

**The Horse Race:  
What Polls Reveal as the Election Campaign Unfolds**

Christopher Wlezien  
Temple University

Robert S. Erikson  
Columbia University

*International Journal of Public Opinion Research, forthcoming*

An earlier version of this paper was presented at the Gallup Symposium on the Science of Pre-Election Polling, Washington, D.C., May, 2002. Portions also were previously presented at the University of Essex, Nuffield College, Trinity College, Dublin, and in the seminar series of the Houston Chapter of the American Statistical Association. The research described herein builds on a larger project that has been supported by grants from the National Science Foundation (SBR-9731308 and SBR-0112856) and the Institute for Social and Economic Research at Columbia University. Special thanks are due to Bruce Carroll, Joe Howard, Jeff May, and Amy Ware for assistance with data collection and management, David Firth, Charles Franklin, Don Green, Michael Marsh, Allan McCutcheon and the anonymous *IJPOR* reviewers for useful comments, along with participants in the Gallup Symposium, including Harold Clarke, Kathleen Frankovic, Simon Jackman, Edward Kaplan, Frank Newport, George Terhanian, and Michael Traugott.

## Abstract

The horse race of election campaigns is of great interest to scholars of public opinion and the public itself (Iyengar, Norpoth, and Hahn, 2004). In the modern day, in most every democratic country, we have a stream of polls registering electoral preferences during campaigns. In some countries, not a day goes by without a new poll release from one survey house or another. We thus have a lot of data at hand to assess the evolution of preferences over time. The problem is that the different survey houses use different methodologies and these differences matter for the results they report. In the presence of such “house effects,” how can we tell what is happening as the campaign unfolds? That is, how can we tell whether differences in poll results capture real change in electoral preferences and not differences in survey practices themselves? In this manuscript, the authors offer an approach for disentangling house effects from poll results during the course of the election campaign. Using the approach, we can effectively watch the horse race as it actually happens, which the authors demonstrate with pre-election polls from the 2000 US presidential race.

For over half a century, pollsters in many different countries have been asking samples of the public about their choices in the next election. The practice now is so common in some countries that hardly a day passes during the official campaign without encountering results of new polls, often from multiple organizations. In these so-called *trial-heat* polls, citizens typically are asked about how they would vote “if the election were held tomorrow.” Although we now have many polls, they give only imperfect readings of electoral preferences—poll results consist of true preferences plus survey error. Part of the error is simple sampling error, which is well-known. There also is error owing to the different methodologies that different survey organizations employ, however. Results may differ across survey houses due to differences in weighting samples, screening for likely voters, and other practices (see Groves, 1989). These effects not only are real, they appear to have increased over time (see Wlezien and Erikson, 2001). Given the evident differences across survey houses, how can we trace the evolution of electoral preferences over time? More critically, perhaps, how can we tell whether and how the horse race changes as the election campaign unfolds before us?

This challenge is the subject of the manuscript. The goal is to extract the underlying electoral preference over the election cycle using the full set of available polls, that is, in the presence of multiple survey organizations and designs. The approach builds on previous research in which poll results from various organizations were combined into one time series after the fact, to capture the history of voter preferences during the 1996 election (Erikson and Wlezien, 1999). We show that it is possible using the technology to at least partially disentangle preferences from poll results as the campaign actually evolves. In effect, we can watch the horse race “live,” as it happens. This is demonstrated here using polls from the U.S. presidential election of 2000. In a concluding section, we consider what polls do and don’t tell

us about preferences during the campaign and on Election Day itself.

### **Disentangling Preferences**

Trial-heat poll results represent a combination of true preferences and survey error. Survey error comes in many forms, the most basic of which is sampling error. All polls contain some degree of sampling error. Thus, changes are observed from poll to poll even when the division of candidate preferences is constant and unchanging. This is well known. The problem is that one cannot literally separate sampling error from reported preferences. Sampling error is random, after all.<sup>1</sup>

All survey results also reflect design effects. These represent the consequences of the departure in practice from simple random sampling that results from clustering, stratifying, and the like (Groves, 1989). When studying pre-election polls, the main source of design effects surrounds the polling universe. It is not easy to determine in advance who will vote on Election Day. All one can do is estimate the voting population. Pollsters attempt to ascertain each respondent's likelihood of voting based on such factors as interest in the election, self-reports of frequency of voting in the past, and the like. With this information about the voting likelihood of individual respondents, pollsters use two different approaches to obtaining likely voter samples. By the probability approach, pollsters weight individual respondents by their estimated probability of voting. Thus, the result of a poll is the sum of weighted votes from the survey. The cut-off approach is simpler. Instead of assigning gradations of voting likelihood, pollsters simply divide their registered respondents into two groups—the likely and unlikely voters. The result of the poll is the preference reported for the set of likely voters. Whether and how an organization constructs their likely voter sample can have meaningful consequences for the results, that is, to the extent polls overestimate participation of one

---

<sup>1</sup> It nevertheless does leave traces over time, which, if we make certain assumptions, allows us to make partial adjustments (Gerber, Green and Deboef, 1999).

segment of the population and underestimate another (see Traugott and Tucker, 1984).<sup>2</sup>

Survey organizations frequently weight respondents for other reasons using demographic variables and political ones too, such as party identification. Regardless of the specific approach used, weighting can have beneficial consequences for an analysis of polls. To the extent survey organizations weight by known population characteristics that predict the vote, such as certain census-based demographics, weighting will reduce the sampling error in poll results over time.<sup>3</sup> To the extent that they weight by characteristics that are not known with certainty, such as the distribution of party identification, weighting can introduce additional error.

When combining polls from different survey organizations, house effects are a problem. These represent the consequences of survey houses employing different methodologies, including design itself. Indeed, much of the observed difference across survey houses may reflect underlying differences in likely voter estimations and weighting, for example, when houses weight using different distributions of party identification. Results can differ across houses for other reasons, however, including data collection mode, interviewer training, procedures for coping with refusals, and the like (see Converse and Traugott, 1986; Lau, 1994; also see Crespi, 1988). Whatever the source, poll results will bounce around from day to day because polls reported on different days are conducted by different houses. This clearly complicates the assessment of electoral preferences.

We can formally summarize these effects of survey error. That is, the results of any poll  $P_i$  on a particular day  $t$  will be equal to the true electoral preference on that day ( $P_t^*$ ) plus any design ( $D$ ) and house ( $H$ ) effects:

$$P_{it} = P_t^* + D_j + H_k + e_{it}, \quad (1)$$

---

<sup>2</sup> It also may cause the polling universe to vary over time, which we consider below.

<sup>3</sup> This makes it more difficult based on the reported  $N$ 's to assess the amount of "true" poll variance and also to make adjustments for sampling error (see note 1).

where  $j$  designates different types of design and  $k$  different survey houses. The component  $e_{it}$  captures sampling error and other sources of error, including questionnaire design itself (especially see McDermott and Frankovic, 2003).<sup>4</sup> In our formulation, design effects shift the level of poll results but not their variance over time, though we do know that the latter is possible, if difficult to assess empirically. The difficulty is that we know only a little about the sampling procedures and weighting used by different survey organizations (Wlezien and Erikson, 2001). Although some organizations, such as Gallup, offer virtually full information, many of the others provide partial information and some only the bare details. This limits our analysis of design effects on the level of preferences, as we cannot take into account the specific differences in the polling universes across survey organizations.

While one cannot neatly control for differences in design, one can control for survey house, of course. Since design is determined by survey houses themselves, this largely solves our problem—by controlling for survey house we can capture *net* design and house effects. To do this, we simply estimate an equation of poll results including the survey date and house. We also can include a variable tapping the general polling universe, i.e., adult, registered, and likely. The coefficients for survey dates reflect estimates of public preferences during the campaign. This is what we used previously (Erikson and Wlezien, 1999) when assessing how preferences changed during the 1996 US presidential campaign. The approach can be applied after the fact to any set of polls where the numbers of survey organizations and survey dates are sufficiently large. The approach also can be applied during the course of campaigns themselves, as we will see. Let us begin with a basic application of the original method to pre-election polls from the 2000 presidential race.

---

<sup>4</sup> Questionnaire effects clearly are important but beyond the scope of this analysis, due to the lack of adequate measures of most of the differences that do appear to matter.

## The Polls in 2000

There were an incredible number of pre-election polls during the 2000 presidential election cycle in the US. For the 2000 election year itself, the [pollingreport.com](http://pollingreport.com) website contains some 524 national polls of the Bush-Gore (-Nader) division reported by different survey organizations. In each of these polls, respondents were asked about how they would vote “if the election were held today” with some differences in question wording. The differences typically matter little (see Lau, 1994) but there are circumstances in which they can (see McDermott and Frankovic, 2003). Of course, there are other differences in the practices of survey organizations and these are quite important, as we will see.

The 524 different polls are not all independent. That is, there is considerable overlap in many polls—typically tracking polls—conducted by the same survey houses for the same reporting organizations. Where an organization operates a tracking poll and reports 3-day moving averages on each day, for example, the result reported on day  $t$  contain information from days  $t-1$ ,  $t-2$ , and  $t-3$ . Clearly, it is necessary to remove this overlap in polls and doing so is straightforward. Using the example from above, where a survey organization operates a tracking poll and reports a 3-day moving average on each day, one simply selects poll results for every third day. Following this procedure from the beginning of the election year to Election Day leaves 295 separate national polls. Where multiple results for different universes are available for the same polling organizations and dates, data for the universe that is supposed to best approximate the actual voting electorate is used, e.g., a sample of likely voters over a sample of registered voters. Wherever possible, respondents who were undecided but leaned toward one of the candidates were included in the tallies.

-- Figure 1 about here --

Figure 1 displays the basic data. It shows Gore's percentage share of the two-party vote (ignoring Nader and Buchanan) for the set of 295 polls. Each observation represents the "vote" from a specific poll. Since most polls are conducted over multiple days, each poll is dated by the middle day of the period the survey is in the field.<sup>5</sup> The 295 polls allow readings for 173 separate days during 2000, 59 of which are during after Labor Day, the unofficial beginning of the general election campaign. We thus have a virtual day-to-day monitoring during that period. It is important to note that polls on successive days still are not truly independent: Although they do not share respondents, they do share overlapping *polling periods*. That is, polls results centered on a particular day  $t$  usually include information collected on adjacent days,  $t-1$  and  $t+1$ . Thus, results on neighboring days capture many of the same things. This is of consequence for any analysis of preference change.

### **What the Polls Reveal about Preferences**

Our poll results combine true preferences and the different expressions of survey error. We have shown that much of the error is due to the practices of survey houses, the net effect of which we can estimate directly. The 295 polls from 2000 were conducted by 36 different survey organizations. These organizations sampled three different polling universes during the year, specifically, adults, registered voters, and likely voters. As we have discussed, there is little information about the more specific differences in likely voter samples, the effects of which will be absorbed into our estimates of house effects. To explicitly control for survey house, results for the 295 polls are regressed on a set 35 (36 - 1) survey house dummy

---

<sup>5</sup> For surveys in the field an even number of days, the fractional midpoint is rounded up to the following day. There is a good amount of variance in the number of days surveys are in the field, ranging from 1 day snapshots to surveys that stretch up to 31 days. The mean number of days in the field is 3.57; the standard deviation is 2.39 days.

variables.<sup>6</sup> Dummy variables also are included for each of the 173 dates with at least one survey observation. (We include two polling universe dummy variables for “registered” and “likely” voters as well.) With no intercept, the coefficients for the survey date variables constitute our estimates of electoral preferences.<sup>7</sup>

-- Tables 1 and 2 about here --

The results of the analysis of variance are shown in Table 1. Here, we can see that the general polling universe did not affect poll results during 2000, controlling for survey house and date. This is not entirely surprising, as the same was true in 1996 (see Erikson and Wlezien, 1999).<sup>8</sup> Table 1 also shows that survey house did matter in 2000. Over the full year, the estimated range of the house effects is over six percentage points, pointing to important differences among at least some survey houses (also see Traugott, 2001). After Labor Day, house effects are even more statistically significant and much greater in size, with a range of approximately eight percentage points. Presumably these differences at least partly reflect differences in survey design, although we cannot be sure about the exact causes. Regardless, the evident differences have consequences for our portrait of preferences during 2000: Poll results will differ from day-to-day merely because different houses report on different days.

For illustrative purposes, Table 2 lists the estimated effects of selected, well-known houses over the full election year and after Labor Day. (Recall that our data include polls from 36 different survey organizations.) The Table displays the degree to which poll results from the different houses overstated or understated Gore’s poll share compared to the median house,

---

<sup>6</sup> To be clear, this controls for each house conducting at least one poll. The approach will tend to exaggerate the effects of houses that conduct only a small number of polls, say, one or two, as the estimates may reflect sampling error more than real house effects. Analysts thus might want to estimate only the effects of houses that have conducted a larger number of polls. It is worth noting that restricting our analysis to the nineteen houses that conducted at least five polls slightly shrinks the range of house effects and increases their reliability.

<sup>7</sup> For the analysis, the 269,974 respondents in the 295 polls comprise the units of analysis in the equation for the full election year, 130,024 respondents in the post-Labor Day equation.

<sup>8</sup> Again, this does not mean that the differences between the different likely voter universes did not matter.

which was ABC.<sup>9</sup> In Table 2 we can see that most of the organizations shown tended to overstate Gore's share *relative to the median house*, especially over the full election year. In other words, the mainstream polling organizations tended to overstate Gore's share by comparison with the many other organizations not shown in Table 2. The important lesson is that poll results differ meaningfully even among those houses that are within the mainstream, and in seemingly understandable ways (e.g., see Wlezien and Erikson, 2001). The differences also make clear why we cannot—or should not—rely on a single survey house to gauge change in voters' preferences.<sup>10</sup> Which one should we pick? Were we to pick Fox, the results would have overstated Gore's share by more than two percentage points on average compared to the base survey house. Were we to pick Voter.com, the results would have understated Gore's share by over one point. By relying on data from different houses, there is a basis for comparison, one that we can explicitly take into account in our analyses. It also gives us more poll readings and larger daily samples. While a benefit for the analyst, it can be confusing for the media consumer, as the results reported on each day and from day to day differ depending on which organizations happen to report.

-- Figures 2 and 3 about here --

Notice from our analysis of variance in Table 1 that the collective effects of the survey date easily meet conventional levels of statistical significance ( $p < .001$ ) during both the full election year and the general election campaign after Labor Day. This is statistical evidence that underlying electoral preferences changed over the course of the campaign. Figure 2 displays the survey date estimates, which constitute our (house-adjusted) estimates of preferences on different days. Of course, because of the differences across survey houses, it was necessary to select a "base" house. This is not a trivial matter obviously. For this

---

<sup>9</sup> Coincidentally, the median *poll* was conducted by ABC/Washington Post.

<sup>10</sup> The estimated house effects after Labor Day correlate quite nicely (Pearson's  $r = 0.79$ ) with the error of the final pre-election polls, using Traugott's (2001) summary.

analysis, the median house from the analysis of variance for the full election year was used. Though a seemingly appropriate decision, it may not be quite right, and the problem is that we cannot tell for sure. It may be tempting to use the house whose poll best predicted the final outcome, but this is even more tenuous. Given that all polls contain sampling error, *getting it right at the end is as much the result of good luck as good survey design.*<sup>11</sup>

Figure 3 displays results for the post-Labor Day period. There still is evidence of noise in these estimates, largely the result of sampling error. There also is reason to think that the series contains other sources of error that are not easily captured statistically because they vary over time, for instance, those associated with some likely voter estimators (Wlezien and Erikson, 2001). This might help explain some of the sharper spikes in poll results during the early fall. The problem is that there is little we can do. We must accept that our series of polls, even adjusted for systematic differences across survey houses, is imperfect.<sup>12</sup>

### **What Polls Reveal as the Campaign Unfolds**

The foregoing analysis is entirely historical. It tells us what happened using data that are available after the election is over. What do the polls tell us in advance of the election? That is, what do they reveal as the campaign itself unfolds? We obviously cannot step back in time to the 2000 election year. We can, however, behave as if we could. We can produce estimates at each point of the cycle using information about poll results up to that date. Specifically, extending the preceding analyses, house-adjusted estimates of preferences can be generated for each day of the election cycle using polls for the particular day and all preceding

---

<sup>11</sup> A subtler approach is to use the house that best predicts the final vote based on the estimated house effects and the date estimate for the last day of the campaign.

<sup>12</sup> It may be possible to at least partly address sampling error. Gerber et al. (1999) have developed a routine that models the negative autocorrelation in poll results that random sampling error will produce, i.e., falling in the wake of big rises and rising in the wake of big falls. Applying the procedure to our poll data makes only small differences, largely because of the relatively large samples with which we are working. Even the evident differences may be generous given that the procedure captures the effects of shocks to true preferences that are very short-lived, and look much like the effects of random sampling error, i.e., those campaign effects that are here today and gone tomorrow. Regardless, it confirms that our house-adjusted estimates are pretty reliable.

days. Doing so seems straightforward. The only question is: Which house do we select as our base? This is of real importance, as we have seen, for the decision affects the estimates of underlying electoral preferences in substantial ways.

After an election, the decision seems relatively easy: Pick the median house or perhaps the one that performed “best” (see note 11). But what do we do during the campaign? There would seem to be two primary criteria for deciding: (1) the past performance of survey houses; (2) their presence in the field throughout the cycle, from beginning to end. Based on these criteria, the obvious choice would seem to be Gallup. It has the longest continuous record of pre-election polling and has performed very well on average. Recent practices do raise concerns about its performance from day to day (Erikson, et al., 2004), though this is a different matter, reflecting on the variance—not the level—of performance. In 2000 and previous years, Gallup also was in the field throughout the campaign and polled much more frequently than any other organization.<sup>13</sup> Of course, we could change the base survey house as the campaign unfolds and we learn about the differences across houses. For instance, heading into the fall general election campaign, we might choose the median house based on analysis using polls from the pre-Labor Day period. For the purposes of our expository analysis here, we use Gallup as the base house throughout.<sup>14</sup>

To generate the house-adjusted estimates of electoral preferences at each point in time, the ANOVA model described in Table 1 is estimated for each date for which we have poll data. For instance, consider a point 200 days before Election Day (*ED*). On that day, *ED*-200, the effects of survey date, survey house and general poll universe are estimated using all poll results for that day and all preceding days, i.e., *ED*-200, *ED*-201, *ED*-202, and so on, using

---

<sup>13</sup> Fully 43 of the 295 separate national polls used in this analysis are from Gallup, 50 percent more than from any other organization.

<sup>14</sup> It is tempting to start off a year using the estimated house effects from the previous election cycle. While tempting, it presumes that the practices of organizations, especially as regards likely voter estimators and weighting strategies, do not change much from year to year. The problem is that they do change. New houses also

Gallup as the base house. The date estimate for *ED-200* is recorded and the process is repeated for day *ED-199* and each successive day for which there are poll data up to Election Day. These rolling house-adjusted estimates are shown in Figure 4 along with the original house-adjusted poll readings for the last 200 days of the campaign.

-- Figures 4 and 5 about here --

In the figure we can see fairly substantial disjuncture between the two series early on, until about 100 days before Election Day. This is to be expected given that house adjustments during this period are based on relatively small numbers of cumulative surveys, and thus are not highly reliable.<sup>15</sup> After that point, the two series closely track each other, with one notable exception about 70 days out. Indeed, through the post-Labor Day period, the rolling estimates virtually parallel the fully adjusted series based on data available after the election. This is clear in Figure 5. The pattern is satisfying: Although we expect the two series to increasingly correspond as we move closer to Election Day, we do not necessarily expect such a high level of consistency throughout the fall.<sup>16</sup> The rolling estimates do track slightly below the original house-adjusted readings, which is as we should expect given the smaller Gallup baseline. Recall that Gallup understated Gore's vote share by 0.42% on average compared with the median house during the fall (see Table 2).

The estimates in Figures 4 and 5 are based on polls aggregated by the mid-date of the reported polling period, e.g., the estimate for day *ED-30* is based on polls centered on that date. Of course, the data are not available at that time, when polls are still in the field. The data are only available when the polling is complete and there often is a lag in the actual release. To produce "current" estimates of preferences for each day, therefore, it is necessary to date polls by the release date itself. Unfortunately, this information is not readily available

---

can emerge, about which we have no prior information whatsoever.

<sup>15</sup> This is true of the final Election Day estimates for some houses. Also see note 6.

<sup>16</sup> The pattern implies that estimated house effects do not change much after Labor Day or that combining the often

now, after the election. We can, however, guess at the release date. For instance, we might assume that the release date is the date after the last day the poll is in the field. As we cannot be sure, it may make more sense to suppose that a poll is released in the two days after coming out of the field. This can be done by pooling data on each pair of days. Polls ending on *ED-30* can be dated as *ED-29* and *ED-28*, and polls ending on *ED-29* can be dated as *ED-28* and *ED-27*, and so on. The estimate for each day  $t$  thus would be the equivalent of an average forecast using polls ending on days  $t-1$  and  $t-2$ . This seems a reasonable approach given that we are forecasting after the fact, without information on poll release; in the future, we can do it right. To generate the estimate for each day, we simply follow the estimation algorithm used to produce the rolling estimates (in Figure 4) using the newly-dated polls. As before, Gallup is the base survey house.

-- Figure 6 about here --

Figure 6 plots the estimate and a 95% confidence interval for each date in the post-Labor Day period. The confidence interval is based on the daily sample size and the poll margin itself.<sup>17</sup> Notice that the forecasts are less variable than the fully house-adjusted estimates or the rolling estimates (see Figure 5), which is the simple result of pooling. The forecasts together with the confidence intervals reveal additional information about what we “could” see as the fall campaign evolved. First, given that the confidence drifts above the 50-50 threshold, it appears fairly certain that Gore entered the fall with a lead, which quickly disappeared. Second, it also appears that Bush really did gain the lead during the last 30 days, though his support ebbed and flowed somewhat over the period, i.e., on some days, we could not say with great (95%) confidence that he was “winning.” Third, in the days leading up to the election, we could forecast a dead heat only at the very last minute, on Election Day itself,

---

multiple readings on each day during the period largely cancels out the evident differences across houses.

<sup>17</sup> Since poll results on each pair of successive days are pooled, the sample size used to generate the confidence interval is the pooled sample divided by 2. Note that the confidence intervals technically are affected slightly by

and just barely at that. This implies one fairly obvious lesson: In a close race, stay in the field until the very end. The polls that did get it right.

## Conclusions and Discussion

Historically, pre-election polling was relatively simple. Survey organizations relied on samples of the registered voting population or its equivalent. While an imperfect approach, it nevertheless provided us a fairly reliable measure of preferences over a campaign. Polling organizations now rely on likely voter estimations and various weighting schemes, both of which have substantial design effects. Indeed, the differences among likely voter estimators appear to have much greater effects than the general differences between object universes, e.g., adults, registered voters, and likely voters. When we use polls from different survey organizations, therefore, results will vary from one day to the next partly because the universes of the reported polls themselves differ. As we have discussed, there are problems even when polls use the same design over time (see, e.g., Erikson, et al., 2004).

Unfortunately, we have limited information about what different survey organizations actually do to produce their likely voter results, though the information we do have is on the rise.<sup>18</sup> We also (usually) have little choice late in the campaign but to rely on results for likely voters—it is difficult and, in many cases, impossible to locate results for samples of registered voters after survey organizations put their likely voter estimators into effect. Although we cannot rely on the results from a single polling organization, at least with much confidence, we can exploit the information available in multiple polls, even in the presence of substantial house-related differences. Specifically, we can directly estimate the effects of survey houses and adjust the polls based on this analysis. This procedure provides a fairly clean time series

---

the loss of degrees of freedom from estimating house and date effects, and we do not take these into account.

<sup>18</sup> See Mark Blumenthal's summary of practices employed by some polling organizations in 2004, at <http://www.mysterypollster.com/>.

of aggregate preferences. Our analysis shows that we can do this after the fact and, with a bit of care, as the campaign itself unfolds

We can introduce other innovations. There are procedures to explicitly model the effects of sampling error, though these make only a small difference, at least given the relatively large sample sizes with which we are working. Other approaches are much closer to the data collection. For example, instead of aggregating results by the midpoint of the polling period, we can pool the results for all polls that span each particular day. If all polls spanned three days, for date  $t$ , we would pool three-day polls that end on date  $t$ , three-day polls that begin on date  $t$ , and three-day polls centered on date  $t$ . By pooling data centered on different days, we decrease the error variance, by definition. We also make the resulting series of preferences more dependent. The approach thus would be of less use for the purposes of studying campaign dynamics. For the purposes of election forecasting, however, such an approach would provide more reliable information.<sup>19</sup>

Of course, pre-election polls are not forecasts per se. They tap preferences at particular points in time and survey error too. What we have provided is a basic method for extracting elements of survey error from poll results reported by different survey organizations as the election cycle evolves. This may help us determine what the likely vote would be on any given day. It does not necessarily help us see what will happen at some point in the future. That is a subject for other research.

---

<sup>19</sup> There are other more powerful techniques, such as locally weighted scatterplot smoothing (lowess), which allows the analyst much more discretion, i.e., it is less strictly determined by the data. For more details, see Jacoby (1997). For an application to pre-election polls, see Erikson and Wlezien (1999).

## References

- Converse, P.E., and M.W. Traugott. (1986). "Assessing the Accuracy of Polls and Surveys." *Science* 234:1094-1098.
- Crespi, I. (1988). *Pre-Election Polling: Sources of Accuracy and Error*. New York: Russell Sage.
- Erikson, R. S., C. Panagopolous, and C. Wlezien, (2004). "Likely (and Unlikely) Voters and the assessment of Campaign Dynamics." *Public Opinion Quarterly* 68:588-601.
- Erikson, R.S., and C. Wlezien, (1999). "Presidential Polls as a Time Series: The Case of 1996." *Public Opinion Quarterly* 63:163-177.
- Green, D. P., A.S. Gerber, and S.L. DeBoef. (1999). "Tracking Opinion over Time." *Public Opinion Quarterly* 63: 178-192.
- Groves, R.M. (1989). *Survey Errors and Survey Costs*. New York: Wiley.
- Heise, D.R. (1969). "Separating Reliability and Stability in Test-Retest Correlations." *American Sociological Review* 34:93-101.
- Iyengar, S., H. Norpoth, K.S. Hahn, (2004). "Consumer Demand for Election News: The Horserace Sells." *Journal of Politics* 66:157-175.
- Jacoby, William. 1997. *Statistical Graphics for Univariate and Bivariate Data*. Thousand Oaks, Calif.: Sage Publications.
- Lau, R. (1994). "An Analysis of the Accuracy of 'Trial Heat' Polls During the 1992 Presidential Election." *Public Opinion Quarterly* 58:2-20.
- McDermott, M.L., and K.A. Frankovic. (2003). "The Polls—Review: Horserace Polling and Survey Method Effects: An Analysis of the 2000 Campaign." *Public Opinion Quarterly* 67:244-264.

Traugott, M.W. 2001. "Assessing Poll Performance in the 2000 Campaign."

*Public Opinion Quarterly* 63:389-419.

Traugott, M.W. and C. Tucker. 1984. "Strategies for Predicting Whether a Citizen Will Vote

and Estimation of Electoral Outcomes." *Public Opinion Quarterly* 48:330-343.

Wlezien, C., and R.S. Erikson. (2002). "The Timeline of Presidential Election Campaigns."

*Journal of Politics* 64:969-993.

----- (2001). "Campaign Effects in Theory and Practice." *American Politics Research*

29:419-437.

Figure 1: All Separate Trial-Heat Presidential Polls by Date, 2000

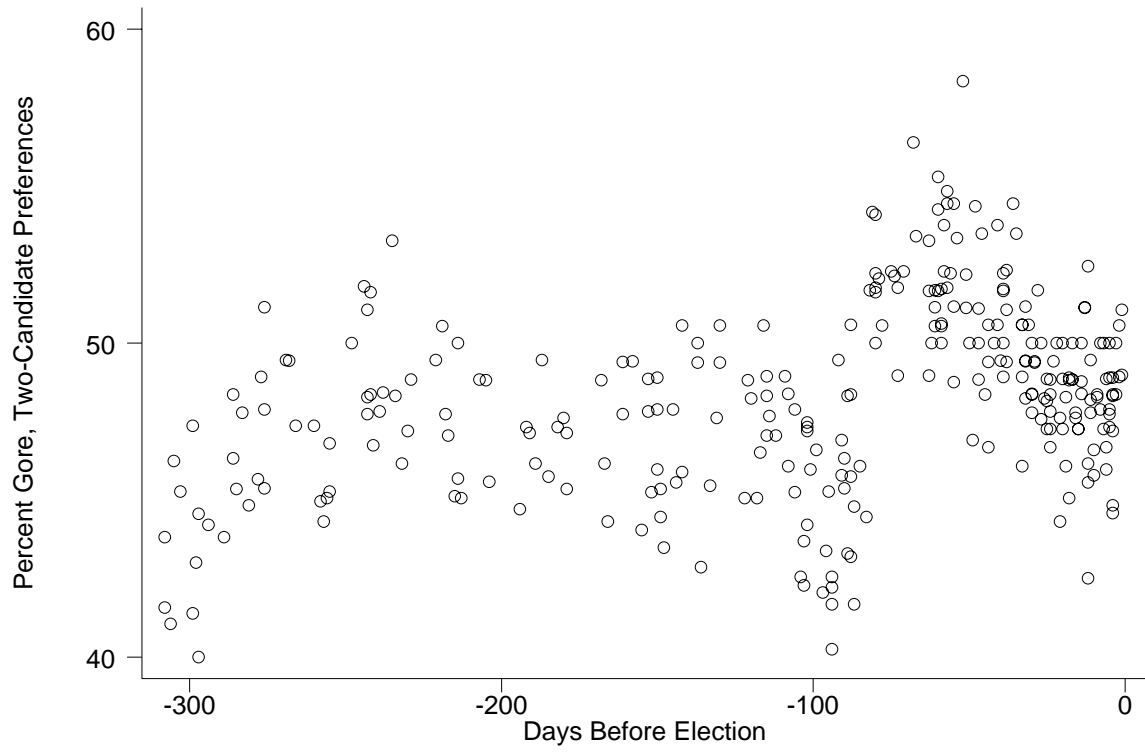


Table 1: An Analysis of General Survey Design and House Effects on  
 Presidential Election Polls, 2000

Variable	Election Year	After Labor Day
Poll Universe	0.12 (0.89)	0.08 (0.78)
Survey House	1.44 (0.09)	2.15 (0.01)
Survey Date	3.88 (0.00)	4.31 (0.00)
<i>R</i> -squared	0.90	0.88
Adjusted <i>R</i> -squared	0.68	0.68
Mean Squared Error	2.79	1.74
Number of polls	295	135
Number of respondents	267,974	130,024

Note: The numbers corresponding to the variables are F-statistics. The numbers in parentheses are two-tailed *p*-values.

Table 2: Estimated Effects of Selected Survey Houses on  
 Trial-Heat Poll Results, 2000 (in percentage points)

---

Polling Organization	Election Year	After Labor Day
Yankelovich/CNN/Time	3.01	3.10
Fox/Opinion Dynamics	2.07	2.62
Pew Research Center	1.26	0.57
CBS	1.18	1.46
Newsweek	1.15	1.76
NBC/Wall Street Journal	0.78	-.12
ABC/Washington Post <sup>a</sup>	0.71	-.19
Reuters/MSNBC	0.69	0.70
ICR <sup>b</sup>	0.64	1.14
ABC <sup>c</sup>	0.00	0.00
CNN/USA Today/Gallup	-.53	-.42
Voter.com/Battleground	-1.12	-1.59

---

<sup>a</sup> Prior to separate ABC and Washington Post tracking poll reports.

<sup>b</sup> International Communications Research.

<sup>c</sup> Based only on tracking poll reports.

Note: Positive numbers indicate a larger share for Gore relative to the median house, negative numbers a larger Bush share.

Source: Pollingreport.com.

Figure 2: House-Adjusted Trial-Heat Presidential Polls by Date, 2000

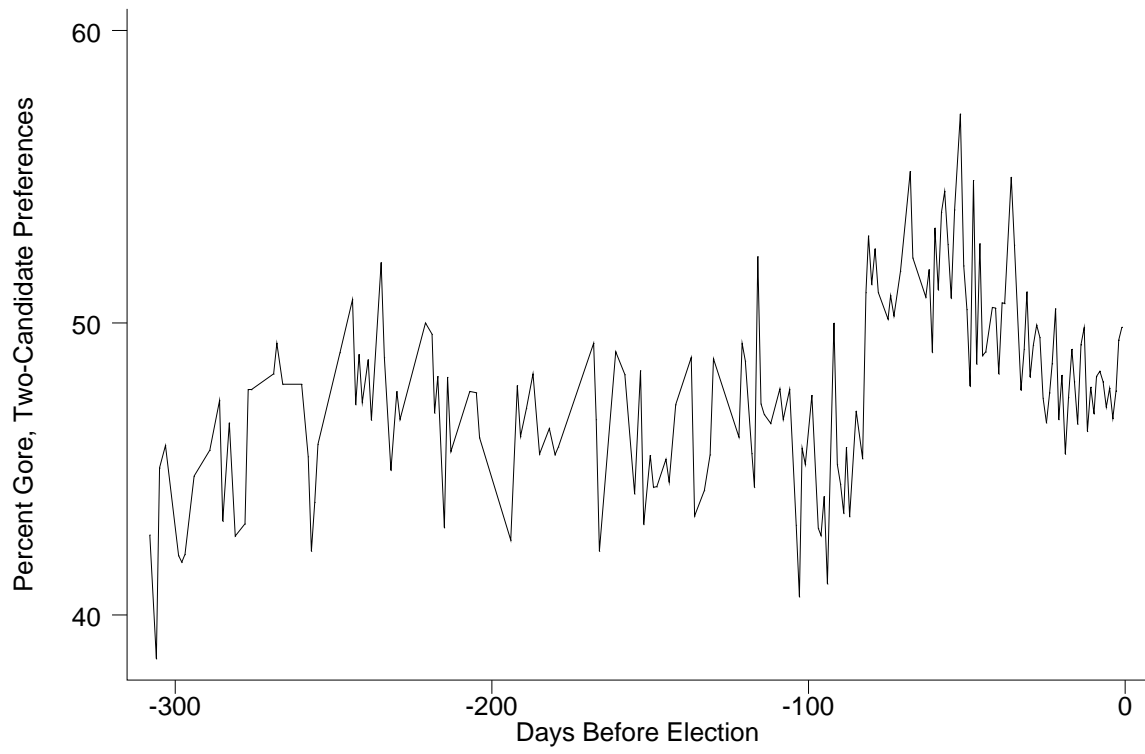


Figure 3: House-Adjusted Trial-Heat Presidential Polls by Date,  
Labor Day to Election Day, 2000

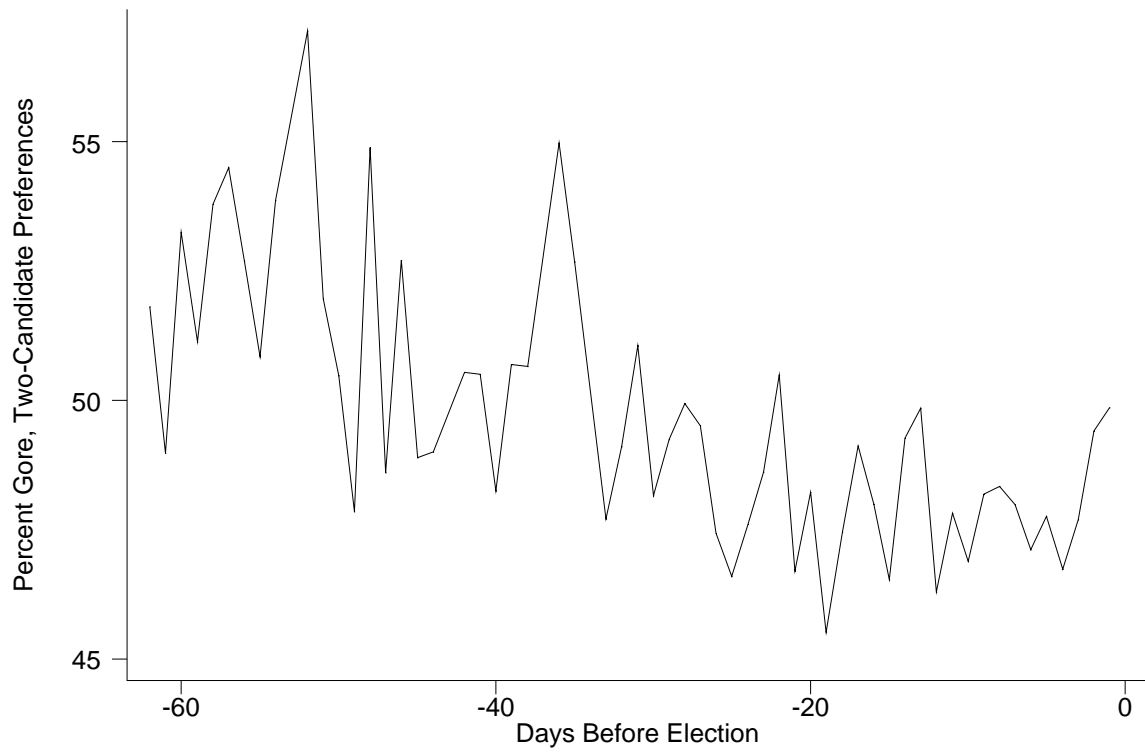


Figure 4: House-Adjusted Poll Readings and Rolling House-Adjusted Estimates, 2000

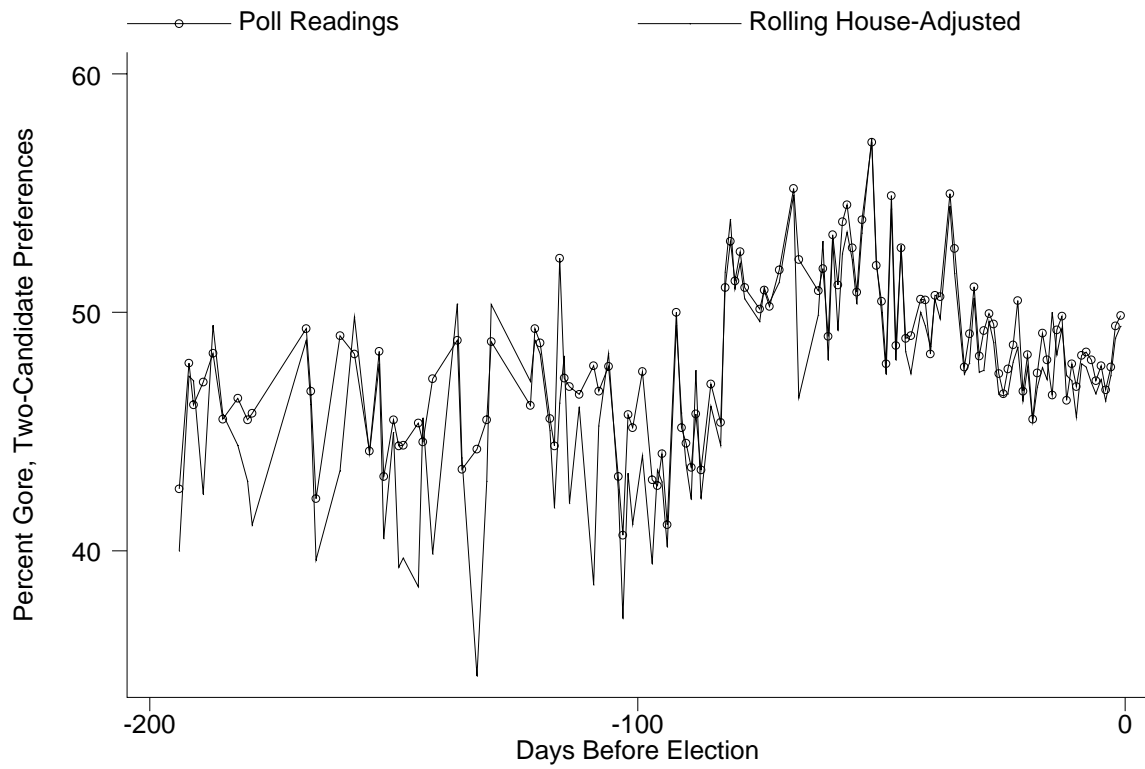


Figure 5: House-Adjusted Poll Readings and Rolling House-Adjusted Estimates, Labor Day to Election Day, 2000

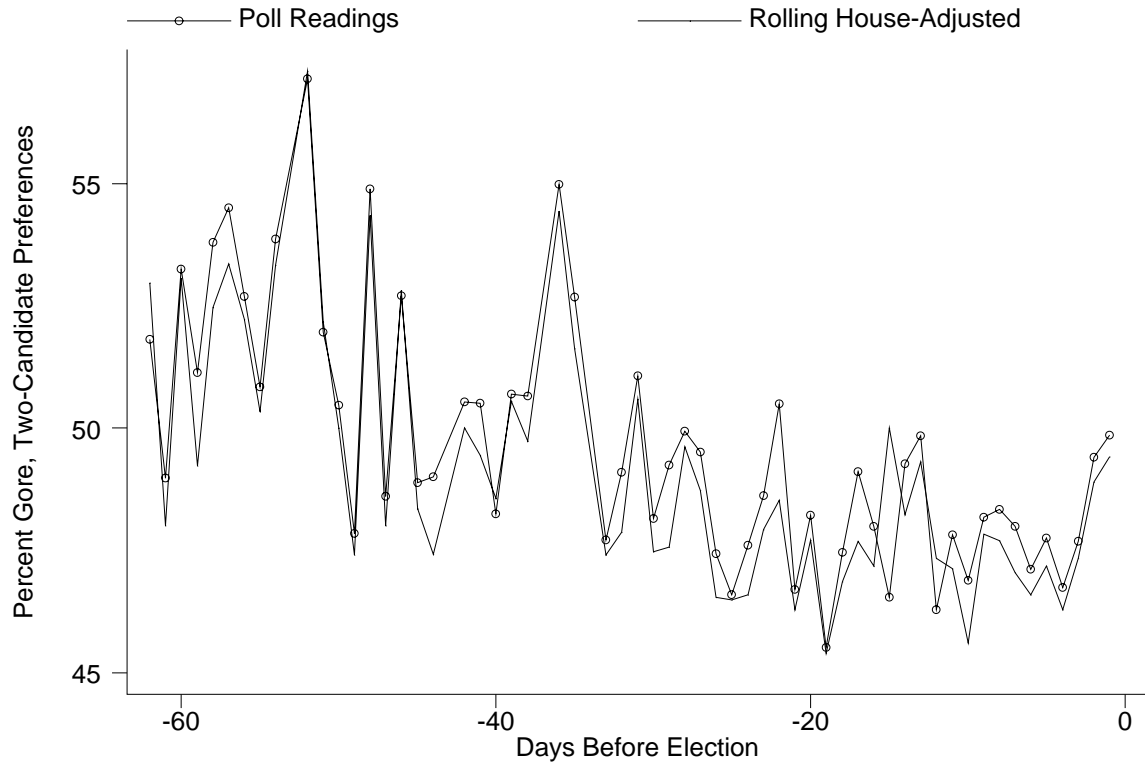
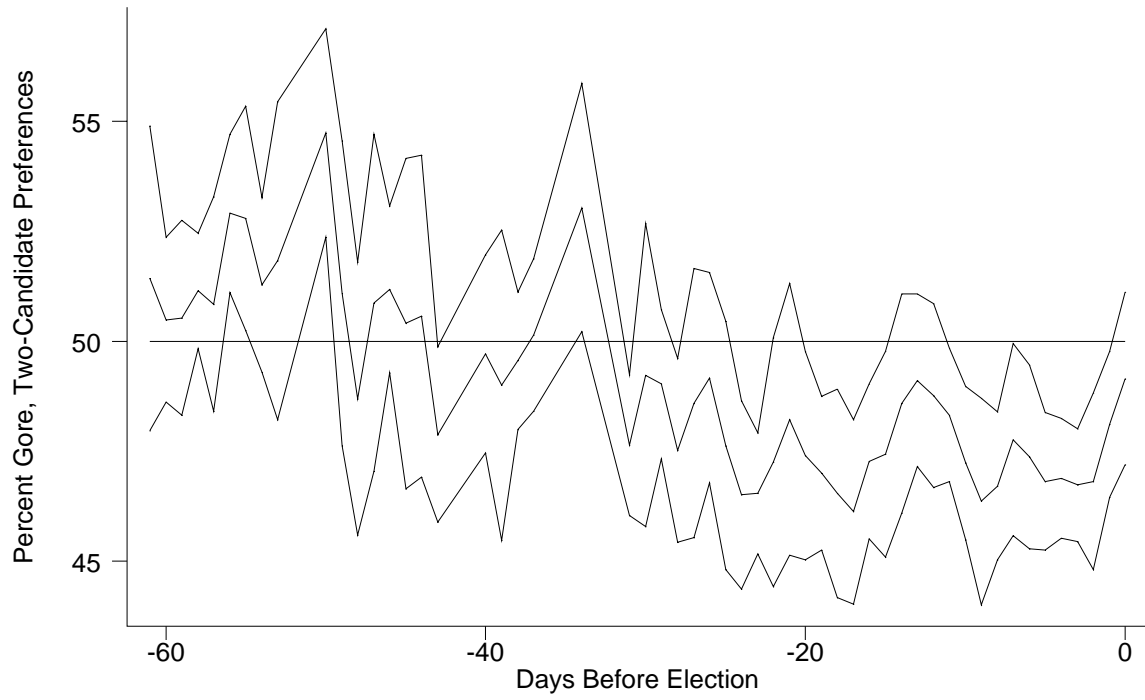


Figure 6: “Current” Rolling House-Adjusted Estimates (with 95% Confidence Intervals), Labor Day to Election Day, 2000



## Biographical Notes

Christopher Wlezien is professor of political science at Temple University. He recently co-edited (with Pippa Norris) a book on the 2005 UK election entitled *Britain Votes* (Oxford University Press).

Robert S. Erikson is professor of political science at Columbia University. He is coauthor of *The Macro Polity* (Cambridge University Press), *Statehouse Democracy* (Cambridge University Press), and *American Public Opinion* (Longman).

Address correspondence to Christopher Wlezien, Department of Political Science, Temple University, Philadelphia, Pennsylvania, 19122-6089 USA. E-mail: [Wlezien@temple.edu](mailto:Wlezien@temple.edu)