

Palmetto State Primaries: An Examination of South Carolina's Nomination Contests

[Accepted for publication in *Journal of Political Science*]

James S. Craven
College of Charleston
cravenjs@g.cofc.edu

Jordan M. Ragusa[†]
College of Charleston
ragusajm@cofc.edu

John-Anthony G. Thevos
College of Charleston
john.thevos@gmail.com

Abstract

Compared to research on general elections, the body of work on presidential primaries is limited. At the same time, the body of work on South Carolina's nomination contests pales in comparison to the volume of research on the Iowa caucus and New Hampshire primary. We seek to reconcile these imbalances by building a statistical model that predicts a candidate's vote share in the South Carolina primary and thus helps explain who wins and losses. We find that the key predictors of a candidate's vote share include: their volume of statewide endorsements, their share of media attention, their race (Democratic contests only), whether they are from a neighboring state, their vote share in New Hampshire, and their vote share in Iowa (Republican contests only). In the end, our model explains 78% of the variation in South Carolina primary outcomes from 1988 to 2016. Because this paper was motivated by the 2016 contests, at the end we discuss the implications of the model for our understanding of the 2016 results.

[†] Corresponding Author. Authors listed alphabetically with equal authorship. We would like to thank Phillip Bridgmon, Cole Blease Graham, Jr., and David Rohde for commenting on this paper at the 2016 meeting of the *Citadel Symposium on Southern Politics* in Charleston, SC. We would also like to thank John Holder for providing feedback at the 2016 meeting of the *South Carolina Political Science Association Conference* at Winthrop University in Rock Hill, SC. Thank you to Gibbs Knotts for his support of this project. Lastly, we would like to thank Adam Chamberlain and the anonymous referees at the *Journal of Political Science* for their suggested revisions. As always, all errors are our own.

[1] Introduction

On February 20th, 2016, roughly 700,000 South Carolinians voted in the GOP primary. Donald Trump won 32.5% of the vote, 10% more than his nearest competitors, Marco Rubio and Ted Cruz, who finished with 22.5% and 22.3% of vote (respectively). Although Trump's victory netted him *all* of South Carolina's 50 delegates, the key storylines were Rubio's surprising second place finish and Jeb Bush's decision to drop out despite the aid of \$130-million in total campaign spending.¹ In the days after the contest, observers debated the value of Governor Nikki Haley's endorsement of Rubio, the effect of media coverage in Trump's victory, why spending seemed to have no effect on Bush's vote share, and the value of Ted Cruz's evangelical support.

On the Democratic side, there was less intrigue. Just under 300,000 South Carolinians voted in the Democratic primary on February 27th. As voters headed to the polls, most observers expected Hillary Clinton to win by a large margin as the "fundamentals" seemed to line up on her side. Clinton made over a dozen visits to South Carolina in the weeks leading up to the election, picking up a coveted endorsement from Representative Jim Clyburn (SC-06), and garnered the lion's share of media attention in both state and national outlets. By comparison, her main opponent, Bernie Sanders, focused his attention on Super Tuesday states outside South Carolina and picked up far fewer statewide endorsements. If there was a key storyline in the Democratic race, it was the size of Clinton's victory: 73.5% to Sanders' 26.0%.

¹ See "How Jeb Bush Spent \$130 Million Running for President With Nothing to Show for It" appearing in the *New York Times* on February 22, 2016.

Our objective in this article is to examine the fundamental dynamics of South Carolina's nomination contests from 1988 to 2016. We were motivated to study this topic given the scarcity of scholarly research on South Carolina's primaries. In fact, we identified just one journal article that focuses specifically on South Carolina's nomination contests (Vinson and Moore, 2007) compared to dozens of books and articles on Iowa and New Hampshire. Our main goal is to identify the factors that explain a candidate's performance and better understand who wins (and loses) in the Palmetto State. We do this by collecting original data on the 63 candidates who ran in nomination contests in South Carolina from 1988 to 2016.

[2] Why Study South Carolina's Nomination Contests?

Compared to research on general elections, political scientists have dedicated far less time to primaries. We believe this is the case for a few reasons. On the one hand, the "usual factors" that reliably predict general election outcomes (economic conditions, presidential approval, etc.) do not have the same power in primaries. At the same time, primary elections are difficult to model given their high degree of uncertainty. With low turnout and multiple candidates from the same party on the ballot, primary elections can hinge on idiosyncratic factors that vary state-to-state and election-to-election. Yet nomination contests *deserve* greater attention for a few general reasons. First, winning a party's nomination is the biggest barrier to winning a presidential election. Second, primaries and caucuses are many voters' first exposure to the presidential candidates. And third, the eventual nominee becomes the defacto party leader and plays a key role in shaping the party's platform.

In addition to their general importance, we believe South Carolina's nomination contests are uniquely important. On the Republican side, South Carolina has held the unique position as "first in the south" since 1980. South Carolina has also correctly picked every Republican

nominee since 1980, with the exception of 2012 when Newt Gingrich defeated Mitt Romney. On the Democratic side, although South Carolina has only been “first in the south” since 2004, it is historically one of the earliest Democratic contests and has a strong predictive record as well having correctly picked every winner since 1992 (with the exception of John Edwards in 2004). In addition, South Carolina is typically the first state on the primary calendar that has a large percentage of minority voters: a key Democratic constituency. Simply put, South Carolina seems to be pivotal in *both* parties’ nomination contests. Despite this, there is a significant imbalance in the literature. Although dozens of books and articles have studied nomination contests in Iowa and New Hampshire, we are aware of just one journal article that focuses specifically on South Carolina’s nomination contests (Vinson and Moore, 2007).

In this scholarly vacuum, journalists and pundits have offered various “keys” to winning the South Carolina primary. Among the important factors, two receive almost universal mention: (1) the power of Evangelicals in Republican contests and (2) the pivotal role of Black voters in Democratic contests. For example, exit polls from 2016 revealed that Evangelicals comprised 67% of the GOP primary electorate while Black voters made up 61% of the Democratic primary electorate.² Yet given the recent victories by Trump (2016) and Gingrich (2012), some commentators have questioned the role of religion in GOP primaries. For example, Scott Huffmon, director of the Winthrop University poll, claimed³ that “evangelicals here behave

² CNN Politics. 2016. *South Carolina Exit Polls Database*. Retrieved from:

<http://www.cnn.com/election/primaries/polls>

³ McCormick, J., & Niquette, M. (2016). “Republicans Take White House Fight to South Carolina After Trump Win.” *Bloomberg Politics*. Retrieved from:

as voters and not church-goers.” Rounding out the “keys” to winning South Carolina contests, journalists and pundits also frequently cite the importance of: (3) navigating South Carolina’s diverse constellation of conservative voters and (4) the value of being from a southern or a neighboring state. On the third factor, commentators frequently note that the GOP primary electorate has an ideological balance between social conservatives in the Upstate, socially moderate economic conservatives in the lowcountry, and rural pro-military conservatives in the Midlands. On the fourth factor, observers sometimes cite the anomalous victories by Edwards (2004) and Gingrich (2012) as evidence that southern identity and geographic proximity are important factors in South Carolina contests.

Of course there are a handful of studies that have identified factors that matter in primaries in general. One of the key findings in the literature is the presence of “momentum” effects. Simply put, researchers have found that candidates who do well in early states and “beat expectations” do well in subsequent states because of their perceived electoral viability (Abramowitz 1989; Bartels 1987, 1988; Cohen, Noel, and Zaller 2004; Norrander 2006; Steger 2007). Another key factor is the role of endorsements. Research has consistently shown that in primaries--low-information contests where reliable cues are absent--endorsements have strong effects on outcomes (Cohen et. al. 2008; Dominguez 2011; Steger and Davis 2000; Steger 2007). A key debate in the literature is whether primaries hinge on the ideological “compatibility” between candidates and voters. Aldrich and Alvarez (1994) find evidence of issue-voting in primaries while Cutler (2002), Stockley (2008) and Gopoiian (1982) find no evidence of ideological voting. In contrast, a number of studies cite sociodemographic factors as filling the

<http://www.bloomberg.com/politics/articles/2016-02-10/presidential-primary-moves-to-politically-steamy-south-carolina>

ideological and partisan gap. Factors such as a candidate's age, race, and gender have been shown to powerfully shape how primary voters decide (Cutler 2002; Gopoian 1982; Stockley, 2008). Additional studies have examined the role of the media (Patterson 1980) and campaign spending (Adkins and Dowdle 2000; Steger 2007) in predicting primary election outcomes.

[3] Election Outcomes: 1988-2016

[3.1] Data and Hypotheses

Our dataset records a candidate's performance in South Carolina's nomination contests from 1988-2016. Within this time period we have an even balance of six contested Democratic elections (1988, 1992, 2000, 2004, 2008, 2016) and six contested Republican elections (1988, 1996, 2000, 2008, 2012, 2016). Our dependent variable is candidate's vote share.⁴ It ranges from 0% (for various candidates) to 92% (for Al Gore in 2000). Our predictor variables fall into three categories: campaign effects, candidate resources, and candidate characteristics. For theoretical reasons, we tested whether four variables (New Hampshire Vote %, Iowa Vote %, Candidate Ideology, and Evangelical) have differential effects in each party's contest. We discuss these factors below and their theoretical justification. None of the other factors had significant interaction effects, indicating that it is appropriate to combine both parties into a single model. As just one example, we did *not* find that the effect of media attention mattered more or less to one party's candidate.

Our key campaign variables are *NH Vote %* and *IA Vote %*, which record a candidate's vote share in the New Hampshire primary and Iowa caucus. Research has identified

⁴ Election results were obtained from the *Vital Statistics of American Politics* volume.

“bandwagon” and “momentum” effects in primaries, where candidates who do well in early states receive a larger share of the vote in subsequent states due (in part) to perceptions about their electoral viability (Abramowitz 1989; Bartels 1987, 1988; Cohen, Noel, and Zaller 2004; Norrander 2006). In addition to testing this effect, we are interested in which earlier state contest best predicts South Carolina’s results and for which party’s candidates. On the one hand, New Hampshire is typically closest in time to South Carolina’s election, making it potentially more salient in the minds of voters on both sides of the aisle. On the other hand, Iowa’s demographics more closely resemble South Carolina’s on the Republican side of the aisle.⁵ For these reasons

⁵ Using the 2012 *Cooperative Congressional Election Survey* (CCES) dataset, we compared Democratic and Republican voters in Iowa, New Hampshire, and South Carolina. Although we were unable to find data on the characteristics of *primary* voters in each state, we think this is a reasonable approximation of what each party’s electorate looks like. We found three key results: on ideology, religion, and race. In Iowa, 72.4% of Republicans describe themselves as “conservative” or “very conservative” compared to 71.3% of South Carolinians and just 57.5% of New Hampshire Republicans. We see the same pattern on the Democratic side. In Iowa, 34.9% of Democrats describe themselves as “liberal” or “very liberal” compared to 31.2% of South Carolina Democrats and 45.5% of New Hampshire Democrats. On religion, 50.7% of Iowa Republicans say they were “born again” (a key tenet of Evangelical faith) compared to 56.7% in South Carolinians and just 17% in New Hampshire Republicans. On the Democratic side, Iowa and New Hampshire voters are alike in terms of religion. Just 14.5% of Iowa Democrats say they were “born again” compared to 5.2% of New Hampshire Democrats and a whopping 42.3% in South Carolina Democrats. South Carolina is also the outlier on the Democratic side of the aisle when it comes to race. A whopping 49.4% of Democrats in South Carolina are African

we interact a candidate's vote share in these states with the variable *GOP*, an indicator for party affiliation coded "1" for Republicans and "0" for Democrats. Based on the above discussion, our expectations are that New Hampshire's results will have an effect on both parties' candidates (due to the proximity of the contests and bandwagon effects) while Iowa's results will have an effect in South Carolina on the Republican side (due to demographic similarity, namely the power of Evangelical voters).

We also control for the race's competitiveness. Our variable *Competitiveness* counts the number of a candidate's rivals who received at least 10% of the vote in both Iowa and New Hampshire. We expect a negative effect on this variable, indicating that the greater the number of competitive candidates on the ballot, the fewer votes a candidate receives. If a candidate ran in a previous contest in South Carolina, the variable *Previous Vote* records his or her vote share in the prior race. Because primary elections are low information contests that can come down to factors like name recognition (Bartels 1988; Kam and Zechmeister 2013; Krasno and Green 1988; but see Abramowitz 1975), we suspect that candidates who campaigned in South Carolina in previous elections have a built-in advantage over Palmetto State newcomers.

We have three variables that tap a candidate's resources. *Media Coverage* is an index that combines the volume of local and national media coverage the candidate garnered in the

American compared to just 2.9% of New Hampshire Democrats and 5.6% of Iowa Democrats. As a whole, these demographics confirm that on the Republican side of the aisle, there are much stronger demographic similarities between South Carolina and Iowa (in terms of ideology and religion). On the Democratic side, although there are ideological similarities between Iowa and South Carolina, on race and religion South Carolina is the outlier.

week leading up to the South Carolina contest.⁶ Higher values indicate greater media attention. We believe media coverage is likely to matter because, unlike general elections, the public pays far less attention to the campaign during a primary (Bartels 1988). Our media coverage variable also accounts for the fact that “momentum” matters in primary campaigns (Bartels 1987, 1988; Norrander 2006). *Endorsements* is an index of candidate’s endorsements. Our variable focuses on endorsements from leading South Carolina figures, with more important endorsements (a gubernatorial endorsement, for example) weighted⁷ more heavily than less important endorsements (a state representative, for example). Research has shown that endorsements provide important voting cues, particularly in low information contests like primaries (Cohen et.

⁶ Because of database availability, the only South Carolina newspaper we were able to search for the entire time period was the historical *Post & Courier* database. The total number of stories appearing in the *Post & Courier* in the week before the primary serves as our measure of the volume of local coverage. For the volume of national media coverage, we used the historical *New York Times*. For both variables we computed the relative percent of media coverage in each source by dividing each candidate’s number of articles by the total number of articles among the candidate’s rivals. We then combined relative local and relative national media coverage by multiplying them together.

⁷ All endorsements were coded during our search of the *Post & Courier*. All endorsements were coded for South Carolina politicians. A gubernatorial endorsement is coded as “1,” a senatorial endorsement is coded as “0.5,” a representative’s endorsement is coded as “0.25,” and a state legislator’s endorsement as “0.10.” Any major endorsement that we came across in our search of the *Post & Courier* that is not included in the above coding scheme (another state’s governor, a former South Carolina politician, a major city’s mayor, etc.) is coded as “0.10.”

al. 2008; Dominguez 2011; Steger and Davis 2000). Ultimately, one of the more salient debates⁸ in 2016 concerns whether the “party decides” who wins the nominations via endorsements and other behind the scenes activities (Cohen, Karol, Noel, and Zaller 2008). Finally, *Campaign Contributions* is total dollar amount of campaign contributions a candidate raised divided by the total volume of campaign contributions for a candidate’s rivals in their party’s nomination contest.⁹ A considerable volume of research on campaign contributions has examined the effect of money in congressional elections (see for example Green and Krasno 1988). Yet a handful of studies have found that campaign spending is a key factor in primaries as well (Gurian and Haynes 1993; Gurian 1986). However, this work has shown that the use and effectiveness of campaign spending is contingent on a variety of campaign factors such as name recognition, momentum, and viability (Haynes, Gurian, and Nichols 1997; Gurian 1990).

We also include five variables that tap a candidate’s characteristics. Some studies suggest that candidate characteristics are the most important factors in primary elections (Gopoian 1982; Stockley, 2008). *South* is a variable coded “1” if the candidate is from a Deep

⁸ See for example an article by Nate Silver discussing the role of endorsements and the book “The Party Decides.” <http://fivethirtyeight.com/features/the-republican-party-may-be-failing/>

⁹ Because campaign finance data is released by the FEC according to filing deadlines, the most up-to-date data available at the time of South Carolina’s election is the year-end fundraising totals. Our variable therefore records total campaign receipts through December 31st the year before the primary. We acknowledge that the variable we use—campaign fundraising—is an imperfect measure and that the ideal variable would be the amount of money *spent* in South Carolina on advertisements and other campaign activities. Given the historical scope of our study (back to 1988), and the focus on primary elections, these data are simply not available.

South state.¹⁰ On the one hand, voters may consciously favor candidates from the South. On the other hand, regional identity may serve as a kind of voting cue. We use this operationalization, as opposed to a variable for all former Confederate states, because the states comprising the Deep South are most “like” South Carolina in a variety of ways. A separate variable, *Neighboring State*, records if a candidate was from Georgia or North Carolina. We suspect that any additional effect of being from a neighboring southern state has to do with greater name recognition. Indeed, a number of commentators have speculated that “proximity” explains Gingrich’s strong performance in 2012 as well as Edwards’ victory in 2004. *Black Democrat* is a variable coded “1” for Black Democratic candidates.¹¹ Given that African Americans are a large demographic group on the Democratic side—comprising over half of the Democratic primary electorate—we hypothesize that African American Democratic candidates will do better all else equal. *Evangelical* is an indicator coded “1” for evangelical candidates and “0” for non-

¹⁰ Deep southern states are South Carolina, Georgia, Alabama, Mississippi, and Louisiana. We also tested a variable for former Confederate states. As we find in the main analysis, this alternative variable is insignificant.

¹¹ We faced a difficult decision on how to model the effect of race on the Republican side. Simply put, there is not enough variation among Black Republicans in the other independent variables (partly a function of the fact that Alan Keyes counts for 2/3 of the Black Republicans). As just one example, every Black Republican is also a non-evangelical, which creates a collinearity problem. For this reason, the effect of a candidate’s in our model is essentially the effect of race in Democratic contests only. We also tested a traditional interaction effect between race and party affiliation. We found the same null result for African American candidates on the Republican side and none of the other results changed.

evangelicals.¹² Given the high percentage of South Carolina citizens who self-identify as having been “born again,” perhaps evangelical candidates outperform candidates of other faiths or denominations. Because the power of evangelical voters may be stronger on the Republican side of the aisle, we interact this variable with the candidate’s party affiliation. *Candidate Quality* is an indicator coded “1” if the candidate previously held federal elective office or was former governor and “0” if he or she was a businessman, state lawmaker, or held a cabinet position. Past work has shown that high quality candidates outperform candidates without such electoral experience (Krasno and Green 1988). Although this body of work is based on congressional elections, we think the same result might exist in primaries as well.

We also tested the effect of a candidate’s ideology. Our variable uses Adam Bonica’s CF Scores (2014), where higher values indicate more conservative lawmakers and lower values more liberal lawmakers. We interacted this measure with party affiliation to account for possible differences in what each party’s “base” favors from a policy standpoint. We found no effect of candidate ideology on either side of the aisle. In the end, the absence of an effect of ideology is consistent with past work (Cutler 2002; Gopoian 1982; Stockley 2008; though see Aldrich and Alvarez 1994). In the models reported below, we do not include this variable because two of the candidates who ran in 2016 (Donald Trump and Ben Carson) do not have CF Scores in Bonica’s

¹² On an individual level, the best marker of evangelical faith is the belief that one has been “born-again.” We were unable to find any reliable way of coding this for each individual candidate. For this reason, we used commonly accepted Christian denominations associated with evangelical faith: Baptists (Southern Baptist, Independent Baptist, National Baptist Convention), Evangelical Methodist, Evangelical Lutheran, Assemblies of God, Church of God.

dataset (because they never ran for prior office). Because Trump won the 2016 GOP primary, we felt it necessary to keep this observation in the analysis.

[3.2] Main Findings

Our dependent variable is the percentage of the vote a candidate received. Although the response is continuous, it is bound between 0 and 1. As Papke and Wooldridge (1996, pg. 619) note, “the bounded nature of such variables and the possibility of observing values at the boundaries raise interesting functional form and inference issues.” Simply put, ordinary least squares (OLS) is not an ideal estimation strategy. On the one hand, the independent variables almost certainly have nonlinear effects. For example, while a few endorsements may improve a candidate’s vote share, there is a threshold at which each additional endorsement yields a smaller return in term of vote share. A linear regression would be inappropriate this context. At the same time, OLS could produce predictions that are less than 0% or more than 100% of the vote. Consistent with recommendations, we therefore estimate a fractional logit model (so-called because it uses a logit link function) which estimates non-linear effects in the independent variables and constrains the predictions to a 0-1 range (Baum 2008; Papke and Wooldridge 1996).

Table 1 presents the model’s results. According to Table 1, the model performs reasonably well, explaining 78% of the variation in South Carolina’s primary and caucus results from 1988 to 2016.¹³ In addition to the coefficients and standard errors, the column dy/dx lists

¹³ Because Stata does not compute r-squared in GLM, our r-squared was computed “by hand” as the squared correlation between the observed response and the predicted values. Diagnostic tests indicate that our fractional logit model is indeed superior to ordinary least squares.

the marginal effect of the significant covariates. For continuous variables, the marginal effect was computed as a standard deviation increase above its mean. For dichotomous variables, the marginal effect was computed as a discrete change. With the marginal effects we can say “how much” each factor affects a candidate’s vote share.

[Table 1 about here]

Looking at the results in Table 1, a number of the predictors are significant. Among the campaign effects, we find that the results in Iowa and New Hampshire are predictive of a candidate’s vote share in South Carolina. Figure 1 plots a candidate’s predicted vote share (holding all other values at their mean) varying their performance in Iowa and New Hampshire from 0% to 50%. With respect to the Iowa Caucuses, Figure 1 reveals that a candidate’s performance in Iowa has a large positive effect on the Republican side of the aisle and a minuscule effect on the Democratic side (in fact, the effect is negative and insignificant for Democrats). As discussed in footnote 5 above, this interaction effect supports our claims about the demographic similarity between the GOP’s primary electorates in Iowa and South Carolina and their dissimilarity on the Democratic side. As the marginal effects in Table 1 show, a Republican candidate who finishes a standard deviation above the mean in Iowa (about 33% of the vote) is predicted to garner an additional 19.0% of the vote in South Carolina. With respect to New Hampshire, Figure 1 reveals that a candidate’s performance in New Hampshire has a modest effect on both sides of the aisle. As the marginal effects show, a Republican candidate who finishes a standard deviation above the mean in New Hampshire (about 31% of the vote) is predicted to garner an additional 4.4% of the vote in South Carolina while a Democratic candidate who finishes a standard deviation above the mean in New Hampshire is predicted to garner an additional 2.9% of the vote. As a whole, the analysis indicates that New Hampshire’s

result has a small effect on a candidate's performance in South Carolina regardless of party affiliation while Iowa's result has a large effect, but only on the Republican side.

[Figure 1 about here]

Among the candidate resource variables, we find that both media attention and endorsements have significant effects on a candidate's vote share in South Carolina. As far as how much these two factors matter, the marginal effects reveal that both have approximately equivalent effect sizes. According to the marginal effects in Table 1, a standard deviation increase in a candidate's media attention increases their vote share by 7.2% while a standard deviation increase in endorsements increases a candidate's vote share by 5.8%. None of the other resource variables affect primary outcomes in South Carolina, according to the model. Perhaps surprisingly, candidates who out-fundraised their opponents do not perform better than candidates with less money on hand. Furthermore, the model indicates that candidates who have run previous campaigns in South Carolina do not perform better than political newcomers.

Lastly, two of the five candidate characteristics have statistically significant effects on a candidate's performance. According to Table 1, candidates from neighboring states and African Americans (on the Democratic side) all perform better in South Carolina. Looking at the marginal effects, a Black candidate in a Democratic contest can expect to earn an additional 23.8% of the vote while a candidate from North Carolina or Georgia is expected to earn an additional 19.3% of the vote. Needless to say, these effects are quite large in magnitude and support research suggesting that candidate characteristics are the most important factors in primary elections (Gopoian 1982; Stockley, 2008). Although the size of the effects may be surprising, they certainly comport with the historical record. After all, an African American Democrat has won three primary elections in South Carolina (Jesse Jackson in 1984 & 1988 and

Barack Obama in 2008) while neighboring state candidates have won two primaries (Newt Gingrich in 2012 and John Edwards in 2004). While being from a neighboring state does have a sizable effect on a candidate's vote share, candidates from the Deep South do not outperform candidates from other regions. Equally surprising is the insignificant effect on the variable for Evangelical candidates. Despite the purported power of Evangelical voters in South Carolina, the results indicate that Evangelical candidates do not outperform candidates of other faiths. Finally, we do not find that high quality candidates (those who have held federal elective office or were a former governor) outperform candidates with little to no political experience.

[Table 2 about here]

Based on the estimates in Table 1, we used the model to predict each candidate's vote. Table 2 presents the predictions alongside a candidate's actual vote share. In terms of the model's predictions, the *average* margin of error is 6.4%. As a whole, the model does very well predicting the winner and the ordinal ranking of the top performers. However, Table 2 reveals that it is plagued by a few large errors. Because of this, our model's *median* error is just 3.6%. Looking at Table 2 the model correctly predicts the winner of every nomination contest in each party's contests (92%) with the exception of the 2008 Republican primary. Our model predicts a victory for Mike Huckabee in the GOP contest that year, with an estimated vote of 27%. Although this is close to Huckabee's actual vote share (he received 30%), the model wrongly estimated that McCain would garner just 20% of the vote (he received 33%). Our model's biggest error occurred in the 2000 Democratic caucus, predicting that Bill Bradley would receive

42% of the vote. As we explain in the footnote below, the error on Bradley is the product of two inherent limitations with our model.¹⁴

[3.3] A Closer Look at Iowa and New Hampshire

Our finding of a connection between Iowa, New Hampshire, and South Carolina warrants further inspection. We speculated that this effect varies, in part, based on the demographics of each state's primary electorate. We believe this is a logical conclusion to draw given the robust demographic similarity between Iowa and South Carolina on the GOP side and the modest demographic similarity between New Hampshire and South Carolina on both sides of the aisle (see footnote 5 above).

¹⁴ Our model's first inherent limitation is that it does not adequately take into consideration the nature of outcomes *within* contests. In this case, the model correctly predicted a landslide victory by Gore but also predicted a large vote share for Bradley (which is impossible in practice). The model gives Bradley such a high vote share because he won 37% of the vote in Iowa and 46% in New Hampshire and had received a large share of the media attention in the week before the South Carolina election. A second inherent limitation with our model is that it pools data for primaries and caucuses. In 2000 the Democratic Party held a caucus, so it is not a surprise that Gore (the clear party favorite) did so well among caucus goers (who tend to be more attuned to party politics). Removing caucuses would improve the model's performance but would leave us with only 41 observations. We therefore decided to keep caucuses and primaries in the model together. A model with just primary elections would produce a number of the same results.

What about candidate momentum? Earlier we noted research showing candidates who do well in early states tend to do well in subsequent states because of perceptions about their electoral viability (Abramowitz 1989; Bartels 1987, 1988; Cohen, Noel, and Zaller 2004; Norrander 2006). We therefore conducted a secondary analysis to examine whether momentum—and the electoral proximity of each state’s contest—also plays a role in South Carolina’s presidential nomination contests. Because South Carolina has been an early state on the primary calendar, one possibility is that momentum effects do not exist. We examined this question using data on the number of days between each state’s contests. Although the most recent contests have occurred in February or late January, in the 1980s and 90s all of South Carolina’s elections were held in March. In other words, there is enough variation in timing of each state’s contest that we should be able to discern if larger scheduling gaps correlate with a smaller effect size between Iowa and New Hampshire and South Carolina.

[Figure 2 about here]

We conducted this secondary analysis by re-estimating the model in Table 1. In this secondary model, we used a *triple* interaction between party, the results in Iowa and New Hampshire, and the number of days until South Carolina’s contest.¹⁵ Figure 2 plots the marginal effect of a standard deviation increase above the mean in each state’s result (33% of the vote in Iowa and 31% in New Hampshire) varying the number of days until South Carolina along a

¹⁵ We do not report the model here because of space constraints but are happy to make it available upon request.

typical range.¹⁶ Figure 2 confirms the presence of momentum effects. We can see that Iowa and New Hampshire's effect on South Carolina's results dissipates the longer the gap between contests. For example, a Republican candidate who receives 33% of the vote in Iowa is predicted by this additional model to garner an additional 19% of the vote in South Carolina if the election is held within three or four weeks. At a six week gap, however, the effect decreases to 10% and is no longer statistically significant. We can see the same pattern for Democrats in Iowa, though as Table 1 reveals, the effect is small in magnitude and always insignificant. In New Hampshire the result is the same for both Democrats and Republicans. At one week, candidates for both parties are predicted to earn an additional 12-15% in South Carolina if they receive 31% of the vote in New Hampshire. By three weeks, however, the effect declines to an additional 2-5% and is insignificant. Although the standard errors in Figure 2 are large, preventing us from drawing firm conclusions, the pattern is consistent with the notion that the connection between the results in Iowa, New Hampshire, and South Carolina declines the longer the scheduling gap between the states.

An additional implication of the conditional effect in Figure 2 is that candidates who do well in Iowa and New Hampshire are poised to increase their lead in South Carolina, provided the scheduling gap is small. When the races are further apart, by comparison, candidates who do well in Iowa and New Hampshire are less likely to increase their lead. At the other end of the spectrum, candidates who *underperform* in the two earlier states have an opportunity to recover

¹⁶ Since 1988 the Iowa caucuses have been held roughly three to six weeks before South Carolina's contest while the New Hampshire primaries have been held roughly one to three weeks before South Carolina. Although the gap in some years has fallen outside these intervals, we use these ranges to make straightforward comparisons.

in South Carolina if the scheduling gap is large. As a whole, Figure 2 reveals that candidates who do well in the two earlier states *lose* momentum with large scheduling gaps while candidates who did poorly are more likely to *recover*.

[4] Explaining the 2016 Results

Looking at the 2016 predictions in Table 2, we can see that model performed fairly well. On the one hand, it correctly predicted a victory for both Hillary Clinton and Donald Trump, coming within 3% of both candidates' actual vote totals. It also correctly predicted the top-3 finishers, in order, on the Republican side. Notably, most polls had Ted Cruz coming in second place and Rubio finishing in a distant third. For example, an average of all publicly available polls from *HuffPost Pollster* had Ted Cruz winning 19.3% of the vote (he received 22.3%) and Marco Rubio winning 15.9% of the vote (he received 22.5%).¹⁷ Despite this, it is important to be clear on the difference between polling and our model: *polling is still a superior approach when it comes to election predictions*. Nonetheless, because the model has some predictive value, we believe it is worthwhile to ask why our estimates were so accurate in 2016 and identify the factors that explain the 2016 results.

On the Republican side, Trump's victory can be explained by his strong second place finish in Iowa and his huge volume of media coverage. Although Ted Cruz won Iowa by a 3-point margin, giving him an advantage heading into South Carolina's contest, Trump dominated his rivals in local and national media coverage in the weeks before the election. Compared to Cruz, for example, Trump appeared in 36% more local news stories and 26% more national news

¹⁷ See: <http://elections.huffingtonpost.com/pollster/2016-south-carolina-presidential-republican-primary>

stories according to the data we compiled. In our model, this difference in media coverage is estimated to have increased Trump's vote share by 6%. In the end, this result is consistent with Nate Silver's (2016) claim that Donald Trump is "running a perpetual attention machine." Adding insult to injury for Ted Cruz, recall that our model finds that Evangelical candidates *do not* perform better than non-Evangelical candidates despite the large volume of born again Christians in the state.

As far as Rubio is concerned, he, too, was aided by a strong finish in Iowa. However, the big question is why Rubio finished in second place when most pre-election polls had him far behind Ted Cruz. According to our model, Rubio's second place finish is attributable (almost entirely) to his slate of statewide endorsements. Indeed, he received the endorsement of Governor Nikki Haley, Senator Tim Scott, and Representative Trey Gowdy.¹⁸ Ultimately, his endorsements far exceeded that of his competitors. According to our model, Haley's endorsement (alone) is estimated to have increased Rubio's vote share by 8.5%. And because Haley's endorsement occurred just days before the contest, it is possible the polls were unable to capture this late breaking effect.

As a final matter, one of the key storylines in the aftermath of the election was Bush's weak fourth place finish and his decision to drop out after South Carolina despite being aided by \$130-million in spending.¹⁹ Unfortunately for Bush, our model suggests that candidate receipts

¹⁸ In the analysis we did not count Joe Wilson's endorsement of Rubio because it came on the day of the election.

¹⁹ See "How Jeb Bush Spent \$130 Million Running for President With Nothing to Show for It" appearing in the *New York Times* on February 22, 2016.

have no effect on primary election results. Bush also lagged far behind his competitors in terms of state and national media attention, according to the data we compiled. And while Bush was aided by a few key endorsements (namely, South Carolina Senator Lindsey Graham), he received just 2.8% of the vote in Iowa (one of the key predictors on the Republican side in our model).

On the Democratic side, the model's implications for the 2016 results are clear. Simply put, almost all of the key factors lined up in Hillary Clinton's favor. In terms of the covariates in our model, only one netted Bernie Sanders an advantage: his victory over Clinton in New Hampshire. However, although unfortunate for Sanders, in our model New Hampshire's predictive power in South Carolina is much smaller than Iowa is on the Republican side. Every other major factor in our model is predicted to have benefited Clinton: in particular an advantage in terms of both endorsements and media attention. According to our model, Clinton's endorsements increased her vote share by 14% while her larger share of media attention is estimated to have increased her vote share by 24%.

[5] Conclusions

Our goal in this paper was to identify the key dynamics of South Carolina's nomination contests. As noted earlier, we believe this is worthwhile research given (1) the low volume of research on primary outcomes compared to general election outcomes and (2) the shortage of research on South Carolina's nomination contests despite its status as the "first in the South." Our analysis was based on an original dataset that recorded each candidate's vote share from 1988 to 2016 as well as the characteristics of the candidates, their resources, and broader campaign dynamics.

We believe there are five conclusions worth highlighting. First, our results show that one of the biggest predictors of candidate performance in South Carolina is their performance in

Iowa. However, our model reveals that this effect only exists on the Republican side of the aisle. On the Democratic side, Iowa has virtually no predictive power. Although dozens of studies have documented “momentum” effects in the order of state primaries, our results show that this effect hinges in South Carolina on state demographics and party affiliation. Indeed, demographic data reveal similarities in the GOP electorates in Iowa and South Carolina that do not exist on the Democratic side. Second, we find that New Hampshire’s results have modest predictive power on South Carolina’s results. Unlike Iowa, however, this effect exists for candidates in both parties. A third finding is that receiving endorsements from prominent state party officials and garnering significant media attention are pivotal candidate resources. Because Donald Trump, Ted Cruz, and Marco Rubio all performed well in Iowa, our model indicates that the 2016 GOP contest hinged on Trump’s media dominance and Rubio’s strong slate of statewide endorsements (thus explaining Cruz’s surprising third place finish). Fourth, candidates from neighboring states and African American candidates (on the Democratic side) perform better in South Carolina on average. Consistent with past research on primary results, the model indicates that these candidate characteristics have quite large effects on outcomes. Fifth, and finally, there are a few perhaps two surprising null results. Although Evangelicals are thought to be a powerful force in South Carolina, our data reveal that Evangelical candidates do not perform better than non-Evangelical candidates. Candidates from the Deep South do not have an advantage, either.

Although this project helps address the lack of research on South Carolina’s primaries, greater work is certainly needed. A key unresolved question is *why* South Carolina has such a successful record picking the eventual GOP nominee. With the exception of 2012, the winner of South Carolina has gone on to secure the Republican nomination every four years since 1980. One plausible explanation for this is that South Carolina boasts a healthy mix of conservatives

and is broadly representative of Republican voters nationwide. According to an article in *The Economist*, South Carolina's mix of conservatives includes "God-and-guns rural voters in the hills of the Upcountry...shrink-the-government fiscal conservatives in the Lowcountry... [and] 400,000 ex-servicemen, who hold robust views on national defence."²⁰ Such an effect may resolve a key puzzle in our model: the fact that the results in Iowa and South Carolina are closely linked yet Iowa winners do not have nearly the same track record when it comes to winning the GOP nomination. We believe this paradox may be due to the more diverse GOP primary electorate in South Carolina as well as the timing of South Carolina's contests. Simply put, while Iowa and New Hampshire may "cull" the initial list of candidates, the tradeoff is that the second and third place finishers have time to recover. After South Carolina, by comparison, the race speeds and exacerbates the momentum effects thus leaving candidates less time to climb to the top. All in all, future work will need to examine the how timing and demographics interact to shape a state's predictive ability. Although our work helps address some unresolved questions, our dataset is not suited to examining these issues.

²⁰ See "The Firewall State," *The Economist*, November 21, 2015.

[6] References

- Abramowitz, A.I. (1975). "Name Familiarity, Reputation, and the Incumbency Effect in a Congressional Election." *Western Political Quarterly*, 28 (4), 668-84.
- Abramowitz, A.I. (1985). "Economic Conditions, Presidential Popularity, and Voting Behavior in Midterm Congressional Elections." *Journal of Politics*, 47, 31-43
- Abramowitz, A. I. (1989). "Viability, Electability, and Candidate Choice in a Presidential Primary Election: A Test of Competing Models." *The Journal of Politics*, 51(4), 977-992. doi:10.2307/2131544
- Adkins, R. E., & Dowdle, A. J. (2000). "Break Out the Mint Juleps? Is New Hampshire the "Primary" Culprit Limiting Presidential Nomination Forecasts?" *American Politics Quarterly*, 28(2), 251-269.
- Aldrich, J.H., Alvarez, R. M. (1994). "Issues and the Presidential Primary Voter." *Political Behavior*, 16 (3), 289-317.
- Bartels, L. M. (1987). "Candidate Choice and the Dynamics of the Presidential Nominating Process." *American Journal of Political Science*, 31(1), 1-30.
- Bartels, L.M. (1988). *Presidential Primaries and the Dynamics of Public Choice*. Princeton, NJ: Princeton University Press.
- Baum, C. F. (2008). "Stata Tip 63: Modeling Proportions." *Stata Journal*, 8(2), 299.
- Bonica, A. (2014). "Mapping the Ideological Marketplace." *American Journal of Political Science*, 58, 367-386.
- Cohen, M., Noel, H., & Zaller, J. (2004). "Constant Rules, Changing Game: A General Model of Presidential Primaries, 1972 to 2000." Annual meeting of the American Political Science Association, Philadelphia.
- Cohen, M., Karol, D., Noel, H., Zaller, J. (2008). *The Party Decides: Presidential Nominations Before and After Reform*. Chicago, IL: University of Chicago Press.
- Cutler, F. (2002). "The Simplest Shortcut of All: Sociodemographic Characteristics and Electoral Choice." *The Journal of Politics*, 64, 466-490.
- Dominguez, C. B. K. (2011). "Does the Party Matter? Endorsements in Congressional Primaries." *Political Research Quarterly*, 64, 534-44.
- Gelman, A., King, G. (1993). "Why Are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable?" *British Journal of Political Science*, 23, 409-451.
- Gopoiian, J.D. (1982). "Issue Preferences and Candidate Choice in Presidential Primaries." *American Journal of Political Science*, 26 (3), 523-546.

- Green, D. P., Krasno, J. (1988). "Salvation for the Spendthrift Incumbent: Estimating the Effects of Campaign Spending in House Elections." *American Journal of Political Science*, 32, 884-907.
- Gurian, P. H. (1986). "Resource Allocation Strategies in Presidential Nomination Campaigns." *American Journal of Political Science*, 30(4), 802-21.
- Gurian, P. H. (1990). "The Influence of Nomination Rules on the Financial Allocations of Presidential Candidates." *The Western Political Quarterly* 43(3), 661-87.
- Gurian, P. H., and A. A. Haynes. (1993). "Campaign Strategy in Presidential Primaries, 1976-88." *American Journal of Political Science*, 37(1), 335-41.
- Haynes, A. A., P H. Gurian, and S. M. Nichols. (1997). "The Role of Candidate Spending in Presidential Nomination Campaigns." *The Journal of Politics* 59(1), 213-25.
- Herrnson, P. S., & Morris, I. L. (2007). "Presidential Campaigning in the 2002 Congressional Elections." *Legislative Studies Quarterly*, 32(4), 629-648.
- Jacobson, G.C. (2006). "Measuring Campaign Spending Effects in U.S. House Elections." In Brady, H.E., Richard, J. (Eds.) *Capturing Campaign Effect*, pgs. 199-220.
- Kam, C.D., Zechmeister, E.J. (2013). "Name Recognition and Candidate Support." *American Journal of Political Science*, 57 (4), 971-986.
- Krasno, J.S. & D. P. Green. (1988). "Preempting Quality Challengers in House Elections". *Journal of Politics*, 50, 920-36.
- Norrander, B. (2006). "The Attrition Game: Initial Resources, Initial Contests and the Exit of Candidates During the US Presidential Primary Season." *British Journal of Political Science*, 36(03), 487-507.
- Papke, L.E., Wooldridge, J. (1996). "Econometric Methods for Fractional Response Variables With An Application to 401(K) Plan Participation Rates." *Journal of Applied Econometrics*, 11, 619-632.
- Patterson, T. E. (1980). *The Mass Media Election: How Americans Choose their President*. Praeger Press.
- Silver, N. (2016). "Donald Trump is Running a Perpetual Attention Machine." *FiveThirtyEight.Com*. Retrieved from: <http://fivethirtyeight.com/datalab/donald-trump-is-running-a-perpetual-attention-machine/>.
- Steger, W. P. (2007). "Who Wins Nominations and Why?: An Updated Forecast of the Presidential Primary Vote." *Political Research Quarterly*, 60(1), 91-99.

Steger, W. P., & Davis, A. (2000). *Candidate Endorsements in Presidential Nomination Campaigns. Voting in America*. Westport, CN: Greenwood Press.

Stockley, J. (2008). "Social Forces and the Primary Vote: Examining Race, Gender, Age, and Class in the 2008 Presidential Primaries." *Race, Gender & Class*, 15, 32-50.

"The Firewall State." (2015, November 21). *The Economist*. Retrieved from:
<http://www.economist.com/news/united-states/21678802-if-republican-campaign-return-normality-it-will-do-so-south-carolina>

Vinson, C. D., Moore, W.V., (2007). "Campaign Disconnect: Media Coverage of the 2000 Republican Presidential Primary in South Carolina." *Political Communication*, 24 (4), 393-413.

Table 1: Predictors of South Carolina Election Outcomes 1988-2016

	<u>Coefficient</u>	<u>SE</u>	<u>dy/dx</u>
Competitive	-0.31**	(0.14)	-3.9%
GOP	-0.30	(0.50)	
Iowa %	-0.54	(1.35)	-1.1% (D)
GOP * Iowa %	6.24***	(1.41)	19.0% (R)
New Hampshire %	2.32*	(1.28)	4.4% (D)
GOP * New Hampshire %	-1.37	(1.32)	2.9% (R)
Previous Vote	0.23	(1.19)	
Media Attention	9.52***	(1.96)	7.2%
Campaign Receipts	-1.15	(0.91)	
Endorsements	0.66***	(0.18)	5.8%
South	0.17	(0.83)	
Neighboring State ¹	1.76**	(0.74)	19.3%
Black Democrat ¹	1.43***	(0.26)	23.8%
Evangelical	0.18	(0.59)	
GOP * Evangelical	0.09	(0.74)	
Quality ¹	-0.14	(0.28)	
Constant	-2.50***	(0.57)	
Observations	63		
R2	0.78		

*** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses.

¹ Marginal effect computed for a discrete (1-0) change. For the continuous factors, the marginal effects were computed for a standard deviation increase above the mean.

Table 2: Actual and Predicted Vote Share

<u>Democratic Contests</u>				<u>Republican Contests</u>			
<u>Candidate</u>	<u>Year</u>	<u>Actual Vote</u>	<u>Predicted Vote</u>	<u>Candidate</u>	<u>Year</u>	<u>Actual Vote</u>	<u>Predicted Vote</u>
Bruce Babbitt	1988	0.00	0.02	Pierre S. du Pont, IV	1988	0.00	0.03
Paul Simon	1988	0.00	0.03	Jack Kemp	1988	0.12	0.05
Dick Gephardt	1988	0.02	0.07	Pat Robertson	1988	0.19	0.17
Michael Dukakis	1988	0.07	0.08	Bob Dole	1988	0.21	0.25
Al Gore	1988	0.18	0.19	George H.W. Bush	1988	0.48	0.54
Jesse Jackson	1988	0.54	0.47	Phil Gramm	1996	0.00	0.02
Eugene McCarthy	1992	0.00	0.05	Richard Lugar	1996	0.00	0.02
Tom Harkin	1992	0.07	0.06	Robert K. Dornan	1996	0.00	0.02
Jerry Brown	1992	0.06	0.07	Alan Keyes	1996	0.02	0.03
Bob Kerrey	1992	0.00	0.08	Steve Forbes	1996	0.13	0.06
Paul Tsongas	1992	0.18	0.14	Lamar Alexander	1996	0.10	0.10
Bill Clinton	1992	0.63	0.33	Pat Buchanan	1996	0.29	0.16
Lyndon LaRouche	2000	0.00	0.04	Bob Dole	1996	0.45	0.61
Bill Bradley	2000	0.02	0.42	Gary Bauer	2000	0.00	0.05
Al Gore	2000	0.92	0.75	Alan Keyes	2000	0.05	0.08
Dick Gephardt	2004	0.00	0.03	Steve Forbes	2000	0.00	0.14
Dennis Kucinich	2004	0.00	0.03	George W. Bush	2000	0.53	0.39
General Wesley Clark	2004	0.07	0.03	John McCain	2000	0.42	0.50
Joe Lieberman	2004	0.02	0.03	Rudy Giuliani	2008	0.02	0.02
Howard Dean	2004	0.05	0.07	Fred Thompson	2008	0.15	0.03
Rev. Al Sharpton	2004	0.10	0.13	Ron Paul	2008	0.04	0.05
John Kerry	2004	0.30	0.15	John McCain	2008	0.33	0.20
John Edwards	2004	0.45	0.27	Mitt Romney	2008	0.15	0.23
Hillary Clinton	2008	0.26	0.15	Mike Huckabee	2008	0.30	0.27
John Edwards	2008	0.18	0.36	Michelle Bachmann	2012	0.00	0.02
Barack Obama	2008	0.55	0.52	Rick Santorum	2012	0.17	0.13
Bernie Sanders	2016	0.26	0.39	Ron Paul	2012	0.13	0.13
Hillary Clinton	2016	0.74	0.71	Mitt Romney	2012	0.28	0.30
				Newt Gingrich	2012	0.40	0.40
				John Kasich	2016	0.07	0.03
				Jeb Bush	2016	0.08	0.06
				Ben Carson	2016	0.07	0.03
				Ted Cruz	2016	0.22	0.24
				Marco Rubio	2016	0.23	0.26
				Donald Trump	2016	0.33	0.30

Figure 1: Effect of Iowa and New Hampshire on South Carolina Vote Share

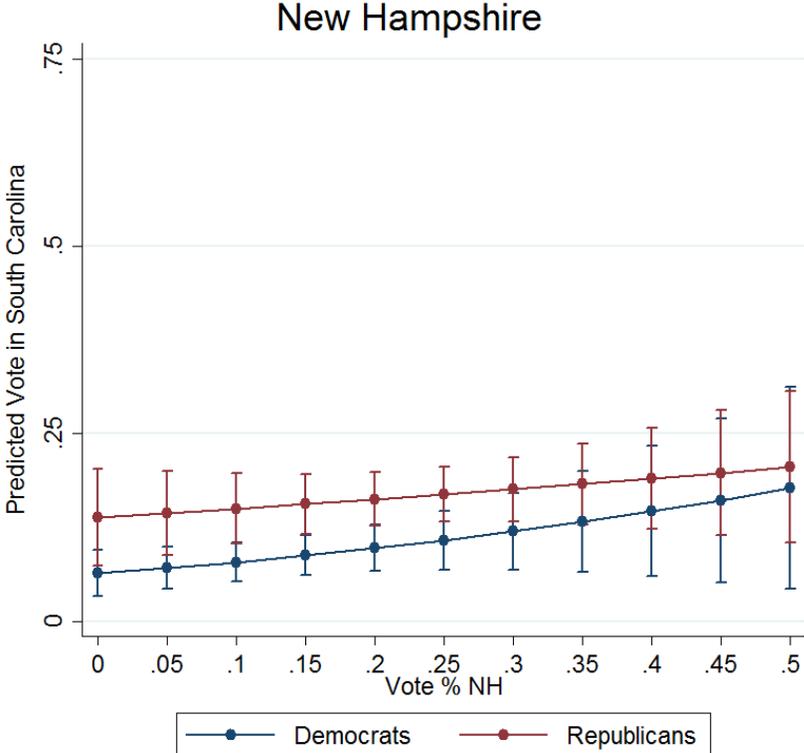
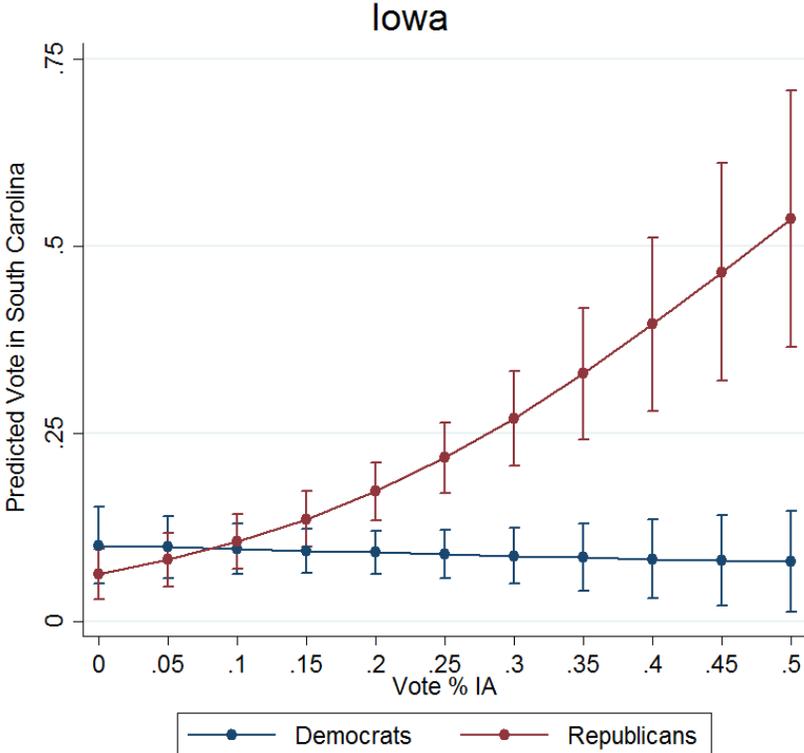


Figure 2: Momentum After Iowa and New Hampshire

