

**An Analysis of the 2016 U.S. Republican Presidential Primary Election****Richard V. Adkisson**: [radkisso@nmsu.edu](mailto:radkisso@nmsu.edu)**James T. Peach**[jpeach@nmsu.edu](mailto:jpeach@nmsu.edu)**New Mexico State University**

---

**Abstract**

This paper examines the 2016 Republican presidential primary vote for Donald Trump. Trump's primary campaign rhetoric emphasized the need to make America great again and proposed that he is the best candidate to make this happen. Counties with older, less educated, and economically stressed populations tended to give strong support to Trump while his potentially offensive rhetoric did not have clear negative consequences in the primary. Conclusions are based on the analysis of data from 2238 counties or county equivalents in the 35 states that held primary elections and held both Republican and Democratic primaries on the same day.

**JEL:** A1, D7, P16**Keywords:** 2016 primary election, voting models, political behavior

---

**Introduction**

There have been strange presidential elections before. McLaughlin (2016) lists the 10 most bizarre. The 2016 presidential primary has been colored by the unexpected success of two less-than-conventional candidates. On the Republican side, brash, outspoken businessman, former Democrat, and untested politician, Donald Trump, handily beat a stable of both conservative and mainstream Republican candidates and eventually won the general election. While he didn't win, independent and socialist Bernie Sanders led a significant grass-roots challenge to the more mainstream Democratic candidate, Hillary Clinton. The 2016 presidential primary election may become a contender for the "bizarre" list in spite of the noteworthy and groundbreaking result that for the first time a woman, Hillary Clinton, was nominated to run for president by one of the two major U.S. political parties.

The purpose of this paper is to explore the primary vote for Donald Trump. Belying expectations Trump won the Republican primary and eventually the general election. While many have been shocked to hear his anti-immigrant, anti-Mexican, anti-globalization, and anti-Muslim rhetoric, he has found a niche among American voters with his campaign theme of making America great again. On

average, at the county level, Trump received 47.8 percent of the Republican primary votes. Support was not universal. County rates ranged from 9.1 percent for Trump (Utah County, UT) to 91.5 percent for Trump (McDowell County, WV). The aim of this paper is to provide insight into his success among Republican primary voters.

**The 2016 Presidential Primary Election**

Primary election rules vary from state to state. Each political party sets its own primary rules. Some states hold primary elections, some hold caucuses. Some hold Republican and Democratic primaries on the same dates, some on different dates. Some states limit primaries to voters registered with the party and others allow crossover voting or last minute changes of party. The rules for assigning delegates to the party conventions vary as well and the primary season stretches over several months. In 2016 the primary period began on February 1 with the Iowa caucus and ended on June 14 with the District of Columbia. All of this makes meaningful analysis of primary elections difficult. To minimize the possible impacts of this variety of rules, this paper focuses on the subset of states that held primary elections (rather than caucuses), and held both the Democratic and Republican primaries on the same day. Thirty-five states meet these criteria.

Trump was the top vote getter in 29 of these states. The states included are shown in Table 1.

Table 1  
States Included in Analysis

New Hampshire	Tennessee	Mississippi	Arizona	Maryland	California
Alabama	Texas*	Florida	Utah*	Pennsylvania	Montana
Arkansas*	Vermont	Illinois	Wisconsin	Rhode Island	New Jersey
Georgia	Virginia***	Missouri	New York	Indiana	New Mexico
Massachusetts	Louisiana	North Carolina	Connecticut	West Virginia	South Dakota
Oklahoma	Michigan	Ohio	Delaware	Oregon	

\* Some data for Carroll County, AR, Zavala County, TX, Beaver, Plate, and Rich Counties, UT, and 58 of 133 Virginia counties were missing. These counties are not included in the data set.

**Brief Literature Review**

Social Scientists, including economists, have long been interested in the electoral process. While several models have been developed to explain general election voting, it has been more difficult to develop reliable empirical models of presidential primary voting (Steger 2007). Primary voters are more ideologically motivated than are general election voters (Norrander 1989). Because primary candidates compete with others in the same party, clear policy distinctions across candidates can be difficult to discern (Aldrich and Alvarez 1994). One result is that policy differences may be exaggerated for political gain leading to divisive primary campaigns within parties that can have dire implications in the general election (Kenney and Rice 1987; Gurian, et al 2016). Primary election voter turnout is typically low and the choices faced by voters evolve as less-successful candidates drop out of the race before the end of the primary season (Aldrich and Alvarez 1994; Norrander 2006). When a sitting president or vice-president is a candidate, primary voters may resort to retrospective voting, basing their choices on past performance rather than future prospects (Mayer 2010).

At the extreme, Dowdle, Adkins, and Steger (2009) suggest that primary results may be determined before the first primary election or caucus is even held with candidates being chosen in an invisible primary where party elites essentially choose the winning candidate for each party before the voting begins. A slightly less deterministic theory proposes that momentum gained (or lost) in the earliest caucuses and primaries strongly influences the result (Steger 2013). Rational choice is seldom mentioned in assessments of primary voting. An

exception is Abramowitz (1989) who analyzes exit polls from the 1988 primary and finds some evidence that primary voters are rational utility maximizers.

Economists who study voting, at least general election voting, often begin by assuming that economic rationality influences voting decisions (Downs 1957; Fair 1996). Rational choice models are typically supplemented with other information to capture influences beyond pure economic rationality (Walker, 2006). Geography can be important as well. Holbrook (1991) finds evidence of a home-region advantage in presidential elections, Adkisson and Peach (1999) find that living near the Mexico-U.S. border influences presidential voting, and Campbell (1992) includes six regional variables to capture evolving regional trends in presidential voting. More recently, Adkisson and Peach (2017) analyzed the 2016 primary election in Texas and found that voters in counties adjacent to the Mexican border and with high shares of recent foreign born were less likely to vote for Donald Trump than were voters in other Texas counties.

The voting literature does not provide clear guidance as to the theoretical or empirical influences on primary voting. There are suggestions that ideology, policy exaggeration, candidate attrition, retrospective voting, voter turnout, momentum, rationality, socioeconomic and political factors, and geography all have a role in primary voter behavior. With the exception of retrospective voting (no sitting candidates in 2016), the model posited below attempts to incorporate all of these potential influences to gain insight into Donald Trump’s primary election success.

**The Empirical Model and Method**

The unit of observation in this study is the county (or equivalent). Twelve continuous independent variables and 34 dummy variables are incorporated to operationalize the potentially influential factors suggested by the literature review. The results do not provide information on specific voters. Rather, they capture the relationships between aggregate variables measured at the county level and the percentage vote for Trump in the county, the

dependent variable. The 12 continuous variables capture variations across counties. The 34 categorical variables are included to capture unobserved state specific effects. The state variables are important because residents live in both a state and a county and are likely to be influenced by conditions at both levels. Many potentially influential factors, tax policy, education policy, infrastructure provision, etc. are largely determined at the state level. Thus, voters are likely to be influenced by state-level as well as county-level economic conditions. The state primaries occur at different times in the primary cycle. Including the state dummies should help account for geography, state-level social and economic conditions, momentum, attrition, varying candidate efforts across states, and varying voter participation rates.

Given the structure of the empirical model, the coefficients on the continuous, county-level, variables should be interpreted as the county effects of these variables on the Trump vote, given overall state conditions. Table 2 shows the variables included in the empirical model. The data are from the most recently available pre-primary time periods. Variance inflation factors were calculated for the continuous variables and indicated that multicollinearity was not severe. Residual diagnostics indicated heteroscedasticity so the hypothesis tests were conducted using heteroscedastic corrected standard errors. The estimated results are shown in Table 3.

Table 2  
Variables Used in Empirical Model

Variable	Description	Source
<b>Dependent Variable</b>		
TRUMP	Percentage of the popular vote for Donald Trump in 2016, by county	1
<b>Independent Variables</b>		
REPPCT12	Percent of popular vote for GOP candidate in 2012 general election	1
EVANRATE	Rates of adherence per 1000 population, evangelical Protestant, 2010	6
VETPCT	Veterans as a percentage of the county's population age 18 and over.	2
A65PLUS	Percent of county population age 65 or more, 2014	3
BAPLUS	Percent of county population 25 years of age and older whose highest educational attainment is a bachelor's degree or higher, 2014	2
POP2014	County annual population estimate as of July 1, 2014	5
FORBORN	Percent of county population that is foreign born, 2014	2
HISPANIC	Percent of county population Hispanic or Latino (of any race), 2014	3
BLACK	Percent of county population reporting Black race only 2014	3
PCPI2014	County per capita personal income, nominal, 2014	4
POVALL	Percentage of population in poverty, all ages, 2014	7
UNEM2015	County unemployment rate 2015	8
AL-WI	State Dummy Variables, New York (NY) is the base state	
<b>Data Sources</b>		
1	Politico.com <a href="http://www.politico.com">http://www.politico.com</a> ,	
2	American Community Survey, 5 year estimates, 2010-2014, Table DP02	
3	American Community Survey, 5 year estimates, 2010-2014, Table DP05	
4	Bureau of Economic Analysis, Regional Data, CA1, Personal income Summary, Personal Income, Population, Per Capita Personal Income	
5	U.S. Census Bureau, Annual Estimates of the Resident Population: April 1, 2010 to July 1, 2014	
6	Association of Religion Data Archives, U.S. Religion Census: Religious Congregations and Membership Study, 2010 (County File)	
7	U.S. Census Bureau, Small Area Income and Poverty Estimates	
8	Bureau of Labor Statistics, Local Area Unemployment Statistics	

Table 3  
Estimated Coefficients

VARIABLE	COEFFICIENT (t-value)	VARIABLE	COEFFICIENT (t-value)	VARIABLE	COEFFICIENT (t-value)
REPPCT12	0.027 (1.536)	POP2014	9.13E-07 (1.572)	POVALL	0.130 (3.636)*
EVANRATE	-7.35E-03 (-6.817)*	FORBORN	0.145 (2.593)*	UNEM2015	0.676 (5.602)*
VETPCT	0.183 (3.435)*	HISPANIC	-0.097 (-4.427)*	CONSTANT (New York)	51.484 (24.290)*
A65PLUS	0.306 (8.834)*	BLACK	-0.011 (-0.603)	R-Square (adj)	0.8936
BAPLUS	-0.392 (-17.500)*	PCPI2014	3.93E-05 (2.022)**		

\* indicates detectable relationship at  $\alpha=0.01$ , \*\* at  $\alpha=0.05$   
Estimates on 34 state dummy variables are available from the authors (discussed below).

## Results

The first five continuous variables are intended to operationalize ideological characteristics. REPPCT12 should reflect the counties general preferences for Republicans or Democrats. The relationship between REPPCT12 and TRUMP is not statistically detectable indicating that a county's recent party preferences in the general presidential election are unrelated to the 2016 primary vote for Trump. EVANRATE, the rate of adherence to evangelical Protestantism, should capture a county's general degree of religious, and likely, social conservatism. As a group Evangelicals are often targeted by candidates as well and candidates often work to make their messages palatable to this group. In this case the higher the value of EVANRATE the lower the vote for Trump. A one-person increase in evangelical adherence relates to a 0.027 percentage point decrease in the Trump primary vote. Veterans do not necessarily vote as a block or represent a particular ideology but as a group, veterans are often assumed by candidates to be relatively

patriotic, concerned about veterans' issues, and perhaps interested in defense issues. VETPCT, the percent of a county's over 18 years old population who are veterans shows a positive and statistically detectable relationship with the Trump vote. A one percentage point increase in a county's veteran population relates to a 0.183 percentage point increase in the Trump vote.

Like veteran status, age does not automatically indicate ideology but Trump's rhetoric refers to America's lost greatness so it seems reasonable that older Americans might have a different perspective on this message than younger Americans. A65PLUS, the percentage of a county's population age 65 or older relates positively to the Trump vote and the relationship is statistically detectable. A one percentage point increase in the value of A65PLUS relates to a 0.306 percentage point increase in the Trump vote. Hypothetically education should enable one to adopt a worldview that appreciates nuances in social and political life and perhaps to eschew overly simple answers to complex problems. BAPLUS, the percentage of a county's population that has attained at least a bachelor's degree, is negatively related to the Trump vote and the relationship is statistically detectable. A one percentage point increase in BAPLUS relates to a 0.392 percentage point decrease in the Trump vote. POP2014, a county's population in 2014, is included to allow for possible ideological differences across the rural-urban spectrum. If these ideological differences exist, they seem to have no statistically detectable relationship with the Trump vote.

FORBORN, HISPANIC, and BLACK are included to capture the impacts of county level social (demographic) characteristics in the face of policy exaggeration. These particular variables are included because much of Trump's campaign rhetoric was overtly anti-immigrant with special emphasis on illegal immigration from Mexico and immigration from Islamic nations. His overall message was labeled racist by some. FORBORN, the percentage of a county's population that is foreign born, is positively related to the Trump vote and the relationship is statistically detectable. A one

percentage point increase in the value of FORBORN relates to a 0.145 percentage point increase in the Trump vote. At first this seems counterintuitive given the candidate's rhetoric but from other perspectives it makes sense. First, the aggregate data cannot tell us whether it is naturalized immigrants or the native born people who share space with them who are voting. Non-citizen foreign born would be unable to vote, whatever their opinion might be. It could also be possible that immigrants who entered the U.S. legally are angered by others who enter illegally and therefore support Trump's strong stand against illegal immigration. Another possibility is that Trump has focused particularly on Mexican and Islamic immigrants leaving other immigrant groups unoffended. The results on HISPANIC, the percentage of a county's population that is Hispanic, suggests that this might be the case. A one percentage point increase in a county's Hispanic population relates to a 0.097 decrease in the Trump vote. BLACK, the percentage of a county's population identifying as Black did not have a statistically detectable relationship with the Trump vote.

To operationalize economic rationality, three variables are included, PCPI2014, county per-capita personal income, POVALL, county poverty rate, and UNEM2015, county unemployment, all show positive and statistically detectable relationships with TRUMP. A one dollar increase in PCPI2014 relates to a .00004 percentage point increase in the Trump vote. Republicans in higher income counties had a slight preference for Trump perhaps suggesting that they relate to his purported business acumen or his promise of an even better tomorrow. Given the results on POVALL and UNEM2015 it could also be the case that higher per-capita incomes in a county make being poor and/or unemployed seem all the worse. The results on POVALL show that a one percentage point increase in a county's poverty rates relates to a 0.130 percentage point boost in the Trump vote. Similarly, a one percentage point increase in UNEM2015 relates to a 0.676 percentage point increase in the Trump vote. These last two results suggest that those struggling economically are

attracted to Trump's message that America has lost its greatness and that globalization, immigration, and politics as usual are to blame.

An analysis of the dummy variables can provide some insight into the geographic dimensions of Trump support, at least for the 35 states included here. The constant term reported in Table 3 is 51.484 and applies only to New York. The interpretation of this number is that if all of the other variables took a zero value, the model predicts that 51.484 percent of New York Republican primary voters would vote for Trump. When statistically different, the coefficients on the state dummy variables are added to the New York value to give an equivalent prediction for each state. Two states, Indiana and Pennsylvania show no statistical difference in their constants when compared to New York. Twenty-one states have constants less than that of New York and 12 have constants greater than that of New York. States at the furthest distance in the negative direction are Utah (-35.627), Texas (-26.537), Oklahoma (-21.603) and Arkansas (-21.603). The states with the greatest positive differences are California (19.338) and West Virginia (16.251). Geographically, Trump did relatively well (compared to New York) in parts of the Northeast and parts of the West. His weakest regions were in the south and the Midwestern states of the Great Lakes region.

The R-squared value indicates that variation in the independent variables included in the model explain 89.36 percent of the variation in the Trump vote. Approximately one-third of the explanatory power comes from county-by-county variation and the other two-thirds comes from state-by-state differences.

### Conclusion

This research reported in this paper attempts to explain the extent to which variation in county-level ideological and socioeconomic conditions can explain variation in the Republican primary vote for Donald Trump, given general state conditions. As predicted by previous work, ideology, as operationalized in this work, does play a role. Counties with large shares of veterans and older

people tended to vote more strongly for Donald Trump while counties with high levels of adherence to evangelical Protestantism and highly educated residents were less likely to vote for Trump. Socio-demographic variables were important too. Counties with higher shares of foreign born residents tended to favor Trump while counties with higher shares of Hispanics were less supportive. Economic rationality seems also to have its role. Counties with higher per-capita personal incomes tended to favor Trump as did counties facing apparent economic stress as measured by high rates of unemployment and poverty. It seems that those in dire economic straits relate their conditions to the political status quo and that they buy Trump's promise to make America great again.

### References

- Abramowitz, Alan I. "Viability, Electability, and Candidate Choice in a Presidential Primary Election: A Test of Competing Models." *The Journal of Politics* 51, 4 (1989): 977-992.
- Adkisson, Richard V., and James Peach. "Voting for President: Elections along the US-Mexican Border." *Journal of Borderlands Studies* 14, 2 (1999): 67-79.
- Adkisson, Richard V. and Jim Peach. "The Determinants of the Vote for Trump: An Analysis of Texas 2016 Primary Results." *Applied Economic Letters*, (2017).  
<http://dx.doi.org/10.1080/13504851.2017.1307917>
- Aldrich, John H., and R. Michael Alvarez. "Issues and the Presidential Primary Voter." *Political Behavior* 16, 3 (1994): 289-317.
- Campbell, James E. "Forecasting the Presidential Vote in the States." *American Journal of Political Science* (1992): 386-407.
- Dowdle, Andrew J., Randall E. Adkins, and Wayne P. Steger. "The Viability Primary Modeling Candidate Support before the Primaries." *Political Research Quarterly* 62, 1 (2009): 77-91.

- Downs, Anthony. "An Economic Theory of Political Action in a Democracy." *The Journal of Political Economy* (1957): 135-150.
- Fair, Ray C. "Econometrics and Presidential Elections." *The Journal of Economic Perspectives* 10, 3 (1996): 89-102.
- Gurian, Paul-Henri, Nathan Burroughs, Lonna Rae Atkeson, Damon Cann, and Audrey A. Haynes. "National Party Division and Divisive State Primaries in US Presidential Elections, 1948–2012." *Political Behavior* (2016): 1-23.
- Holbrook, Thomas M. "Presidential Elections in Space and Time." *American Journal of Political Science* (1991): 91-109.
- Kenney, Patrick J., and Tom W. Rice. "The Relationship between Divisive Primaries and General Election Outcomes." *American Journal of Political Science* (1987): 31-44.
- McLaughlin, Elliott C. "10 of the Most Bizarre Elections in American History." CNN Politics (March 14, 2016)  
<http://www.cnn.com/2015/10/30/politics/interesting-u-s-elections/>
- Mayer, William G. "Retrospective Voting in Presidential Primaries." *Presidential Studies Quarterly* 40, 4 (2010): 660-685.
- Norrander, Barbara. "Ideological Representativeness of Presidential Primary Voters." *American Journal of Political Science* (1989): 570-587.
- Norrander, Barbara. "The Attrition Game: Initial Resources, Initial Contests and the Exit of Candidates during the US Presidential Primary Season." *British Journal of Political Science* 36, 3 (2006): 487-507.
- Steger, Wayne. "Polls and Elections: Two Paradigms of Presidential Nominations." *Presidential Studies Quarterly* 43, 2 (2013): 377-387.
- Steger, Wayne P. "Who Wins Nominations and Why? An Updated Forecast of the Presidential Primary Vote." *Political Research Quarterly* 60, 1 (2007): 91-99.
- Walker, David A. "Predicting Presidential Election Results." *Applied Economics* 38, 5 (2006): 483-490.