What Does "Do Campaigns Matter?" Mean?

Joseph Bafumi[†]
Andrew Gelman[‡]
David K. Park[‖]

January 26, 2004

Abstract
Scholars disagree over the extent to which presidential campaigns activate predispositions in voters or create vote preferences that could not be predicted. When campaign related information flows activate predispositions, election results are largely predetermined given balanced resources. They can be accurately forecast well before a campaign has run its course. Alternatively, campaigns may change vote outcomes beyond forcing predispositions to some equilibrium level. We find most evidence for the former: opinion poll data are consistent with Presidential campaigns activating predispositions, with fundamental variables increasing in importance as a presidential election draws near.

Keywords
Campaign effects, Enlightened preferences, Mean reversion, Presidential elections, Random Walk

---

We thank Robert S. Erikson for helpful comments and the National Science Foundation for financial support.
[†]Department of Political Science, Columbia University, jb878@columbia.edu
[‡]Department of Statistics and Department of Political Science, Columbia University, gelman@stat.columbia.edu
[‖]Department of Political Science, Washington University, dpark@artsci.wustl.edu

Introduction

The debate over how presidential campaigns influence votes remains unresolved. When campaign related information flows activate predispositions, election results are largely predetermined given balanced resources. They can be accurately forecast well before a campaign has run its course. Alternatively, campaigns may change vote outcomes beyond forcing predispositions to some equilibrium level. A recent crop of scholars has sought to tackle this enduring question. Anderson, Tilley and Heath (2003) go beyond past work (e.g. Gelman and King, 1993) by demonstrating the effectiveness of fundamental variables in predicting the vote at both the individual and the aggregate level. They find voting patterns to be predictable despite the variability in polls leading up to Election Day, with campaigns working to move voters toward their personal equilibrium vote choices.

From a different perspective, Hillygus and Jackman (2003) employ a new source of panel data to gain a deeper understanding of the relationship between campaign effects and public opinion. They claim to find evidence of both disequilibrating campaign effects and activation of predispositions where fundamentals explain vote choice. They label the former effect as "persuasion." However, a voter may be persuaded toward or away from their expected vote by campaign related information flows. If voters are persuaded toward their predicted equilibrium attitudes, then Hillygus and Jackman's results are consistent with the enlightened-preferences hypotheses of Gelman and King and Anderson et al. The findings of Hillygus and Jackman do not clearly point to campaign effects beyond activation of predispositions. Thus, it is not clear that they move beyond the conventional wisdom of minimal (disequilibrating) campaign effects.

In this note, we provide two analyses to determine the explanatory power of activation on the vote. First, we fit a series of logistic regressions predicting vote choice based on the fundamentals in the months leading to a presidential election. Such a model has pointed to evidence of activation in past presidential election years (Gelman and King, 1993). We provide further evidence by applying the model to the more recent 2000 presidential election. Second, we offer a new algorithm for investigating the predictive power of the fundamentals free from disequilibrating campaign effects. Specifically, we use the coefficients from a model predicting the vote by the fundamentals in 1996 to predict the vote over the course of the 2000 presidential election campaign season.

This paper is a response to Hillygus and Jackman in the sense that they imply disequilibrating campaign effects when their data do not necessarily confirm such effects. To make our point, we study the same time period as Hillygus and Jackman using another source of data.

Mean Reversion or Random Walk Campaign Effects

Scholars have long argued over the effect of campaigns in presidential elections (see Hillygus and Jackman, 2003, for an extensive review). Despite the extensive research in the area, no one has offered a complete explanation of what does "Do campaigns matter?" mean. We consider aggregate preferences as a time series leading up to an election, following the example of Wlezien and Erikson (2001). This time series

may be stationary, so that preferences have a mean reversion quality where final votes are predicted by predispositions. Or preferences may be better characterized as the result of a random walk where campaigns change or create preferences apart from how they would be predisposed to vote. To know which is true, a study of the mean reversion versus the random walk theory of elections is required.

In the mean reversion model, voters have predispositions that can determine their vote even before they know what that vote will be. These predispositions are driven by the so-called fundamentals that determine election outcomes. Here, we shall estimate the effects of age (indicators for 18-29, 30-44, 45-64, 65+), region of residence (indicator for south versus non-south), income (five category continuous), sex, race (indicator for black versus non-black) and education (four category continuous) as fundamentals. We include ideology and party identification as fundamentals as well.[1] They were normalized to account for the differing codes in the surveys employed.

Using similar inputs, Gelman and King (1993) show convincing evidence that fundamentals are activated by information flows as campaigns progress and elections draw near. For the 1988 presidential campaign, they show an increasing effect of race and ideology but a decreasing effect of gender and region in determining the vote. The predispositions that grow in importance are those that are out of equilibrium early in the election season. Anderson et al. (2003) similarly find that media coverage during election years helps voters select parties that better represent their ideological ideal points. In 2000, with relative certainty surrounding the eventual major party nominees coupled with the fact that they were well known among the American public, we would expect the fundamentals to be close to equilibrium even well before the election—but to the extent that the campaign changed attitudes, we would expect the fundamentals to become more important toward the end of the campaign.

Increasing Importance of Fundamentals as Campaign Progresses

Using the 2000 National Annenberg Election Survey, we fit a series of logistic regressions predicting vote choice based on the fundamentals in the months leading to the election. The time interval represented by each snapshots is nearly two weeks for the surveys early in the election and decreases to days in the period just before the election when the survey was conducted more frequently. For each snapshot *t*, we fit a model of the form,

$$\Pr(\mathbf{Y}_{it}=1) = \text{logit}^{-1}(\beta_t \mathbf{X}_{it}) \qquad (1)$$

where individuals are subscripted *i* within cross-section *t*; $\mathbf{Y}_{it}$ equals 1 or 0 for supporters of Bush and Gore, respectively; $\mathbf{X}_{it}$ represents a matrix of covariates (the fundamentals); and β_t represents a vector of coefficients for that cross-section. The total sample size of

[1] Debate exists on whether party identification and ideology should be considered fundamental in this sense. For example, Andersen et al. (2003) argue that party identification is not a fundamental variable but a short cut for other fundamentals such as issue positions. We accept this line of thinking. Nonetheless, we include ideology and party identification since it is more feasible than including all of their parts (a series of issue and socialization variables).

the Annenberg surveys employed is 29,544, which we divide into snapshots of approximately 1,000 respondents each.

FIGURE 1 HERE

Figure 1 plots the coefficient estimates (and standard errors) for each predictor over time. Most of the fundamentals show little change in their ability to predict the vote. Race tends to become a more powerful predictor with blacks more likely to support the Democratic candidate, Al Gore. In the early months of the primary election and again after securing delegates for the Republican Party nomination, George W. Bush worked very hard to show his compassionate conservatism by appealing to black Americans. This may have moved black votes out of equilibrium but it did not last long. Further, the relationship between ideology and the vote strengthened dramatically as the election drew near. As voters learned more about the issue stances of each candidate, ideology better separated Gore from Bush voters.

We can also show improvement in model fit as the election approaches. Figure 2 plots the deviance per data point for each of the fitted logistic regression equations over time.[2] Deviance is a general measure of error for generalized linear models (McCullagh and Nelder, 1989). This estimate is stable for large sample studies. It is a measure of misfit. A lower deviance implies the model is fitting better. To yield an average measure of fit, analogous to mean squared error, we divide the deviance by the total sample size of each cross-section studied. Figure 2 shows the deviance per respondent for each of our cross-sections of the Annenberg survey. The fundamentals are stronger predictors of vote choice as the election nears, with deviance per observation dropping from nearly 0.8 at the beginning of the series to below 0.7 just before the election.

FIGURE 2 HERE

Do we know that these findings can be attributed to activation of predispositions by campaign related information flows? No. It may be that no equilibrium level exists for these fundamentals. Rather than move toward equilibrium, they could move in a random walk strengthening when properly shocked. In this scenario, the effect of race and ideology cannot be predetermined. They strengthen (or remain dormant) heading toward Election Day based on the ability of one or more campaign(s) to manipulate these voter attributes. Presumably, then, the campaigns could have chosen to make income or region more important predictors over the course of the election season while ignoring race and ideology.

Or, elections may be largely predetermined by the fundamentals in a process where predispositions are activated by campaign related information flows. Then, we expect the weight given to the fundamentals by the average voter during a survey just before one presidential election to become increasingly similar to the weights given the same fundamentals as the next presidential Election Day approaches. If true, the fundamentals will have been shown to revert back to equilibrium during high information campaign seasons. While they may deviate from this mean when elections are not salient,

[2] Deviance is –2*log likelihood. Deviance per observation is $-2*\sum^{n}_{(i=1)} p(y_i|x_i,beta)/n$; that is, the average value of the log likelihood of the fitted model, averaging over the n data points.

they return to equilibrium levels when called upon.[3] We investigate this hypothesis using poll data from the 1996 and 2000 Presidential elections.

Increasing Predictability of Individual Opinions as Campaign Progresses

In 1996, Bob Dole, the Republican nominee, challenged incumbent Democrat Bill Clinton for the presidency. Each ran a well-funded campaign and received considerable press coverage. Campaign related information increased to voters as the election neared. The same story can be told of the 2000 presidential election except with Bush and Gore in leading roles. With mean reversion, we expect voters' predispositions to have activated as the 1996 election approached, crystallizing on Election Day when information is at its highest. These predispositions then loosen as Election Day passes. They begin to crystallize again as the 2000 presidential election looms.[4]

Under the mean reversion model, the process by which predispositions are activated in 1996 should be the same as the process in 2000 (at least to first approximation) with disregard to the differing candidacies and campaign strategies in those years (that which would promote disequilibrating campaign effects). We can test if this is true. Suppose the coefficients of the fundamentals from a 1996 vote choice equation are shown to explain the relationship between 2000 fundamentals and the 2000 vote better as Election Day approaches. This would show evidence of activation unencumbered from year specific campaign effects. To see how much of vote choice is predetermined by the fundamentals, free from random walk campaign effects, we fit a logistic regression using the 1996 National Election Survey (NES), which was conducted shortly before the November election, to predict the vote preference from the fundamentals. Formally:

$$\Pr(\mathbf{Y}_i^{(1996)}=1) = \text{logit}^{-1}(\mathbf{X}_i^{(1996)}\,\boldsymbol{\beta}^{(1996)}) , \qquad (2)$$

where individuals are subscripted *i*; $\mathbf{Y}_i^{(1996)}$ equals 1 or 0 for Dole or Clinton supporters, respectively; $\mathbf{X}_i^{(1996)}$ represents a matrix of covariates (the fundamentals); and $\boldsymbol{\beta}^{(1996)}$ represents a vector of coefficients for the 1996 presidential election.

To apply this model to the 2000 election, we take the estimated vector $\boldsymbol{\beta}^{(1996)}$ and multiply it by the corresponding variables from each snapshot *t* of the 2000 Annenberg surveys ($\mathbf{X}_{it}^{(2000)}$) to yield a linear predictor ($\mathbf{Z}_{it}$).

$$\mathbf{Z}_{it} = \mathbf{X}_i^{(2000)}\,\boldsymbol{\beta}_t^{(1996)} \qquad (3)$$

We use this as a predictor in a new logistic regression for the 2000 Annenberg vote choice item for the cross-sections conducted from early April to just before the election:

$$\Pr(\mathbf{Y}_{it}^{(2000)}=1) \;=\; \text{logit}^{-1}(\alpha_{0t} + \alpha_{1t}\mathbf{Z}_{it}) , \qquad (4)$$

[3] For example, Bafumi (2003) has found evidence that high information elections activate predispositions in political sophisticates that give them significantly higher reliability scores for ideological self-placement responses as compared to non-sophisticates. In low information contexts, the two groups look no different.
[4] For simplicity, we leave aside the effect of midterm elections for now.

where $\mathbf{Y}_{it}^{(2000)}$ equals 1 for Bush supporters. We fit this regression to each of our snapshots from the Annenberg surveys. We expect α_{1t}, the coefficient of the prediction based on 1996, to increase toward 1 as Election Day 2000 nears. For comparison, we also fit the model for the 2000 NES also using the 1996 NES weights.

FIGURE 3 HERE

Figure 3 shows the trend. The weights assigned to the fundamentals by registered voters in 1996 apply more strongly to vote choice in 2000 as the election draws near. Both the final Annenberg and NES surveys confirm that about 80% of the linear predictor from the 1996 fundamentals applies to the 2000 election. Naturally, one would ask about the remaining 20%. This residual could be accounted for by fundamentals not included in our model[5] or by campaign effects that move beyond the activation of predispositions. Further research is needed to confirm or deny the existence of such random walk campaign effects in presidential elections. The evidence above, however, shows a clear and substantial effect for mean reverting campaign effects.

To date, the search for such random walk campaign effects has been clouded by an unclear definition of what "Do campaigns matter?" means. This paper serves to help clarify the meaning of this question. It further creates a benchmark from which further research can be conducted.

Conclusion

Political candidates invest time, energy, and money into their campaigns. The effect of these campaigns may be to activate predispositions that achieve predictable votes or to force vote outcomes that would not have otherwise happened. In campaigns that are won based on local issues, where resources are severely unbalanced between the candidates and information flows are minimal, we expect both sorts of campaign effects to occur. Presidential election campaigns tend to be nationalized, highly competitive, and with high information.

We posit that campaigns may activate predispositions so that votes are mean reverting or that they may engender out-of-equilibrium or random walk vote outcomes when shocked by biasing campaign information. Our findings from the 2000 election campaign are consistent with Gelman and King (1993) and Anderson et al. (2003) in finding a growing importance for the fundamentals as an election draws near and information flows activate predispositions.

The fact that aggregate election outcomes and individual vote choice both move toward predictable outcomes, provides strong support for the findings of mean reversion campaign effects. This is consistent with the campaign effects found by Hillygus and Jackman (2003). Thus, we do not criticize the empirical results of Hillygus and Jackman (2003), or their general perspective, but rather their belief that they have found definitive evidence of something beyond mean reversion. It may be that presidential election

[5] These variables may be at the individual level or at the macro level such as economic performance, to the effect that these have different effects on different voters.

campaigns create or change vote preferences beyond activation of various predispositions but it remains to scholars to find convincing evidence of such an effect.

References

Andersen, Robert. 2003. "Do newspapers enlighten preferences? Personal ideology, party choice and the electoral cycle." *Canadian Journal of Political Science*, 36: 601-620.

Andersen, Robert, James R. Tilley and Anthony F. Heath. 2003 "Political Knowledge and Enlightened Preferences: Party Preference through the Electoral Cycle." Technical Report, Department of Sociology, McMaster University.

Bafumi, Joseph. 2003. "The Regress of Ideological Sophisticates." Presented at the 2003 meeting of the American Political Science Association.

Cleveland, W.S. 1979. "Robust Locally Weighted Regression and Smoothing Scatterplots." *Journal of the American Statistical Association*, 74: 829-836.

Gelman, Andrew and Gary King. 2003. "Why Are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable?" *British Journal of Political Science*, 23: 409-451.

Hillygus, D. Sunshine and Simon Jackman. 2003. "Voter Decision Making in Election 2000: Campaign Effects, Partisan Activation, and the Clinton Legacy." *American Journal of Political Science*, 47: 583-596.

McCullagh, P. and J.A. Nelder. 1989. *Generalized Linear Models*, Second Edition. London: Chapman and Hall.

Romer, Daniel, Kate Kenski, Paul Waldman, Christopher Adasiewicz and Kathleen Hall Jamieson. 2004. *Capturing Campaign Dynamics: The National Annenberg Election Survey*. New York: Oxford University Press.

Wlezien, Christopher and Robert S. Erikson. 2002. "The Timeline of Presidential Election Campaigns." *Journal of Politics*, 64: 969-993.

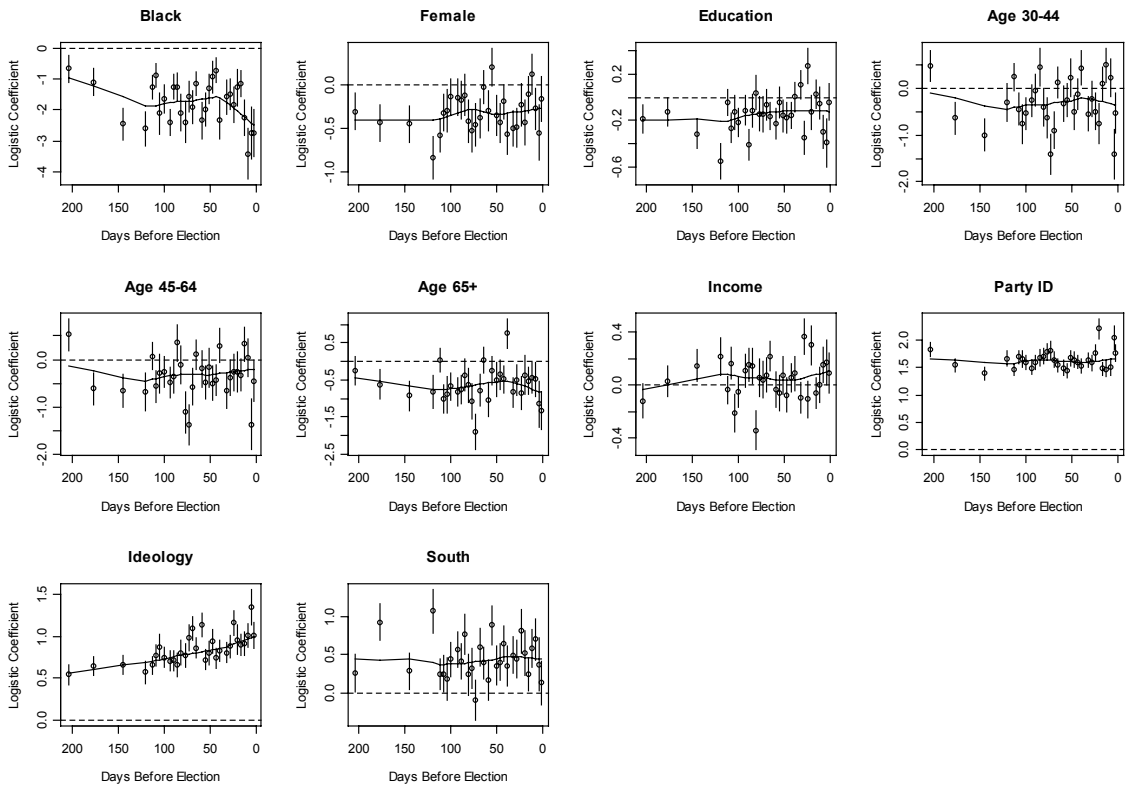

Figure 1: Coefficients with standard error bars for a series of logistic regression using the fundamentals to predict the probability of supporting Bush (among those supporting Bush or Gore) in the 200 days before the 2000 presidential election. Education is on a 1-4 scale, income is on a 1-5 scale and party identification and ideology are represented as z-scores to adjust for different question wordings in different surveys. For each graph, a dotted line indicates zero effect. Lowess smoothed regression lines (Cleveland, 1979) show the trend in the coefficient for each predictor: race and ideology increase in importance during the campaign, and the others are basically stable.

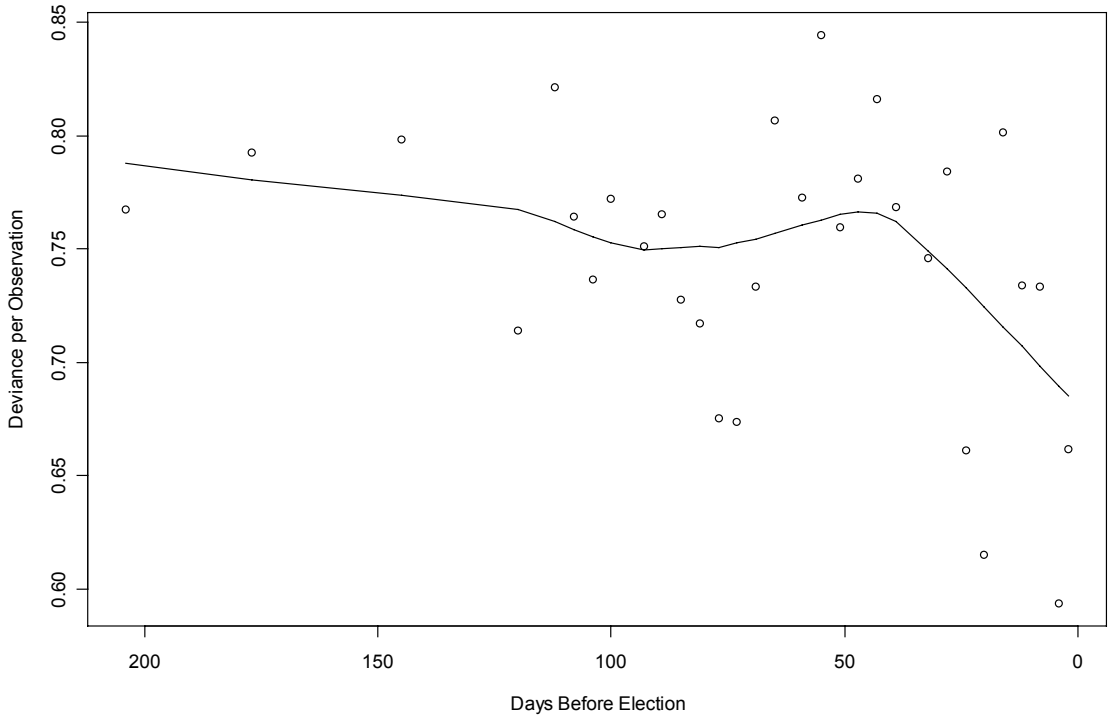

Figure 2: The deviance per observation summarizes the error in the logistic regression model predicting vote choice, as fit to each cross-section of the Annenberg survey (coefficient estimates are displayed in figure 1). As shown by the lowess line, the model fit improves (that is, the deviance decreases) as the election draws near, indicating the increasing predictive power of the fundamental variables.

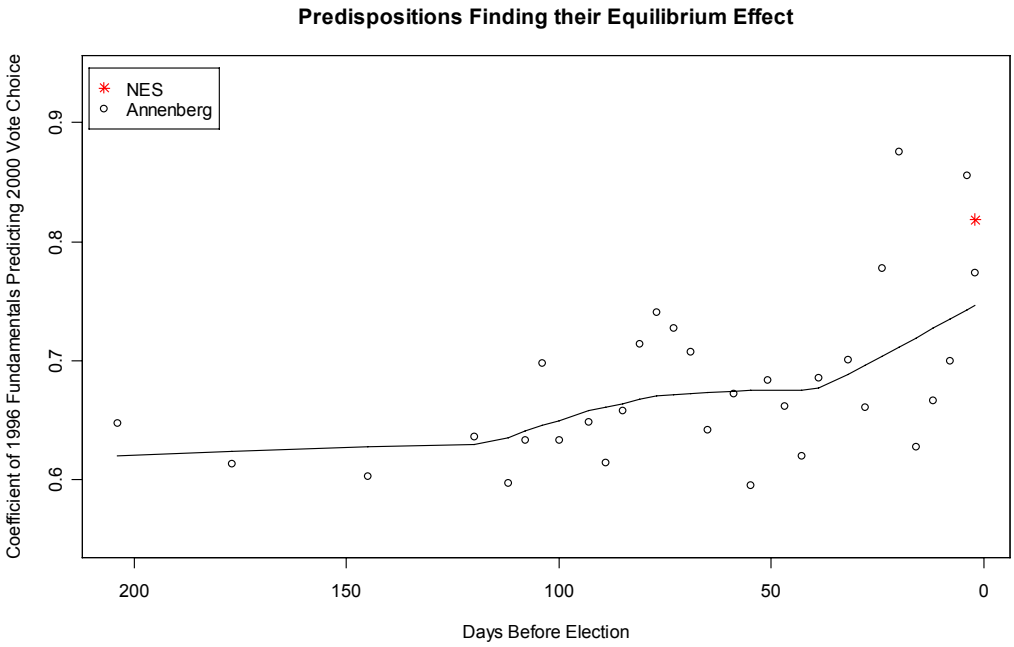

Figure 3: A prediction is generated using the coefficients from a 1996 vote choice equation and covariates from 2000 studies conducted up to 200 days out from the election. This prediction is then used as a covariate predicting 2000 vote choice. The coefficients yielded are plotted above. Vote preference in 2000 can be better predicted by the 1996 fundamentals as the election draws near. A lowess smoother shows the trend.