abstract   Only in recent years has the "likely voter" technology been extended to polls well in advance of an election. In the case of the 2000 U.S. presidential election, CNN/*USA Today*/Gallup tracking polls indicated considerable fluctuations in likely voter preferences, greater than among the larger pool of registered voters surveyed. This article explores how Gallup's likely voter model exaggerates the reported volatility of voter preferences during the campaign. Much of the reported variation in candidate preference reported by Gallup in that election is not due to actual voter shifts in preference but rather to changes in the composition of Gallup's likely voter pool. The findings highlight dangers of relying on samples of likely voters when polling well before Election Day.

Prologue

Following the first presidential debate of 2000 on October 3, candidate George W. Bush received a remarkable boost in the polls, a boost so large that it became the centerpiece of campaign coverage in the media. Of the many polls, the largest shift to Bush occurred in what was one of the most closely-watched indicators—the daily CNN/*USA Today*/Gallup tracking poll.[1] In the

ROBERT S. ERIKSON is a professor of political science at Columbia University. COSTAS PANAGOPOULOS is a Ph.D. candidate in the Department of Politics at New York University, where he is also executive director of the Political Campaign Management Program. CHRISTOPHER WLEZIEN is a Reader of Comparative Government and Fellow of Nuffield College, Oxford University. A previous version of this paper was presented at the Meeting of the World Association of Public Opinion Research Annual Conference, Prague, Czech Republic, September 17–19, 2003. The research is supported by grants from the National Science Foundation (SBR-9731308 and SBR-0112856). For useful and stimulating comments, we are most grateful to Robert Shapiro and the anonymous reviewers. Address correspondence to Robert S. Erikson; e-mail: rse14@columbia.edu.

1. For convenience, subsequent mentions of the CNN/*USA Today*/Gallup tracking poll often refer to it simply by its familiar "Gallup" label.

October 2–4 tracking poll, which centered around the October 3 debate (and reported on October 5), Bush lagged Al Gore by a seemingly formidable 51-40 margin among "likely voters." By October 5–7 (the next independent segment of the tracking poll), Bush surged ahead with a 49-41 point lead. Seemingly, in a matter of only three days, Bush gained 9 points and Gore lost 10 points, for a 19-point swing. With this example being the most extreme instance, the Gallup tracking poll attracted considerable attention throughout the 2000 campaign for its exceptional volatility. The extraordinary volatility of Gallup's 2000 tracking poll is well documented (Traugott 2001) and has drawn skepticism from veteran poll-watchers (e.g., Kohut 2001; Norris 2001).

Often suspected as responsible for the volatility in Gallup's tracking poll is Gallup's selective reporting of only the vote intentions of respondents deemed as "likely" voters, even early in the campaign. Gallup's screen for detecting "likely voters" is admittedly sensitive to respondent enthusiasm (Newport 2000a). Thus, some of the early-October surge for Bush may have represented a shift of suddenly energized Bush voters into the likely voter pool, while suddenly dispirited Gore voters moved out. We can see more than a hint of this from an examination of the shift among Gallup's larger pool of registered voters (likely plus unlikely voters) in early October 2000. For the same three-day period of the 19-point swing to Bush among likely voters, the swing among registered voters was "only" 10 points—from a Gore lead of 48-38 to a 43-43 percent tie. It follows that if likely voters surged twice as much for Bush as registered voters did, that unlikely voters must have been going the other way. Indeed, this is what happened. Our analysis of Gallup's data reveals a three-day swing of 11 points toward Gore among "unlikely" voters.[2] Among those registered to vote but seemingly too uninterested to vote, Gore's lead actually grew from 42-36 to 49-32 over the three-day period following the first debate.

How can we account for the disparate movements of likely and unlikely voters? Of course, it could be the case that inattentive voters liked what they saw in Gore's performance and surged to him, while more discerning attentive voters were persuaded by the media buzz that Gore's first debate performance was too heated. More plausible, however, is the rival hypothesis of the changing composition of likely and unlikely voters as Gore and Bush voters shifted their enthusiasm, thus inflating Gallup's report of a likely voter surge to Bush.

The General Problem

When predicting the vote, polling organizations must concern themselves with voter turnout, as they know that a large percentage of respondents are not going to actually vote. The problem, of course, is that the candidate preferences that nonvoters express in interviews may be different from those of

2. See below for a discussion of the methodology for separating likely and unlikely (registered) voters.

actual voters. Typically in the United States, nonvoters are more likely than voters to select the Democrat when asked to choose their preferred candidate. This makes some sort of voter screen essential. Almost always in the United States, pollsters first ask respondents whether they are registered to vote, and among the registered, they present a series of questions designed to separate voters from nonvoters. These screening questions involve whether the respondent has been electorally active in the past and knowledge of such matters as the location of his or her polling place. They also involve questions of political interest, as voters who are more excited about the campaign are more prone to vote. Based on their score on the screening instrument, registered respondents might be assigned a probability of voting, which is then used as a weight when tallying the projected vote. The more typical solution—used by Gallup and many other pollsters—is to divide registered respondents into two groups. Respondents who score beyond a specified cutoff are designated as "likely" voters, whose choices are then counted in the tally. The choices of those scoring below the cutoff are discarded (Asher 2001; Crespi 1988; Daves 2000).

This article presents no general quarrel with the goal of selecting likely voters when polling on the eve of an election. And it remains agnostic about the methodological details of doing so. The proof is in the results. Pollsters know that to maintain their credibility they must forecast accurately, and their actual record of accuracy is excellent (Traugott 2001).

The intent here is to highlight the problems with likely voter screens when they are applied weeks or months in advance of the election. While polling organizations once screened for likely voters only near the end of an election campaign, in recent years polling organizations have applied their likely voter technologies to polls well in advance of Election Day. For instance, in the 2000 presidential race, the Gallup Poll measured the opinions of likely voters throughout the campaign. Gallup chose likely voters not directly on whether respondents would turn out on Election Day, but rather on whether they would turn out for a snap election on the date of the interview. As Frank Newport, the Gallup Poll's editor-in-chief, explains the methodology and its purpose:

> It is important to remember that the results of preelection polls are not intended to predict how the election will turn out (with the exception of the very last poll conducted the weekend before the election, which usually is a good predictor of the actual election results). Instead, polls are conducted to indicate who would win "if the election were held today" (Newport 2000b).

The goal of this article is to explore some perplexing issues about the use of likely voter samples when polling weeks or months in advance of the election. Our argument explores a terrain that should be familiar to the polling community: On the one hand, there is good reason to identify likely voters on the grounds that registered respondents who are less likely to vote are disproportionately likely to express Democratic preferences. To ignore this frequent (but irregular) pattern is to overestimate support for Democratic candidates.

On the other hand, estimates of who may be likely to vote in the weeks and months prior to Election Day in large part reflect transient political interest on the day of the poll, which might have little bearing on voter interests on the day of the election. Likely voters early in the campaign do not necessarily represent likely voters on Election Day. Early likely voter samples might well represent the pool of potential voters sufficiently excited to vote if a snap election were to be called on the day of the poll. But these are not necessarily the same people motivated to vote on Election Day.

For this analysis, we make use of the CNN/*USA Today*/Gallup's daily tracking polls in the 2000 presidential campaign, using the individual-level data available from the Roper Center. (We are grateful to the Gallup organization for making these data available [via the Roper Center] in remarkably transparent fashion.) This allows us to reconstruct the polls' pools of registered voters and their two components of likely and unlikely voters with remarkable precision.

Analyzing the Fall Campaign: Data and Methodology

Labor Day typically signals the start of the fall campaign. In 2000, Americans celebrated Labor Day on September 4. Our analysis examines the Gallup tracking poll (from Labor Day forward) that monitored respondents' answers to the "trial-heat" presidential vote intention question: "If the election were held today. . . ."[3] We examine a total of 60, 3-day moving averages of sentiment between September 4–6 and November 2–4.[4]

Gallup's *reported* number of likely voters typically equals about four-fifths of its pool of registrants and a bit less than two-thirds of its total number of respondents. Likely voters are designated, however, to reflect half the electorate and half the (weighted) respondents. Gallup works down its likely voter screen, incorporating as likely those in the top half. In order to represent the 50 percent most likely among adults, some respondents scoring in the mid-range of likelihood receive fractional weights for the likely voter pool. In other words, registered voters are weighted at "1" as likely voters, others at "0" as unlikely, and a small third group weighted with a number in between. "Unlikely" voters consist of registered respondents who are given weights

---

3. Here is the full question wording: "Now, suppose that the presidential election were being held today, and it included Al Gore and Joe Lieberman as the Democratic candidates, George W. Bush and Dick Cheney as the Republican candidates, Pat Buchanan and Ezola Foster as the Reform Party candidates, and Ralph Nader and Winona LaDuke as the Green Party candidates. Would you vote for—[READ AND ROTATE] Al Gore and Joe Lieberman, the Democrats; George W. Bush and Dick Cheney, the Republicans; [READ AND ROTATE] Pat Buchanan and Ezola Foster, the Reform Party candidates; or Ralph Nader and Winona LaDuke, the Green Party candidates?"
4. After November 4, Gallup switched from three-day to two-day surveys. We ignore these in our analysis.

Table 1. Sizes of Three-Day Samples in 2000 CNN/*USA Today*/Gallup Tracking Poll, by type of Respondent

|	Minimum	Maximum	Mean
All respondents	1153	4186	1618
Registered[a]	864	3474	1263
With recorded vote preference	780	3266	1183
With major-party preference	783	3241	1103[b]
Likely voters[c]	563	2119	778[c]
Unlikely voters	216	1117	333

[a] Including voters who say they do not need to register to vote.
[b] Total is not exactly equal to the sum of likely voters and unlikely voters, due to the difference in mean representative weights of likely and unlikely voters.
[c] For likely voters, the mean is slightly less than the mean of Gallup's reported likely voter *N*s (833) due to Gallup's full counting of respondents weighted fractionally for the likely voter pool. See footnote 6.

"1 minus the likely voter weight." (Some respondents then are fractionally in both the likely and unlikely voter pools.)[5]

As our measure of the projected vote, throughout we assess only the relative strengths of Bush and Gore, ignoring Nader, Buchanan, and undecided responses. Thus, we summarize sentiment over a particular time segment for a particular group as simply "percentage for Bush" among those who say they will vote for Bush or Gore. Table 1 summarizes the data in terms of sample size measured variously for the 60, 3-day periods.[6]

Figures 1 and 2 display separately presidential choice over the fall campaign for likely, unlikely, and the full set of registered voters. Figure 1 shows the 60, 3-day moving averages that Gallup normally presents. Figure 2

5. Typically, Gallup includes as likely all respondents scoring 7 on its 7-point screen plus a fraction of those scoring at 6, where those scoring at 6 are given fractional weights so that the weighted sum of 6s and 7s equals half the sample. Also see footnote 6.
6. Gallup employs poststratification to weight its respondents for representativeness of the adult sample. (Each respondent may have a different weight for each of the three days for which he or she is part of a three-day survey.) In calculating group *N*s we employ Gallup's weight factor, with a further multiplier so that the weighted *N* for each subsample in each survey equals the actual number of cases for the group. For instance, the mean weight for major-party voters using Gallup's weights is .96, reflecting that people with a voting preference are easier to interview than those without. Thus, our *N*s for major-party registered voters are slightly larger (reflecting observed *N*s) than would be imputed from Gallup's occasional reports of preferences among registered voters. For likely voter samples, we norm Gallup's weights so that each survey's weighted sum of fully (not fractionally) likely voters equals its observed number of fully likely voters. Similarly, weights for voters coded as fully (not fractionally) unlikely are normed so that each three-day survey's weighted sum equals its observed number of fully unlikely voters. Voters weighted fractionally likely and fractionally unlikely may have different representative weights in the likely and unlikely pools. Their weights are a product of the representative weights, the appropriate fractional weight between 1 and 0 as a likely or unlikely voter, and the norming coefficient.
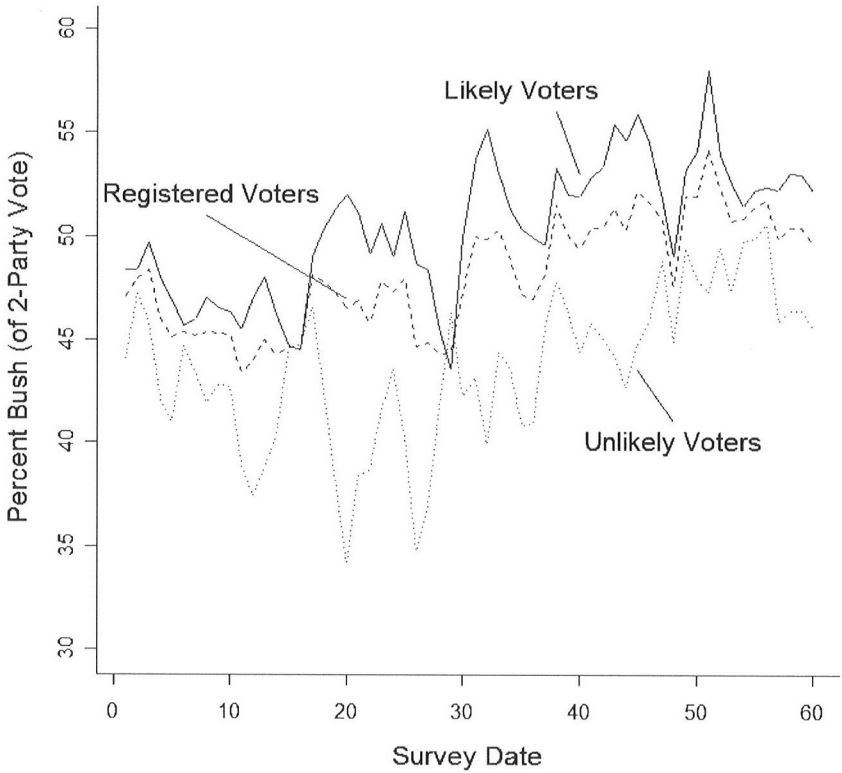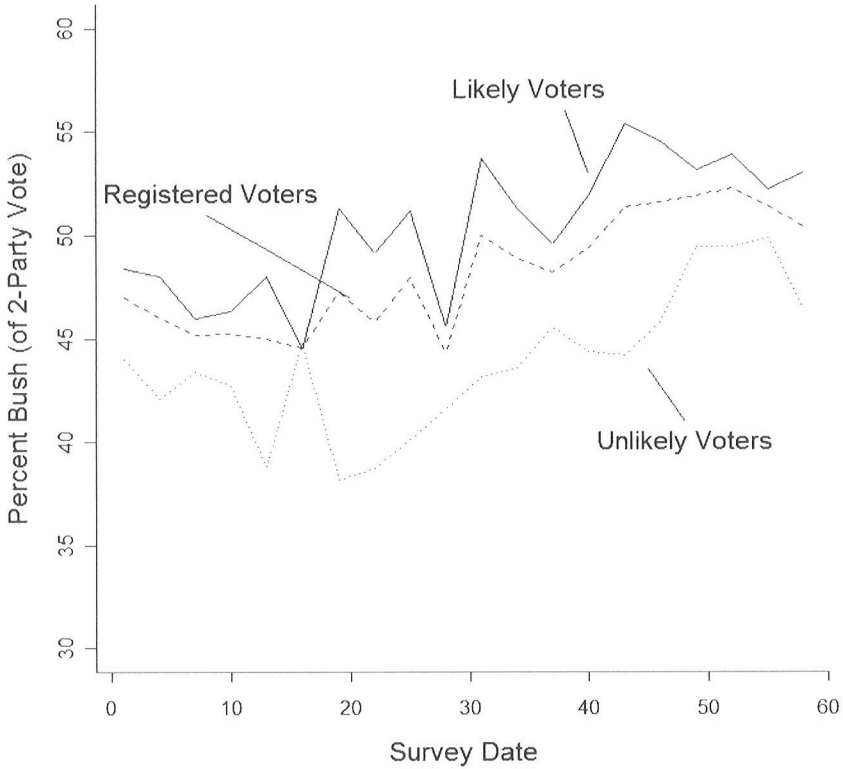
Figure 1. CNN/*USA Today*/Gallup three-day tracking polls, September 4–November 4, 2000: Daily readings for likely voters, unlikely voters, and registered voters. The 60 survey dates range from September 4, 5, 6 (1) to November 2, 3, 4 (60).

displays a series of 20, 3-day readings (taken every third day), which represent non-overlapping 3-day bands of time (i.e., September 4, 5, 6; September 7, 8, 9; September 10, 11, 12; etc., through November 2, 3, 4).

The figures demonstrate several points. First, likely voters were more favorable toward Bush than were unlikely voters. Clearly the likely voter screen resulted in estimates of the vote that were more Republican than if no screen other than voter registration were used. There are no surprises there.

Second, figures 1 and 2 show far more volatility among likely voters and unlikely voters than among the larger pool of registered voters, a pattern that persists even when we adjust for the different degrees of sampling error for the different groups due to their different sample sizes (see below.) Preferences

Figure 2. CNN/*USA Today*/Gallup three-day tracking polls, September 4–November 4, 2000: Every-third-day readings for likely voters, unlikely voters, and registered voters. The 60 survey dates range from September 4, 5, 6 (1) to November 2, 3, 4 (60).

are seemingly more stable for registered voters as a group than for the two component parts (likely and unlikely voters) treated separately.

Third, figures 1 and 2 reveal that *changes* in the preferences of the likely voters do not necessarily parallel changes in the unlikely voter samples. More often than not, when pro-Bush sentiment strengthens among likely voters over a three-day span, it weakens in the unlikely voter sample over the same span. In fact, changes in the preferences expressed by unlikely voters are *negatively* correlated with the changes among likely voters (Pearson's $r = -.27$)—and this is so even though some respondents have part of their weight in each of the two voter pools.

What can we make of this negative correlation? A normal expectation would be that the preferences of likely and unlikely voters to fluctuate in the

same direction. Whatever stimuli occur in the information environment ought, in principle, to either affect both voters and nonvoters positively or affect both negatively, although not necessarily to the same degree. Maybe voters are more (or less) sensitive than nonvoters to campaign stimuli, but at the very least one would not expect a negative correlation. At first glance, then, the observed negative correlation seemingly suggests a counterintuitive result: When attentive, interested (and therefore likely) voters are increasingly drawn to one candidate, their less attentive and less interested counterparts—so inattentive and uninterested as to display little interest in voting—shift their preferences in the opposite direction.

There is another possible explanation. The contrary likely and unlikely voter shifts could be due to the frequent short-term changes in relative partisan excitement, as changing enthusiasm of Democrats and Republicans generates movement in the Gallup monitoring instrument. At one time, Democratic voters may be excited and therefore appear more likely to vote than usual. The next period the Republicans may appear more excited and eager to vote. As Gallup's likely voter screen absorbs these signals of partisan energy, the party with the surging interest gains in the likely voter vote. As compensation, the party with sagging interest must decline in the likely voter totals.

Our analysis proceeds by first trying to answer the following question: How much of the observed variation in the polls is genuine daily variation in public sentiment rather than sampling error? We then inquire whether the true variation is larger in the pool of registered voters or the likely voter samples. To begin with, we compare the amount of variance in the polls with the amount that would be observed—given the survey Ns and random sampling—if public sentiment were constant (and evenly divided between Bush and Gore) throughout the duration of the fall campaign. The surplus variance is indicative of the true variance in public sentiment over the campaign (Erikson and Wlezien 1999).

True Variance

Imagine that the national division of presidential preferences was constant and unchanging from Labor Day until Election Day in 2000. The pollsters would observe changes in their results from poll to poll, with measurement error masquerading as campaign dynamics. Given a set of polls with frequencies and sample sizes from the 2000 election, we can impute the amount of variance that would be observed in poll results if there were indeed no real change. Assuming simple random sampling, we can compute this easily and determine how closely the sampling error variance approximates the observed variance in the actual polls. The surplus in the observed variance is then attributed to true change. We can conduct this exercise for our three types of voter

Table 2. Estimated Properties of CNN/*USA Today*/Gallup Tracking Poll Three-Day Samples of Voter Preferences, 2000

	Variance			Reliability (True Variance/ Observed Variance)
Sample	Observed (Total)	Error[a] (Sampling)	True (Observed minus Error)	
Registered voters	7.40	2.64	4.76	.64
Likely voters	10.37	3.67	6.70	.65
Unlikely voters	14.19	9.02	5.17	.36

NOTE.—For 60 observations of 3-day intervals of tracking polls.
[a] Estimated as the mean of .25/N where N is the N for the three-day period.

samples: registered, likely, and unlikely. Details of this exercise are shown in table 2.[7]

Theoretically, the sampling error variance in each of the post–Labor Day polls (given the daily Ns, simple random sampling, and no change) is $p(1-p)/N$, where p is the proportion voting for Bush over Gore. Substituting .5 for p, the average error variance of the percentage for Bush among likely voters is 3.67 for likely voters and 2.64 for registered voters. The difference between the two numbers (3.67 and 2.64) reflects the simple fact that likely voters on any given date comprise only about 72 percent of the registered voters. With more registered than likely respondents, the sampling error is least for registered voters. The standard error for likely voters, the square root of the error variance, equals 1.92, which translates into a daily confidence interval of plus-or-minus 3.8 percentage points around the observed percent support for Bush.

Next, we compute the observed variance in the post–Labor Day polls: 10.37 percent for likely and 7.40 for registered voters. These observed variances are about three times the error variances expected given simple random sampling and no dynamics. The imputed ratios (for both registered and likely voters) of true (not error) variance to observed (or total) variance are about 2 to 3. The precise estimates of the statistical reliabilities are .64 for registered voters and .65 for likely voters. Thus, whether monitoring likely or registered voters, almost two-thirds of the observed variance in reported daily preferences is true rather than (sampling) error variance.

The most interesting difference between likely and registered voter samples is the implied disparity in their true variances. Adjusted for sampling error, the variance of presidential preferences is 40 percent greater among likely voters

7. The assumption of simple random sampling is of course only an approximation. Inefficiencies due to departures from simple random sampling may be more than made up by efficiency in the poststratification weighting. All estimates are based on the full set of 60, 30-day surveys. The overlap of the surveys as 3-day moving averages does not pose a problem for this analysis.

Table 3. Estimated Properties of CNN/*USA Today*/Gallup Tracking Poll Three-Day Samples of Change (First Differences) in Voter Preferences, 2000

Sample	Variance			Reliability (True Variance/ Observed Variance)
	Observed (Total)	Error[a] (Sampling)	True (Observed minus Error)	
Registered voters	5.68	5.28	0.40	.07
Likely voters	12.68	7.34	5.34	.42
Unlikely voters	19.23	18.04	1.19	.06

Note.—For 57 observations of changes in 3-day intervals of tracking polls.
[a] Twice the estimate for the error in levels, from table 2.

than among registered voters. Thus, tracking poll readings for likely voters show greater volatility than for registered voters for two reasons: (1) the smaller N for likely voters generates more random error; and (2) likely voters seem to exhibit more real movement.

Also of interest are the statistics for unlikely voters. This group has the most sampling error variance, due to its low sample size. Adjusted for sampling error, the estimated true variance of preferences among unlikely voters is lower than for their likely voter counterparts, but slightly higher than for registered voters. The amount of volatility in the "unlikely" vote compared to the "likely" vote should trigger surprise, given that a major way that respondents get classified as unlikely is by showing signs of being inattentive and uninterested in the campaign.[8]

Further insights are gained from analyzing the variance of preferences as they change over time, that is, focusing on first differences of the poll readings rather than the levels. The details are shown in table 3. When using change scores, the error variance doubles because there is random error for both the "before" and the "after" measures. The result is to depress the reliabilities. Most observed change in preferences is random error rather than true change. For registered voters, the estimated reliability of reported change in percent Bush (from one three-day period to the next) is a miniscule .07, suggesting that virtually all change from poll to poll is error. But for likely voters, the estimated reliability of change scores reaches the relatively lofty level of .42, as if close to half the variance in observed change over three days is real change.

8. Of course, although unlikely voters are least likely to expose themselves to campaign information, as uninformed voters they may be more influenced by the information that does reach them (Zaller 1992).

For likely voters, the imputed true variance in first differences in preferences is 13 times larger than for registered voters. Even for unlikely voters, the imputed true variance in first differences is 3 times that for registered voters. What does this mean? Our interpretation again is that most of the true change in the percentage for Bush for likely voters is not change due to voter conversion from one side to the other but rather, simply, changes in group composition. To the extent that change scores are real, they are due to shifts in the type of respondents who score as "likely voters" or "unlikely voters" from one period to the next. One day, Bush voters score as more enthusiastic than usual about voting, and three days later the enthusiasm may be located among Gore voters, and so on. This is not a change in the preferences of the electorate who will be voting in November; it is only a short-term change in enthusiasm as scored by the likely voter instrument.

The proof for this argument is in the low estimate of the variance in true change scores for the registered electorate. Unlike "likely" and "unlikely" voters—who together comprise the registered electorate—the set of people who are registered to vote is essentially constant throughout the campaign. For registered voters, shifting composition is not an issue. For this compositionally stable group, almost all observed change is error.

If registered voters are stable, then how can there be much more real change among the two components—the "likelies" and the "unlikelies"? Logically, there is only one way. Trends within each group must cancel out—that is, correlate negatively with each other. Earlier we noted a modest negative correlation of −.27 between likely voter change and unlikely voter change. When corrected for reliability, the imputed true correlation in these change scores actually falls below its theoretical limit of −1.00. A possible correlation in the −1.00 range is only plausible if most change is compositional. For instance, when enthusiasm waxes for Bush voters and wanes for Gore voters, Bush gains among likely voters and Gore gains among the "unlikelies," even if there is no net change of preferences.

Stability of Preferences

Let us now examine the dynamics of measured preferences for the different groups of voters. We want to know how much preferences from one three-day period correlate with preferences in the preceding three-day period. More specifically, we want to know whether and to what extent these autocorrelations differ for registered and likely voters.

Our analysis shows that the preferences of registered voters are stable in a way that those of the ever-changing pool of likely voters are not. For registered voter samples, the raw correlation between the current and lagged percentages for Bush is .63. Although not high by itself, it almost matches the imputed reliability coefficient (.64) for registered voter

preferences.[9] Dividing the former (.63) by the latter (.64) yields the estimation of a true three-day correlation between current and lagged values of .98.[10] Thus, the readings for registered voters are perfectly consistent with a model of extremely stable preferences, as if the net presidential choice among registered voters changes very little over the short run of a few days. This result should not be treated as a surprise. Net voter preferences do change, but they change slowly (Wlezien and Erikson 2002).

For likely voters, the over-time correlation reveals a different pattern. Over the three-day interval of one period, the observed over-time correlation is a mere .41 while (from table 2) the reliability is .65. Dividing the former by the latter yields the estimation of a true three-day correlation between current and lagged values of .63. This represents considerable change over a three-day span—even after adjusting for sampling error, less than half of the variance ($.63^2 = .40$) in preferences can be explained by preferences three days earlier.

A mere .63 autocorrelation is not plausible for a constant pool of voters over a three-day span. Mass preferences do not churn this much. The instability of likely voter preferences must be due to composition effects—changes in which kinds of voters are deemed as "likely." As we saw, for the larger set of registered voters, which we know to be virtually stable in its composition, the imputed correlation approaches a perfect 1.0.[11] The results further support the conjecture that the shifting composition of likely (versus unlikely) voters is the major source of change in the 2000 Gallup tracking poll. Indeed, if the tracking poll had reported the preferences among registered voters, there would have been less evident change.

Likely Voting on Election Day

Let us assume that Gallup's (and other pollsters') classifications of their respondents as likely and unlikely voters are reasonably correct regarding the likelihood of voting if an election were to be held in the immediate future. If so, the best way to measure preferences if the election were held the day of the poll would be to poll only likely voters. But which type of sample (likely or registered) is most apt before Election Day? Does the bounce of excitement on the day of the poll have any bearing on predicting Election Day behavior? If the vote intentions of the registered voter sample remain the same on September

9. For our three-day polls we employ a lag of three days, with the *N* reducing from 60 to 57. Technically, we should take into account both the reliabilities of measured preferences at time *t* and time *t*-1, which differ slightly. Doing so makes very little difference.
10. For those who prefer regression coefficients to correlations, we note that (auto)regression coefficients predicting *t* values from *t*-1 values differ only trivially from the corresponding auto-correlations.
11. For unlikely voters, as for registered voters, the imputed reliability adjusted over-time three-day correlation of vote preferences approaches 1.00. It is estimated as a hefty .97. This estimate is fragile, however, as it is based on a ratio of two very low coefficients.

20 and 23—picking two arbitrary dates—but the likely voters' vote division changes, does this mean anything relevant to outcomes on Election Day, or is the short-term bubble extinguished quickly?

If, say, Republicans are more excited than usual on a given campaign date, will this translate even to the immediate future, let alone to Election Day? As a final test, we can readily measure the stability over time of the observed gap between the preferences of likely and unlikely vote. We can measure the gap as the Bush vote among likely voters minus the Bush vote among unlikely voters and observe the temporal stability of this indicator. The decisive answer is that variation in the likely-unlikely voter gap is temporary. Over three days—the time gap between independent samples—the autocorrelation of the likely-unlikely voter gap is actually negative at −.19. Observed differences in the preferences of likely and unlikely voters do not even last for three days. They can hardly be expected to carry over to Election Day.

Conclusion

When polling on the eve of an election, estimating which respondents are likely to vote is an essential aspect of the art. This article has pointed to dangers of relying on samples of likely voters when polling well before Election Day. Our evidence suggests that shifts in voter classification as likely or unlikely account for more observed change in the preferences of likely voters than do actual changes in voters' candidate preferences. Much of the change (certainly not all) recorded in the 2000 CNN/*USA Today*/Gallup tracking polls is an artifact of classification.

How much this critique extends beyond the Gallup tracking poll to all pollsters who screen for likely voters well in advance of Election Day is an open question. Monitoring the division between likely and unlikely voters during the campaign helps to avoid the partisan bias that would result from counting all registered voters equally. But by doing so, the danger is that pollsters mistake shifts in the excitement level of the two candidates' core supporters for real, lasting changes in preferences. Rather than trying to measure enthusiasm for voting at the moment of the poll, our advice to pollsters is to concentrate on advance estimation of respondents' likelihood of voting when the likelihood matters—on Election Day.

References

Asher, Herbert. 2001. *Polling and the Public: What Every Citizen Should Know*. Washington DC: CQ Press.
Crespi, Irving. 1988. *Pre-Election Polling: Sources of Accuracy and Error*. Nework: Russell Sage Foundation.

Daves, Robert P. 2000. "Who Will Vote? Ascertaining Likelihood to Vote and Modeling a Probable Electorate in Preelection Polls." In *Election Polls, the News Media, and Democracy,* ed. Paul J. Lavrakas and Michael W. Traugott, pp. 206–23. New York: Chatham House.

Erikson, Robert, and Christopher Wlezien. 1999. "Presidential Polls as a Time Series: The Case of 1996." *Public Opinion Quarterly* 63:163–77.

Kohut, Andrew. 2001. "Low Marks for the Polls, Media." *Columbia Journalism Review,* January/February, http://archives.cjr.org/year/01/1/kohut.asp (accessed August 2004).

Newport, Frank. 2000a. "How Do You Define Likely Voters?" http://www.gallup.com/content/login.aspx?ci=4636 (posted May 23, 2000; accessed August 22, 2004).

———. 2000b. "Questions Answered about Gallup's Presidential Election Tracking Poll." http://www.gallup.com/content/default.aspx?ci=4651&pg=1 (posted October 4, 2000; accessed August 22, 2004).

Norris, Pippa. 2001. "Too Close to Call: Opinion Polls in 2000." *Harvard International Journal of Press/Politics.* Available online at http://ksghome.harvard.edu/~.pnorris.shorenstein.ksg/acrobat/Campaign2000Editorial.pdf (accessed August 22, 2004).

Traugott, Michael W. 2001. "Assessing Poll Performance in the 2000 Campaign." *Public Opinion Quarterly* 65:389–419.

Wlezien, Christopher, and Robert Erikson. 2002. "The Timeline of Presidential Election Campaigns." *Journal of Politics* 64:969–93.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion.* Cambridge, UK: Cambridge University Press.