

Forecasting Presidential Nominations in 2016: #WePredictedClintonANDTrump

Andrew J. Dowdle, *University of Arkansas*

Randall E. Adkins, *University of Nebraska at Omaha*

Karen Sebold, *University of Arkansas*

Jarred Cuellar, *University of Arkansas*

ABSTRACT

A number of scholars successfully modeled and predicted presidential nomination outcomes from 1996–2008. However, dramatic changes occurred in subsequent years that would seem to make replicating these results challenging at best. Building on those earlier studies, we utilize a series of OLS models that included measures of preprimary resources and early campaign successes or failures to forecast that Hillary Clinton and Donald Trump would win the Democratic and Republican presidential nominations in 2016. This outcome suggests that some fundamental factors governing nomination outcomes have not changed despite the conventional wisdom.

Numerous models forecast general election outcomes by employing a variety of economic and political measures to make accurate predictions about whether the party in control of the White House will retain or lose the presidency (for an overview see Campbell 2012). In many ways forecasting presidential nominations presents a more challenging task. Important individual-level cues such as partisanship or systemic-level factors such as economic growth or the popularity of the incumbent are helpful in understanding why a voter might choose Bill Clinton over George W. Bush in 1992. Unfortunately, they are not useful in explaining why the same individual picked Paul Tsongas over Bill Clinton or Tom Harkin nine months earlier in the New Hampshire primary (Steger, Dowdle, and Adkins 2012).

While the McGovern-Fraser reform movement of the early 1970s created a new system of presidential nominations designed to increase the role of voters in picking party nominees, a period of stability in the nomination process of both parties' emerged by the end of the 1980s (Barilleaux and Adkins 1993). As these contests became more routinized, a number of scholars attempted

to forecast the results of the presidential primary season by utilizing factors such as polling, financial resources, and elite support (Adkins and Dowdle 2000, 2001a, 2001b, 2005; Mayer 1996; Steger 2000; see Steger 2008 for a comparison of the forecasts generated by the different models). Momentum from performing well in early primaries was also found to play an important role in determining nomination outcomes (Bartels 1988), though there is some controversy about the precise effect of particular contests (Adkins and Dowdle 2001a; Christenson and Smidt 2012; Hull 2008).

At first glance, current events appear to have altered this equilibrium in at least two important ways. First, super PACs, a relatively new type of political committee that arose from the *Speechnow v FEC* and *Citizens United v FEC* court decisions in 2010, should alter the impact of traditional sources of campaign finance (Dwyre and Braz 2015). Second, the Republican elite has arguably fragmented in recent years, which should affect elite support on the process (Steger 2015). Since traditional forecasting models encountered difficulty predicting the 2004 Democratic nomination correctly (Steger 2008), these new factors should make predicting recent nomination outcomes even more challenging.

MODEL SPECIFICATION

To forecast presidential nomination outcomes this research employs two OLS regression models that use the “open” presidential nomination contests from 1980–2012 and then applies the estimates to the 2016 Democratic and Republican presidential nomination contests to create forecasts for each.¹ The models examine Democratic and Republican contests from 1980 to 2012 inclusive, with the exception of the 1980, 1996 and 2012 Democratic

Andrew J. Dowdle is a professor in the department of political science at the University of Arkansas. He can be reached at adowdle@uark.edu.

Randall E. Adkins is a professor in the department of political science and associate dean of the College of Arts & Sciences at the University of Nebraska at Omaha. He can be reached at radkins@unomaha.edu.

Karen Sebold is a clinical assistant professor in the department of political science at the University of Arkansas. She can be reached at ksebold@uark.edu.

Jarred Cuellar is a graduate student at the University of Arkansas. He can be reached at jcuellar@uark.edu.

racess and the 1984, 1992 and 2004 Republican contests. Generally, nomination races with sitting presidents are foregone conclusions and bias the predictions by inflating the R-square statistic and thereby skewing the model's results (Adkins and Dowdle 2000).²

The models use a number of preprimary indicators (e.g. a candidate's standing in the preprimary Gallup preference polls, percentage of party endorsements, and fundraising success) and indicators of early campaign success (e.g. finishes in the Iowa caucuses and New Hampshire primary) to generate two forecasts: a preprimary forecast and post-New Hampshire primary forecast.³ In each model the dependent variable is the APV (percentage of the aggregate primary vote) received by each candidate for the Democratic and Republican presidential nominations, excluding the results of the New Hampshire primary.

To forecast presidential nomination outcomes this research employs two OLS regression models that use the "open" presidential nomination contests from 1980–2012 and then applies the estimates to the 2016 Democratic and Republican presidential nomination contests to create forecasts for each.

Poll Results

Scholars have found a strong relationship between poll standing in the preprimary season and predicting presidential nomination outcomes (Mayer 1996).

H¹: The greater a candidate's standing in the preprimary preference polls, the higher the percentage of the aggregate primary vote that a candidate will receive.

The *Poll Results* variable represents each candidate's support among self-identified partisans in the average of preference polls taken for their party's nominee for the 1980–2012 races during the fourth quarter of the year prior to the start of the nomination (e.g., Gallup poll averages for October 1–December 31, 1979 for the 1980 Republican nomination contest).⁴

Campaign Expenditures

Candidates who raised the most money in the preprimary period typically won their party's nomination, but in some instances candidates do underperform (Adkins and Dowdle 2002; 2008). To better measure the differential effects of money spent during the preprimary period and the start of the primary season, and to lessen problems with multicollinearity, we separated total fundraising into two variables: campaign expenditures and cash reserves.

H²: The larger the amount of money spent in the preprimary season relative to their opponents, the higher the percentage of the aggregate primary vote that a candidate will receive.

The *Campaign Expenditures* variable is calculated as a percentage of the campaign funds that each candidate spent during the preprimary period, relative to the total raised by all candidates during that time.⁵

Cash Reserves

Cash reserves represent unrealized potential of the campaign to affect the candidate's performance in the future (Adkins and Dowdle 2001b).

H³: The larger the amount of unspent campaign funds at the end of the preprimary period that candidates have relative to their opponents, the higher the percentage of the aggregate primary vote that a candidate will receive.

While early spending often has a tenuous relationship with nomination outcomes, a number of studies (Adkins and Dowdle 2000; Steger 2002) find a strong positive relationship between the amount of financial reserves a campaign has at the end of the preprimary and success during the primary season. In order to control for both inflation and the context of individual election cycles, *Cash Reserves* are calculated as a percentage of the unspent money that each candidate has available at the end of the fourth quarter of the year prior to the election,

relative to the cash reserves of the entire nomination field at the end of that same period.⁶

Endorsements

Despite the changes following the McGovern-Fraser reforms, party elites still manage to play a crucial role in shaping nomination outcomes (Cohen et al. 2008; Steger 2007).

H⁴: The greater number of elite endorsements that candidates have relative to their opponents by the end of the preprimary season, the higher the percentage of the aggregate primary vote that a candidate will receive.

Endorsements represents the endorsements, defined as the unweighted total of House, Senate and gubernatorial endorsements, a candidate has as a percentage of the total endorsements in that contest by December 31 of the year prior to the election.⁷

Iowa

The Iowa caucuses and New Hampshire primary are important early tests of candidate strength. Since Carter's victory in 1976, many candidates spend resources disproportionate to the numbers of convention delegates awarded trying to win support of voters candidates in these two states or at least to try to beat popular expectations (Steger, Dowdle, and Adkins 2004). The first variable to measure candidate strength in Iowa represents whether candidates won the caucuses. Buell (2000) contends that recent winners there and in New Hampshire receive a "bounce" going into the next round of primaries and caucuses.

H⁵: The winner of the Iowa caucuses will receive a higher percentage of the aggregate primary vote than other candidates in the field.

Iowa Win takes the form of a dummy variable with the winner coded as a "1" and the remainder of the cases coded as "0." The second measure is the candidates' share of the vote in the Iowa caucuses, which reflects the variation in candidate performance.

H⁶: *The higher percentage of the vote that a candidate receives in the Iowa caucuses, the higher the percentage of the aggregate primary vote that a candidate will receive.*

New Hampshire

Comparing the effects of Iowa and New Hampshire in nomination forecasts, Adkins and Dowdle (2001b) found that the results of the New Hampshire primary produced a statistically

The preprimary model correctly predicts 10 of the 12 primary vote winners and all eight of the winners from 1980–2000. It misses Kerry in 2004 by a wide margin and predicts that McCain will finish third in 2008. Further, while it is technically correct that Clinton won the primary vote in 2008 and that Obama finished second, it did not predict a particularly close contest between the two. The post-New Hampshire model only makes one mistake, picking Tsongas to win in 1992 and Clinton to be the runner-up

Early popular support, preprimary elite endorsements and a large campaign war chest entering the formal campaign remain strong predictors of successful nominees.

significant impact on predictive capacity, but the Iowa caucuses did not.

H⁷: *The winner of the New Hampshire primary will receive a higher percentage of the aggregate primary vote than other candidates in the field.*

New Hampshire Win takes the form of a dummy variable where the winner coded as a “1” and the remainder of the cases as “0.” The second measure is the candidates’ share of the New Hampshire primary vote, which reflects the variation in candidate performance.

H⁸: *The higher percentage of the vote that a candidate receives in the New Hampshire primary, the higher the percentage of the aggregate primary vote that a candidate will receive.*

DATA ANALYSIS

Despite the changes that occurred in the nomination process over the past few years, the traditional dynamics of models that forecast presidential nomination outcomes persist. The results of the two OLS models are presented in table 1. The *Poll Results* and the *Endorsements* indicators are significant in both models, which echoes the results of previous works on the topic (see Steger 2008). *Cash Reserves*, which had been significant in prior studies (Adkins and Dowdle 2001a; 2001b), is not significant in the preprimary model but is in the post-New Hampshire model. *Campaign Expenditures* is not significant in either model, which suggests higher levels of spending in the preprimary period do not correlate with winning more primary votes when other factors are accounted for in a multivariate model. This finding reminds us that some hopefuls (such as Phil Gramm in 1996) who perform poorly in early polls continue to fare poorly when voters begin to cast ballots in spite of spending large amounts of money in the preprimary period.⁸

The second model in table 1 also measures whether a candidate won Iowa or New Hampshire along with the impact of the votes that each candidate received in each contest. The positive effects of winning Iowa bolsters Hull’s (2008) claim that Iowa does play an important role in nomination outcomes, but the actual vote percentage from Iowa is not significantly correlated with the overall primary results, which is consistent with prior research (Adkins and Dowdle 2001a). On the other hand, both ordinal and interval-level finishes in New Hampshire are positively correlated with nomination outcomes as previous studies have indicated (Steger, Dowdle, and Adkins 2004).

In table 2 we also demonstrate the two models’ abilities to make ordinal predictions for the open races from 1980–2016.

Table 1
OLS Forecasting Models of Aggregate Primary Vote, 1980–2012

	PrePrimary	Post-New Hampshire
Poll Results	.52** (2.70) [.35]	.27* (2.19) [.18]
Campaign Spending	.04 (.24) [.03]	-.30 (-1.98) [-.18]
Cash Reserves	.15 (1.34) [.16]	.16* (2.15) [.16]
Endorsements	.32** (1.74) [.21]	.21* (2.58) [.21]
Iowa Win		11.81** (2.78) [.21]
Iowa Percent		-.06 (-.51) [-.05]
NH Win		16.17** (3.80) [.29]
NH Percent		.52** (3.78) [.38]
Constant	.89 (.40)	-.37 (-.26)
R ²	.60	.85
Adjusted R ²	.58	.84
F	28.80	53.31
SEE	12.99	8.06
N	82	82

Note: Coefficients are unstandardized ordinary least squares (OLS) coefficients; t scores are in parentheses (), standardized beta coefficients are in brackets []. SEE=standard error estimate. Significant at *p < .05, **p < .01.

Table 2
Combined Model Predicted and Actual Finish, 1980–2016

Year	Party	Stage	First	Second	Third
1980	R	Preprimary	<u>Reagan</u>	Connally	Baker
1980	R	Post-NH	<u>Reagan</u>	<u>Bush</u>	Baker
1984	D	Preprimary	<u>Mondale</u>	Glenn	Cranston
1984	D	Post-NH	<u>Mondale</u>	Hart	Glenn
1988	D	Preprimary	<u>Dukakis</u>	<u>Jackson</u>	Gephardt
1988	D	Post-NH	<u>Dukakis</u>	Gephardt	Jackson
1988	R	Preprimary	<u>Bush</u>	<u>Dole</u>	Kemp
1988	R	Post-NH	<u>Bush</u>	<u>Dole</u>	Kemp
1992	D	Preprimary	<u>B. Clinton</u>	Kerrey	Brown
1992	D	Post-NH	Tsongas	B. Clinton	Kerrey
1996	R	Preprimary	<u>Dole</u>	Gramm	Alexander
1996	R	Post-NH	<u>Dole</u>	<u>Buchanan</u>	Alexander
2000	D	Preprimary	<u>Gore</u>	<u>Bradley</u>	
2000	D	Post-NH	<u>Gore</u>	<u>Bradley</u>	
2000	R	Preprimary	<u>Bush</u>	<u>McCain</u>	Forbes
2000	R	Post-NH	<u>Bush</u>	<u>McCain</u>	<u>Keyes</u>
2004	D	Preprimary	Dean	Gephardt	Clark
2004	D	Post-NH	<u>Kerry</u>	Dean	Clark
2008	R	Preprimary	Giuliani	F. Thompson	McCain
2008	R	Post-NH	<u>McCain</u>	Huckabee	Giuliani
2008	D	Preprimary	<u>H. Clinton</u>	<u>Obama</u>	<u>Edwards</u>
2008	D	Post-NH	<u>H. Clinton</u>	<u>Obama</u>	<u>Edwards</u>
2012	R	Preprimary	<u>Romney</u>	Paul	<u>Gingrich</u>
2012	R	Post-NH	<u>Romney</u>	<u>Santorum</u>	Paul
2016	D	Preprimary	H. Clinton	Sanders	O'Malley
2016	D	Post-NH	H. Clinton	Sanders	O'Malley
2016	R	Preprimary	Trump	Rubio	Cruz
2016	R	Post-NH	Trump	Cruz	Bush

Note: Underlined names indicate a correct ordinal forecast in terms of percent of primary vote. Manuscript submitted before end of 2016 primaries so no final results are available yet.

(likely because of his inability to win either Iowa or New Hampshire that year). On the other hand, the two models predict a decisive Romney victory in 2012. Overall, the forecasts are less accurate for second- and third-place finishers even though the post-New Hampshire model correctly predicts 8 of the 12 runner-ups.

In applying the model to the 2016 election cycle, both models forecast victories for Hillary Clinton and Donald Trump. Clinton receives 76% and 58% of the aggregate primary vote in the preprimary and post-New Hampshire primary models, respectively, with Bernie Sanders' New Hampshire victory tightening the race. Trump barely beats Rubio (21 to 17%) in the preprimary forecast of a crowded Republican field, but easily surpasses Cruz (41 to 21%) in the post-New Hampshire model because of Trump's strong finish there.

FINDINGS

Early popular support, preprimary elite endorsements and a large campaign war chest entering the formal campaign remain strong predictors of successful nominees. The results of the Iowa caucuses

and the New Hampshire primary also influence both the eventual nominees and the margins of victory. Clearly, in spite of recent changes such as the increase in outside money and the fragmentation of support among party elites, the traditional models of forecasting presidential nominees are sufficient if they can correctly forecast a conventional insider like Hillary Clinton and an unorthodox outsider like Donald Trump.

ACKNOWLEDGMENTS

The authors would like to thank Dino Christenson, Tom Cronin, Wayne Steger, and two anonymous reviewers for their comments. A previous version of this manuscript was presented at the 2016 Western Political Science Association Annual Meeting. ■

NOTES

1. While the factors that constitute a "serious candidacy" might be open for debate, a consensus has developed that includes the following individuals: for 1980 Anderson, Baker, G. H. W. Bush, Connally, Crane, R. Dole, and Reagan; for 1984 Askew, Cranston, Glenn, Hart, Hollings, J. Jackson, McGovern, and Mondale; for 1988 (Democratic Party) Babbitt, Dukakis, Gephardt, Gore, Hart, J. Jackson, and Simon; for 1988 (Republican Party) G. H. W. Bush, R. Dole, DuPont, Haig, Kemp, Robertson; for 1992 Brown, B. Clinton, Harkin, Kerrey, Tsongas; for 1996 Alexander, Buchanan, R. Dole, Dorman, Forbes, Gramm, Keyes, and Lugar; for 2000 (Democratic Party) Bradley and Gore; for 2000 (Republican Party) Bauer, G. W. Bush, Forbes, Hatch, Key, and McCain; for 2004 Clark, Dean, Edwards, Gephardt, Kerry, Kucinich, Lieberman and Sharpton; for 2008 (Democratic Party) Biden, H. Clinton, Dodd, Edwards, Gravel, Kucinich, Obama, Richardson; for 2008 (Republican Party) Giuliani; Huckabee, Hunter, McCain, Ron Paul, Romney, F. Thompson; for 2012 Gingrich, Huntsman, Johnson, Ron Paul, Perry, Romney, Santorum; for 2016 (Democratic Party) H. Clinton, O'Malley, Sanders; for 2016 (Republican Party) J. Bush, Carson, Christie, Cruz, Fiorina, Graham, Huckabee, Kasich, Rand Paul, Rubio, Santorum, and Trump.
2. For an alternative approach, see Mayer (1996) where contested nominations with incumbent presidents Jimmy Carter in 1980 and George H. W. Bush in 1992 were included.
3. We define the preprimary period as the year prior to the presidential election. While there is obviously activity in the weeks of the year of the election that occurs prior to the start of the Iowa caucuses, the lack of a uniform starting date makes it difficult to consistently measure that activity.
4. The data was gathered from monthly editions of *The Gallup Report* or annual editions of *The Gallup Poll* and Gallup.com from 1979 to 2003. Results after 2003 came from pollingreport.com. Gallup discontinued the process of polling for preprimary "horse race" preferences in 2015 so we utilized the average percent of approval the candidates received in the CNN poll to generate estimates for the 2016 contest.
5. The results are taken from Line 9 "Total Disbursements This Period" of an individual presidential candidate's Federal Election Commission (FEC)

- “Reports of Receipts and Disbursements” (form 3P) for the year prior to the election.
6. The results are taken from Line 10 “Cash on Hand at the End of the Reporting Period” of an individual presidential candidate’s FEC “Reports of Receipts and Disbursements” (form 3P) for the Year-end Report at the end of the preprimary period. Cash reserves can be problematic when candidates such as Steve Forbes have the ability to loan or donate money to their campaigns and this ability is not reflected on the FEC reports. Unfortunately parsimonious generalizable models cannot account for every possibility, but it is worth noting that the most of the candidates with this ability were under-predicted at the interval-level, though not necessarily at the ordinal-level (e.g. Kerry in 2004, Romney in 2012, Trump in 2016).
 7. We would like to thank Wayne Steger for the endorsement figures from 1980–2012. Endorsements for 2016 are from fivethirtyeight.com. Any errors in the interpretation of this data rest solely on this paper’s authors.
 8. To measure the effect of super PAC spending in 2012, we performed a Pearson’s correlation of the unstandardized residuals of the two models in table 1 for 2012 with two measures collected by the FEC: (1) the percentage of outside money that was spent prior to January 1, 2012 by outside groups supporting a candidate and (2) the percentage of outside money that was spent prior to January 1, 2012 by outside groups attacking a candidate. Since it represents one race, the correlation between the two measures and the residuals of the dependent variable was insignificant.

REFERENCES

- Adkins, Randall E. and Andrew J. Dowdle. 2000. “Break Out the Mint Juleps in New Hampshire? Is New Hampshire the ‘Primary’ Culprit Limiting Presidential Nomination Forecasts?” *American Politics Quarterly* 28 (2): 251–269.
- . 2001a. “How Important Are Iowa and New Hampshire to Winning Post-Reform Presidential Nominations?” *Political Research Quarterly* 54 (2): 431–444.
- . 2001b. “Is the Exhibition Season Becoming More Important to Forecasting Presidential Nominations?” *American Politics Research* 29 (3): 283–288.
- . 2002. “The Money Primary: What Influences the Outcome of Pre-Primary Presidential Nomination Fundraising?” *Presidential Studies Quarterly* 32 (2): 256–275.
- . 2005. “Do Early Birds Get the Worm? Improving Timeliness of Presidential Nomination Forecasts.” *Presidential Studies Quarterly* 35 (4): 646–660.
- . 2008. “Continuity and Change in the Presidential Money Primary.” *American Review of Politics* 28 (Winter): 319–341.
- Barilleaux, Ryan J. and Randall E. Adkins. 1993. “The Nomination: Process and Patterns.” In *The Elections of 1992*, ed. Michael Nelson. Washington, DC: Congressional Quarterly, 21–56.
- Bartels, Larry M. 1988. *Presidential Primaries and the Dynamics of Public Choice*. Princeton, NJ: Princeton University.
- Buell, Emmett H., Jr. 2000. “The Changing Face of the New Hampshire Primary.” In *In Pursuit of the White House*, ed. William G. Mayer. Chatham, NJ: Chatham House, 85–144.
- Campbell, James. 2012. “Forecasting the 2012 American National Elections.” *PS: Political Science & Politics* 45 (4): 610–613.
- Christenson, Dino P. and Corwin D. Smidt. 2012. “Still Part of the Conservation: Iowa and New Hampshire’s Say within the Invisible Primary.” *Presidential Studies Quarterly* 42 (3): 597–621.
- Cohen, Marty, David Karol, Hans Noel and John Zaller. 2008. *The Party Decides: Presidential Nominations Before and After Reform*. Chicago, IL: University of Chicago.
- Dwyre, Diana and Evelyn Braz. 2015. “Super PAC Spending Strategies and Goals.” *The Forum*, 13 (2): 245–267.
- Hull, Christopher. 2008. *Grassroots Rules: How the Iowa Caucus Helps Elect American Presidents*. Stanford, CA: Stanford University Press.
- Mayer, William G. 1996. “Forecasting Nominations.” In *In Pursuit of the White House: How We Choose Our Presidential Nominees*, ed. William G. Mayer. Chatham, NJ: Chatham House, 44–71.
- Steger, Wayne P. 2000. “Do Primary Voters Draw from a Stacked Deck? Presidential Nominations in an Era of Candidate-Centered Campaigns.” *Presidential Studies Quarterly* 30: 727–53.
- . 2002. “A Quarter Century of Network News Coverage of Candidates in Presidential Nomination Campaigns.” *Journal of Political Marketing* 1 (1): 91–116.
- . 2007. “Who Wins Nominations and Why?” *Political Research Quarterly* 60 (1): 91–99.
- . 2008. “Forecasting the Presidential Primary Vote: Viability, Ideology and Momentum.” *International Journal of Forecasting* 24 (2): 193–208.
- . 2015. *A Citizen’s Guide to Presidential Nominations: The Competition for Leadership*. New York, NY: Routledge.
- Steger, Wayne P., Andrew J. Dowdle and Randall E. Adkins. 2004. “The New Hampshire Effect in Presidential Primaries.” *Political Research Quarterly* 57 (3): 375–390.
- . 2012. “Why are Presidential Nomination Forecasts Difficult to Predict?” In *The Making of Presidential Candidates, 2012*, eds. William G. Mayer and Jonathan Bernstein. Lanham, MD: Rowman and Littlefield, 1–22.