Daily Tracking Polls
Too much "noise" or revealed insights?

By David W. Moore

In 1996, several organizations conducted tracking polls for the last two months of the presidential campaign, providing in theory a more detailed look at the dynamics of the contest. Many critics at the time, however, argued that the daily results provided nothing more than "noise"—daily variations that were more a reflection of sampling error than of real change in the relative standings of the candidates. And, in a recent study of all the major pre-election polls from 1996, Erikson and Wlezien conclude that:

During the fall, when political activity and media attention were at their peaks, aggregate presidential preferences remained largely unchanged. To the extent that campaign events influenced the underlying division of preferences, the effects were small and short-lived. Thus, our findings are consistent with the interpretation that the electoral verdict is already in place before the general election campaign begins.[1]

It should be noted that these two authors do not argue that the polls were "wrong" or irrelevant, nor do they conclude that any specific tracking poll or series of polls by one organization overestimated variability in vote preferences. However, at least inferentially, their conclusions do lend credence to those who argue against the utility of tracking polls, suggesting that the daily updates supposedly about the standing of the candidates were really updates on the variation in sampling error.

What this view fails to take into account is a number of statistically significant fluctuations in the size of Bill Clinton's lead over Bob Dole which loosely followed actual campaign events. An analysis of the Gallup Poll trackings, conducted in partnership with CNN and *USA Today*, shows subtle changes in the dynamics of public opinion that would be obscured by aggregating the polling data into longer time periods.

O n September 2, 1996, the CNN/*USA Today*/Gallup Poll began the first phase of its daily tracking of the presidential race, which lasted through September 30. Nightly interviews were conducted with approximately 400 national adults, and data were reported based on three-day rolling averages. The second phase extended from October 1 through November 2 and included nightly interviews with approximately 600 national adults, with data reported on the basis of two-day rolling averages. The third phase, covering the last two days before the election, included interviews with approximately 1000 national adults each night that were combined to report the final pre-election vote estimate.

The data for every 3-day or 2-day period being reported were weighted to conform to US Census population figures for gender, age, region, education and race. Data were also weighted to account for the number of phone lines available in the respondents' households to receive calls. Once the data were weighted on the basis of all national adults interviewed, a separate weighting procedure was used to define likely voters. Results were reported on the basis of likely voters, an average of about 730 over each 3-day period in September and 2-day period in October.

Random-digit-dial telephone interviews were conducted, giving every household in America a nearly equal chance of being selected in the survey. A random selection grid was used to choose respondents within households, and a minimum of five call-backs were made to reach the person selected. New phone numbers were released into the CATI (computer-aided telephone interviewing) system each night to replace numbers which had been resolved, so that on any given night a roughly equal proportion of 1st, 2nd, 3rd, 4th, and 5th calls were made. Regional interviewing quotas were set each day to ensure that every one-day sample was geographically representative of the continental US adult population. The full interviewing period available each day was used in each time-zone to fill regional interviewing quotas.

> "An analysis of the Gallup Poll trackings shows subtle changes in the dynamics of public opinion that would be obscured by aggregating the polling data into longer time periods."

Gallup's final estimate of the vote, based on the last two days of polling and on the application of a "cut-off" likely voter model, was 52% for Clinton, 41% for Dole, and 7% for Perot. The actual national vote (excluding the vote for "other") was 50% for Clinton, 41% for Dole, and 9% for Perot. During the campaign, Gallup reported its figures based on a "probability" likely voter model, changing to the cut-off model for the final estimate. The likely voter figures analyzed here are those based on the probability model, which on average gave Clinton a 5-point greater lead than did the cut-off model. (The data from the final two days of polling are not included in this analysis.)

The data plotted in Figure 1 are based on a vote preference question that included the Perot/Choate ticket. For purposes of

Clinton vs. Dole

1996 CNN/*USA Today*/Gallup Tracking Poll

Figure 1

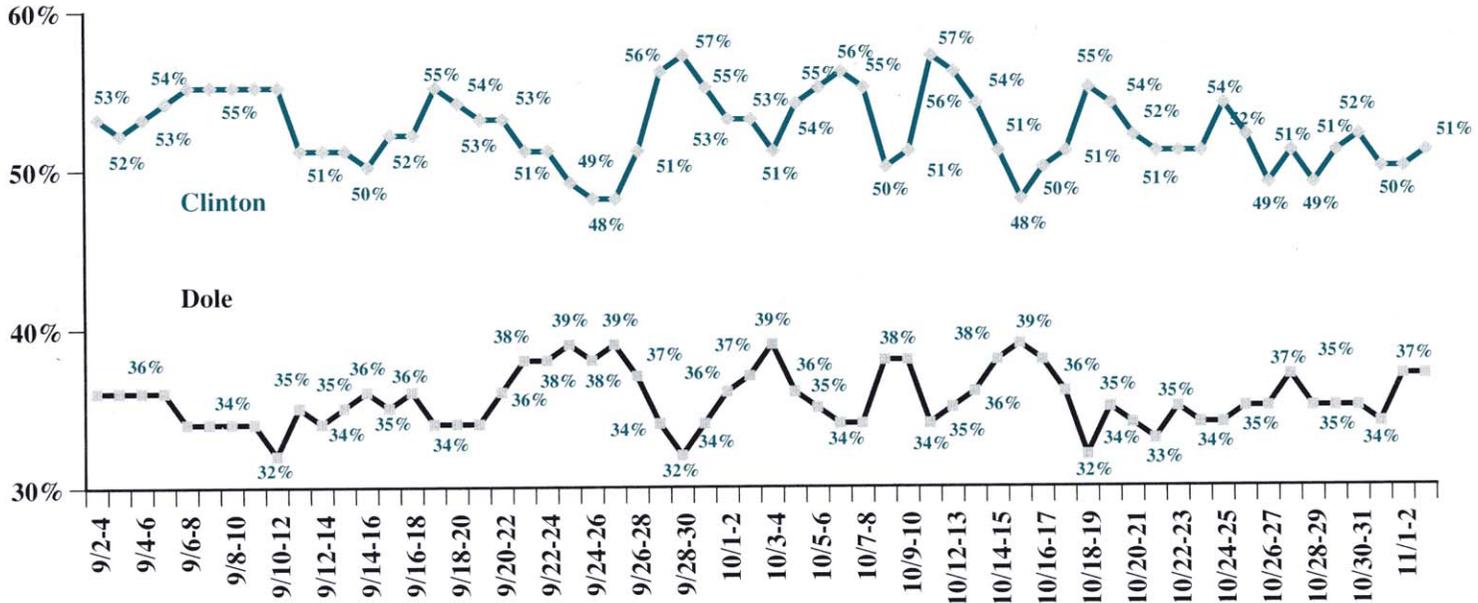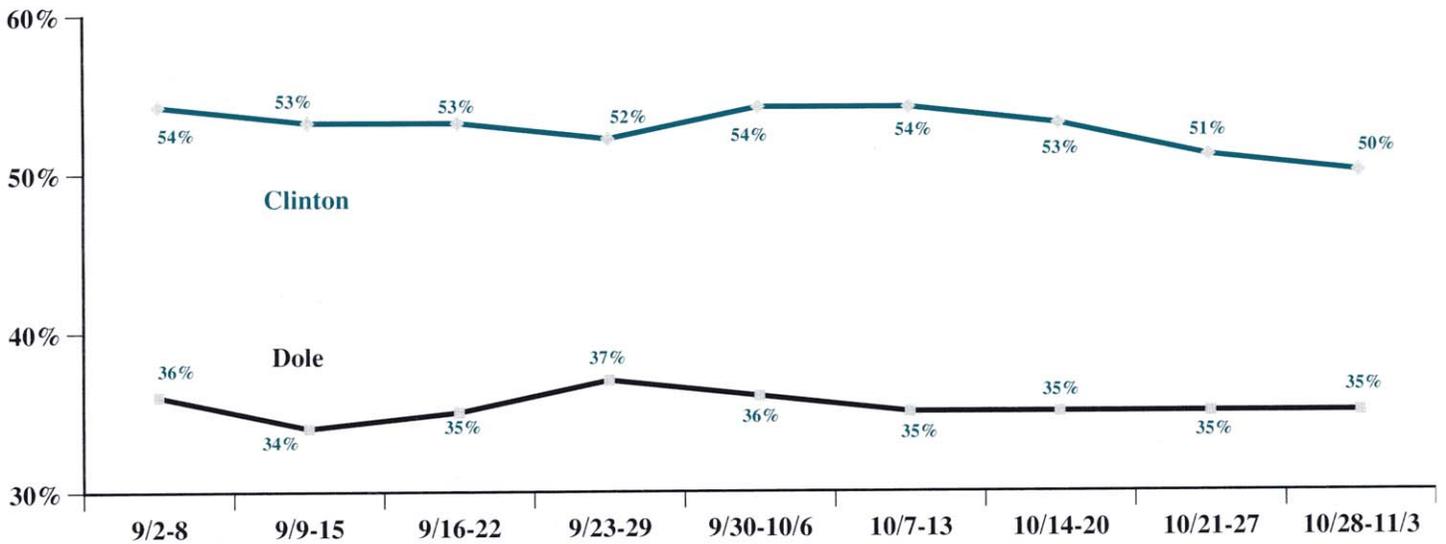## Using 3-Day Rolling Averages

Figure 2

Using Weekly Averages

display, however, only the figures for Clinton/Gore and Dole/Kemp are included. Note the volatility of the vote preference, with Clinton's lead ranging from a high of 25 points (September 28-30) to a low of 9 points (September 25-27 and October 15-16).

When the data are aggregated by week, a completely different picture of the contest emerges. As shown in Figure 2, Clinton's lead over the 9-week period barely varies from week to week—with a high of 19 points in the second and sixth weeks to 15 points in the fourth and ninth weeks. If a polling organization were to have conducted just one poll a week, lasting from Sunday through Saturday, these would have been the results—and they would have suggested that there was virtually no change in the standing of the two candidates.

The weekly aggregate of polling results suggests that very little happened during the campaign, that at least from Labor Day until the final two days before the election, voters consistently expressed substantially greater support for Clinton than Dole by about the same margin.

Another way to aggregate the tracking data is to use 3-day *discrete* (as opposed to *rolling*) averages. The rolling average is reported each day and includes the results averaged over the previous three days. Any given day will be included in three daily reports. Thus, for example, the October 4 results would be reported in the averages of October 2, 3 and 4; October 3, 4 and 5; and October 4, 5 and 6. With a discrete average, October 4 is included only with the results from October 2 to 4. October 5 and 6 are included with the next discrete time period, October 5 to 7. The rolling averages allow daily reporting, while the discrete averages allow reporting only every three days.

The advantage of the rolling average is that it is more sensitive to short-term changes than the discrete average. As the discrete *weekly* averages shown in Figure 2 demonstrate, whenever results are averaged over a several-day period, the day-to-day fluctuations may be lost. And this is true even for the discrete vs. rolling 3-day averages.

The results of rolling and discrete averages are shown in Figures 3 and 4, respectively. Aggregating the data in either manner suggests a much more complex race than shown by Figure 2, with at least six significant changes in Clinton's lead, and perhaps as many as thirteen changes or campaign phases overall. Either approach in aggregating the data suggests a much more dynamic campaign than that suggested by a weekly aggregate. These phases are also listed in Table 1, along with brief descriptions of the major events being reported in the news at the time of the polling.

(A cautionary note: the list of events in Table 1 is only suggestive of what could be driving the variance. Much of the

campaign was waged in selected parts of the country with the use of targeted TV and mail ads. While these mini ad campaigns could have had a major impact on the national figures measuring vote preference, they were not necessarily reported in the national press. Thus, the national news may not provide any insight into what might have affected public opinion at the times of some of the changes in Clinton's lead. The polling figures can still be used to track the fluctuations in public attitudes, however, even when the national news reports are not sufficient to provide the reasons for the changes.)

Contrary to the findings of Erikson and Wlezien, the analysis of the daily tracking polls conducted here shows there were indeed some major changes in the standing of the two candidates during the fall campaign. It is true that some of the changes were short-lived, but they were not necessarily small. And how long the changes lasted hardly seems relevant—the fact that there was significant volatility is by itself newsworthy.

The tracking poll clearly provides a useful description of the campaign *after the campaign is over*, and it can be used by scholars and observers to obtain a more detailed look at what occurred than is possible from the discrete 3-day or 5-day polls that are typically conducted. During the actual course of the campaign, however, interpreting the changes—even determining whether there *are* significant changes—can be difficult. Although there were several distinct phases to the campaign, some of the fluctuations in Clinton's lead within phases were indeed due to random error. Furthermore, changes that are statistically significant do not always appear that way at first, since a modest change may not be statistically significant until enough cases have accumulated—which means that a new lower (or higher) figure cannot necessarily be interpreted *at the time* as a significant change. This uncertainty is clearly a problem for news organizations.
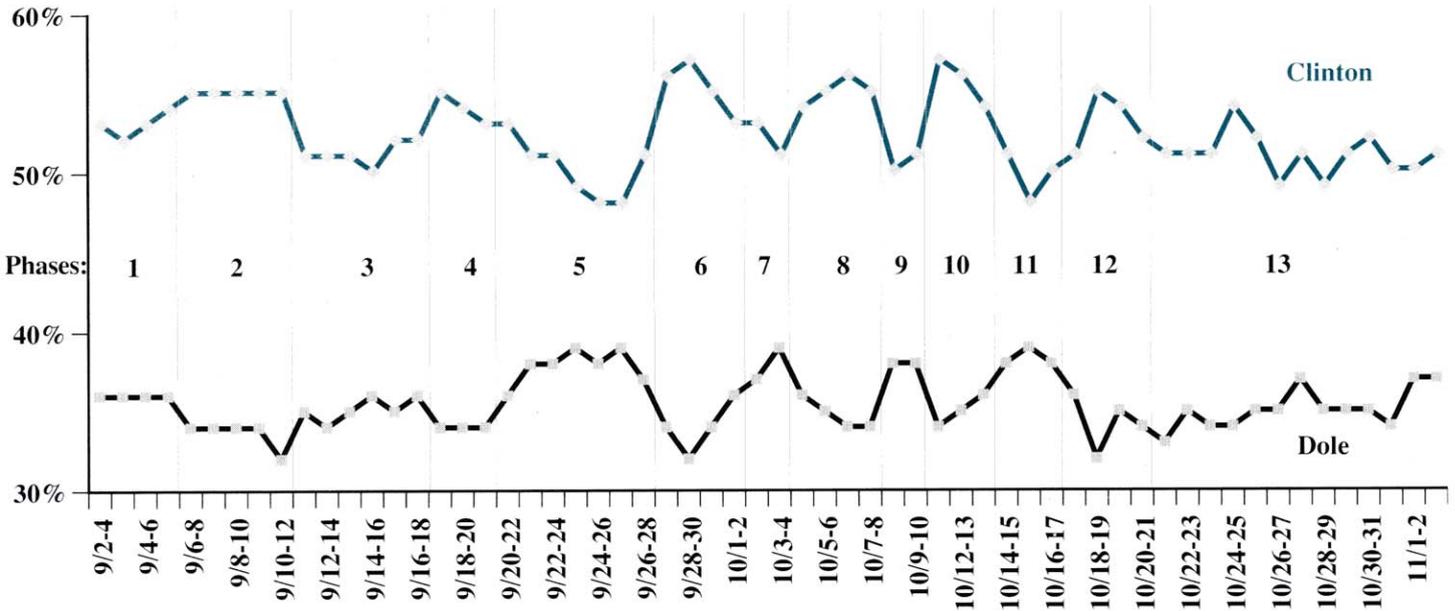
Even if there is a statistically significant change in poll results, the reason for that change is often not immediately obvious to the reporter or pundit. Much of the media advertising by the candidates is targeted to different parts of the country, and—as Dick Morris revealed in his account of his White House experience—often the ads are specifically *not* aired in the Washington market so reporters will be unaware of what the candidate is telling the voters.[2] Thus, reporters may be skeptical that the polls are truly reflecting changes in voter preferences, or at the very least they may be confused by the results, since there may be no evident explanation for them. The underlying causes may not be clear until a more thorough, post-election analysis of campaign strategies is conducted.

As a consequence of the greater fluctuations measured by the 3-day rolling averages, as well as the difficulty of analyzing the *results on a contemporaneous basis, it may be tempting to* conduct a series of 3-day or 5-day polls on a *discrete* basis. In

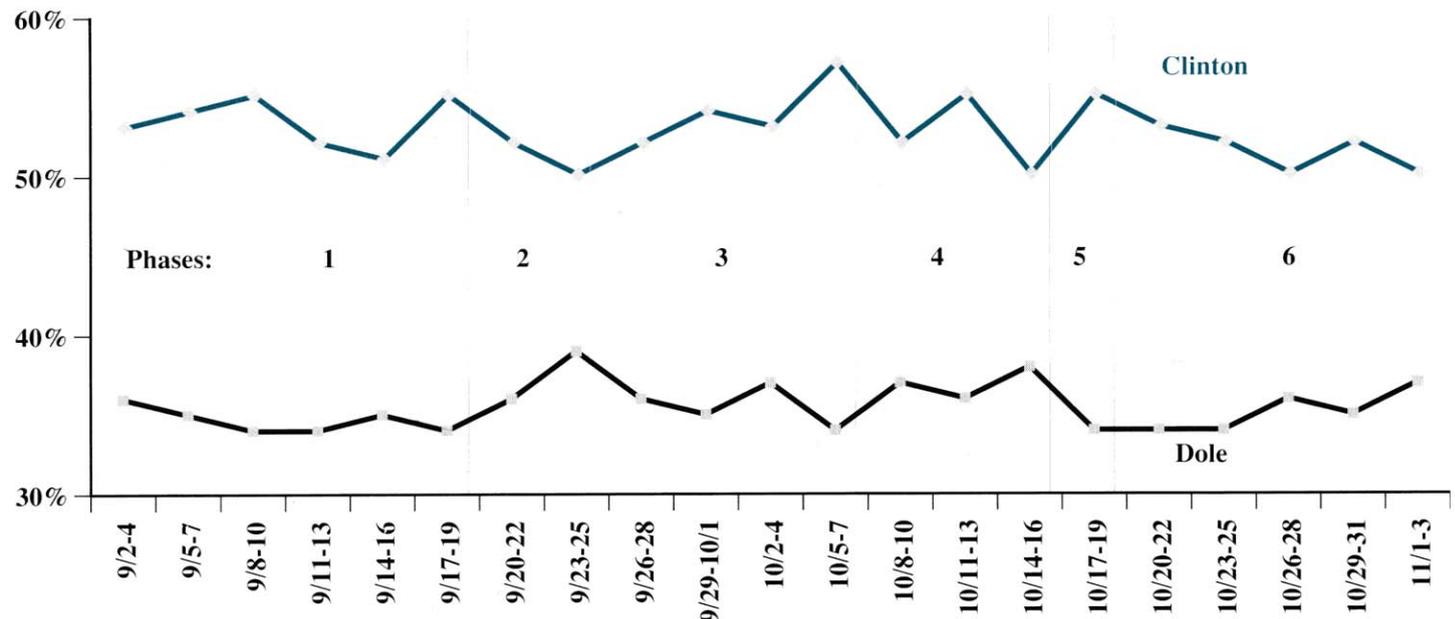# Potential Phases of the Presidential Campaign

Figure 3

Using 3-Day Rolling Averages

Clinton

Dole

Phases: 1   2   3   4   5   6   7   8   9   10   11   12   13

9/2-4 9/4-6 9/6-8 9/8-10 9/10-12 9/12-14 9/14-16 9/16-18 9/18-20 9/20-22 9/22-24 9/24-26 9/26-28 9/28-30 10/1-2 10/3-4 10/5-6 10/7-8 10/9-10 10/12-13 10/14-15 10/16-17 10/18-19 10/20-21 10/22-23 10/24-25 10/26-27 10/28-29 10/30-31 11/1-2

Note: New phase is determined when lead change is statistically significant.

Figure 4

Using 3-Day Discrete Averages

Clinton

Dole

Phases: 1   2   3   4   5   6

9/2-4 9/5-7 9/8-10 9/11-13 9/14-16 9/17-19 9/20-22 9/23-25 9/26-28 9/29-10/1 10/2-4 10/5-7 10/8-10 10/11-13 10/14-16 10/17-19 10/20-22 10/23-25 10/26-28 10/29-31 11/1-3

Note: New phase is determined when lead change is statistically significant.

Table 1

News Events That *Potentially* Influenced Clinton Lead

Daily Tracking Poll Phases	Discrete 3-Day Phases	Dates	Clinton Lead	Activities that May Be Related to the Statistically Significant Changes in Clinton's Lead
1	1	9/2-6	17 pts.	Clinton bounce from the Democratic Convention.
2		9/7-10	21 pts.	Bombing of Iraqi targets.
3		9/11-16	16 pts.	Iraqi fires at US jets; no US response.
4		9/17-20	20 pts.	Dole falls off stage; Perot denied debate participation (Clinton in favor of Perot participation; Dole against).
5	2	9/21-26	12 pts.	Dole ads attack Clinton on drugs; show Clinton joking on MTV about his first efforts to smoke marijuana.
6	3	9/27-30	24 pts.	Clinton's ads counter Dole's drug ads; Mideast summit in Washington.
7		10/1-4	16 pts.	Clinton announces failure of Mideast talks.
8		10/5-7	20 pts.	Clinton Rose Garden signing of bills fighting crime, keeping track of sex offenders; new ad linking Gingrich and Dole.
9	4	10/8-10	12 pts.	Reaction to first presidential debate.
10		10/11-13	21 pts.	Reaction to vice-presidential debate.
11		10/14-16	12 pts.	New attacks/ads by Dole on Clinton's character.
12	5	10/17-20	21 pts.	Reaction to second presidential debate.
13	6	10/21-25	16 pts.	New anti-Clinton ads focusing on character; new questions about DNC finances and foreign contributions.
		10/26-11/2	14 pts.	Dole exclusive focus on ethics and financing issues.

1996, several news organizations followed this approach. However, as Table 1 makes clear, such an approach may simplify the picture to the point where some of the most interesting campaign dynamics are not detected. And polls that are somewhat longer—such as ones yielding the weekly results analyzed here—may be even worse, so that no significant changes in the standing of the candidates are found, even when they are actually occurring.

There is no easy answer to the dilemma news organizations face when considering how to cover an election. The tracking polls maximize the ability to monitor the short-term changes in the campaign, but in the short run—during the actual campaign—the tracking polls may suggest a more volatile public than really exists, or may present results that are difficult to interpret at the time they are occurring. On the other hand, the discrete, multi-day polls tend to be slow in picking up valid changes in the campaign, and may even obscure altogether some important campaign dynamics. The latter approach may be "neater," but the former is likely to be more informative.

Endnotes

[1]Robert S. Erikson, and Christopher Wlezien, "Presidential Polls as a Time Series: The Case of 1996," *Public Opinion Quarterly*, forthcoming 1999.

[2]Dick Morris, *Behind the Oval Office* (New York: Random House, 1997).

David W. Moore is vice president, the Gallup Organization.