That’s Why the Lady Lost to the Trump: Demographics and the 2016 Presidential Election

AARON C. WEINSCHENK
Political Science University of Wisconsin-Green Bay, Green Bay, WI, USA

I examine the relationship between demographics and vote choice using pre-election tracking polls, exit polls, and data on the American states. A number of important findings emerge: (1) there was a fair amount of preference stability in 2016; (2) Clinton underperformed Obama in 83% of demographic groups examined; (3) at the state level, the percent of whites with low levels of education and the size of the rural population had important effects on vote choice, and both of those variables were more strongly correlated with Democratic vote share in 2016 than in 2012; and (4) while state turnout rates were moderately correlated with Clinton’s vote share, changes in state turnout rates from 2012 to 2016 did not have a substantial impact on Clinton’s performance.

KEYWORDS demographics, electoral behavior, vote choice, 2016 presidential election

INTRODUCTION

The election of Donald Trump in 2016 sent shockwaves across the United States and the world. Indeed, many political scientists, pundits, pollsters, political observers, campaign professionals, and ordinary citizens anticipated that Hillary Clinton would win the election.\(^1\) Prior to the election, it was not difficult to find articles—some published months or even years before Election Day—declaring that Democrat Hillary Clinton would prevail...
over Republican Donald Trump. A May 2016 article published by Salon had the title “Donald Trump will not be president: History, polling data, and demographics all point to a single result.” In June 2016, The Guardian ran an article with the headline “Trump won’t win. In fact, the US could be on the brink of a liberal renaissance.” The US News & World Report published an article in March of 2016 with the headline “Why Trump Won’t Win.” A June 2014 article in The Washington Post, though careful not to forecast a winner, noted, “The simple fact is that, in 2016, the Democratic candidate—I assume that will be Hillary but it could be someone else and not make a huge difference—starts at a significantly higher point in terms of electoral votes than whomever Republicans nominate.”

Arguments about why Trump was likely to lose the election focused on a variety of different ideas. For example, some argued that Trump was an inexperienced candidate facing a highly experienced opponent, which put him at a severe disadvantage. Others noted that there were large campaign spending disparities between the Trump and Clinton campaigns, with the Clinton campaign having an advantage over the Trump campaign. A number of political observers pointed out that the Clinton campaign had a far more organized ground game apparatus, especially in battleground states. Many people noted that Donald Trump held inconsistent positions on many issues, which would likely make it difficult to “rally the base” around his candidacy. Furthermore, questions about Trump’s temperament and conduct were raised throughout the campaign. On top of these arguments was the observation that most polls conducted during the general election indicated that Clinton would be victorious.

It is important to note that many explanations about the likely outcome on Election Day 2016 also focused on the role of demographic factors. The goal of this paper is to shed light on the role of demographic factors in 2016. Interestingly, the phrase “demography is destiny” was used repeatedly during the 2016 election. It was well established that the 2016 electorate would be the most racially and ethnically diverse in the history of the United States, and given the high levels of support among many minority groups for Democratic candidates—Obama won 93% of the African American vote, 71% of the Hispanic vote, and 73% of the Asian vote in 2012—the increasing diversity of the US electorate was thought to be a positive development for the Democratic Party in 2016. Indeed, some political analysts noted that the Democratic Party had a “stranglehold” on the Electoral College due, in large part, to the increasing diversity across the United States. Democrats, it was thought, had developed a “blue wall” of support that put Republicans at a natural disadvantage before the 2016 election even began.

Despite discussions about the positive implications of growing diversity for Democrats, some political observers expressed trepidation
about extent to which Hillary Clinton would be able to maintain or improve upon the support of voters who made up the “Obama coalition.” Indeed, Obama was able to bring together a unique mix of different voting groups (e.g., women, Hispanic/Latinos, African-Americans, young people, whites living in Northern states), which undoubtedly helped him secure victory in 2012 and 2008. Throughout the 2016 election, pre-election polls indicated that across many different demographic groups, Clinton was struggling to attain the levels of support that Obama did in the previous two elections.8

The title of an article published by US News & World Report just two months prior to the election read “Hillary Clinton Is No Barack Obama: The Obama coalition may not turn out for his chosen successor.” Come Election Day 2016, would Clinton be able to match or improve upon the performance of Barack Obama in 2012 among key demographic groups in the electorate?

In this paper, I analyze the role of demographics in shaping vote choice during the 2016 election. Political scientists have long demonstrated that demographics play an important role in how people vote and that in some cases demographic attributes become more closely connected to vote choice over the course of campaigns (Kaufmann 2006; Olson and Green 2006; Stonecash 2006; Gimpel and Karnes 2006; McClurg and Holbrook 2009; Kaplan, Park, and Gelman 2012; Erikson, Panagopoulos, and Wlezien 2010). Building on previous research on the connection between demographics and vote choice, I provide a look at the preferences of different groups of voters in 2016. The first part of this paper is descriptive in nature. Using pre-election tracking polls conducted throughout the campaign, I provide a look at levels of support for the candidates among different demographic groups and whether the preferences of different groups were relatively stable during the campaign or whether they shifted as the election unfolded. In addition, I use exit poll data to compare the performance of Hillary Clinton in 2016 to the performance of Barack Obama in 2012 among key demographic groups, with the goal of identifying where Clinton over and underperformed relative to 2012. In the second part of the paper, I develop a state-level model in order to examine how recent demographic and economic changes in the American states impacted candidate performance in 2016 relative to 2012. Before proceeding, it is worth pointing out that prior to the 2016 election, one political reporter noted that “It’s not clear—at least to me—how Republicans will pick the Democratic lock on the electoral college but history suggests they will, eventually, find a way.” As we now know, the Republican Party was able to find a way to pick the Democratic lock—though it seems fair to say, given the widespread view that Clinton would be victorious in 2016, that it happened much earlier than many expected. Although elections are multifaceted events and the outcomes are shaped by numerous forces,
examining the role of demographics will help tell part of the story of how Donald Trump was able to defy expectations and win the 2016 presidential election.

DEMOGRAPHICS AND VOTE CHOICE DURING THE 2016 ELECTION: STABILITY OR CHANGE?

A natural starting place for understanding the impact of demographics on voter preferences during the 2016 election is to consider pre-election polling data collected at different points throughout the course of the campaign. Fortunately, pre-election tracking polls conducted by The Washington Post-ABC News allow for an analysis of what the preferences of different demographics looked like at numerous times during the election. The Washington Post-ABC News conducted polls in June (20–23), July (11–14), August (1–4), September (5–8 and 19–22), and October (10–13). In Figure 1, I plot levels of support for Hillary Clinton (percent saying they will vote for Clinton) for each of the aforementioned dates. Importantly, I plot voter preferences for a variety of different groups in the

The graphs in Figure 1 capture a number of key attributes, though certainly not all, that have been shown to influence how people vote. I supplement the six Washington Post-ABC News tracking polls with data on the preferences of each group on Election Day (using data from the 2016 national exit polls). This provides a more complete picture of where the preferences of different groups started (in early summer) and where they ended up on Election Day. The purpose of Figure 1 is twofold. First, it provides a look at which groups were the most and least supportive of Hillary Clinton. Second, it provides a sense of whether the preferences of different groups were characterized by stability throughout the campaign or whether preferences shifted as the election unfolded. Previous research on the evolution of vote preferences among different groups during presidential campaigns (see, e.g., Erikson, Panagopoulos, and Wlezien 2010) has found that there is a fairly high amount of stability in vote choice, though preferences are sometimes responsive to campaigns, events, or political conditions (Panagopoulos 2012, 2013). In part, the stability in voter preferences stems from the fact that vote choice is heavily influenced by fundamental factors like partisanship and ideology (Erikson and Wlezien 2010; Kaplan, Park, and Gelman 2012). It is worth noting, though, that 2016 election was different than previous elections in a number of important respects. Perhaps the most notable difference was the fact that Donald Trump was not a typical Republican presidential candidate—he had no political experience, lacked a consistent ideology, and lacked backing from many within the Republican establishment. Trump’s status as an atypical Republican candidate had the potential to alter how some groups responded to the Republican Party in 2016. For example, given that Trump’s conservative credentials were questioned by numerous party elites and conservative media personalities, would conservative voters support Trump on Election Day? Information acquired about candidate attributes, policy positions, and ideological views as campaigns unfold can surely lead voters to decide to support a particular candidate or to change their minds about who to vote for on Election Day (Hillygus and Jackman 2003).

A look at Figure 1 illustrates that some groups expressed high levels of support for Hillary Clinton while others did not. For example, Democrats, liberals, people living in urban areas, and non-whites consistently expressed high levels of support for Clinton (minimum level of support among these groups was 58 percent, with the range running from 58 to 92 percent). On the other hand, Republicans, conservatives, and whites without college degrees consistently expressed fairly low levels of support for Hillary Clinton. A number of different groups—Independents, moderates, and 18–39 year olds, for example—fell somewhere in the middle in terms of support for Clinton. Among Independents, for example, the highest level of support for Clinton in the data series shown in Figure 1 was
45 percent (October 13 poll) and among 18–39 year olds, the highest level of support for Clinton was 55 percent (October 13 poll). In general, Figure 1 paints a picture of fairly high levels of stability in terms of the preferences of different groups, although it is clear that the preferences of some groups were more variable than others. The preferences of urban voters, Democrats, whites, and whites without college degrees, for example, hardly moved during the campaign. Other groups exhibited more movement in their preferences during the campaign, though levels of support among the groups shown in Figure 1 generally do not exhibit wild swings during the course of the election. It is worth noting that some groups clearly moved away from Clinton (and toward Trump) during the campaign. For example, the trend for conservatives clearly illustrates a change over the six months shown in the graph. In the June 23 Washington Post-ABC News tracking poll, 29 percent of conservatives said that they intended to vote for Clinton. That number dropped by about 13 percentage points over the course of the 2016 election. In the end, 81 percent of conservatives said they voted for Trump (and 16 percent said they voted for Clinton). Although aggregate level data do not allow for an analysis of why conservatives became more supportive of Trump as Election Day approached, it seems possible that there may have been a “bandwagon effect” among conservatives. In short, even though many conservative voters did not view Trump as a true conservative, they preferred to jump on the Trump bandwagon rather than cast a vote for Clinton. As one Republican voter—quoted in a Washington Post article—noted during the campaign, “I didn’t want to vote for Trump. With Trump, you just hold your nose.” The preferences of suburban voters also moved around a bit during the course of the campaign (level of support ranged from 40 percent to 54 percent). In the June 23 tracking poll, 54 percent of voters in this category said they intended to vote for Clinton. On Election Day, exit poll data revealed that 45 percent of voters in the suburban category voted for Clinton.

The results shown in Figure 1 are consistent with previous research on the dynamics of vote preferences. Using individual-level survey data from the 2008 election and a series of probit models of vote choice, Erikson, Panagopoulos, and Wlezien (2010) find that “the estimated effects of most demographic characteristics remained stable over the duration of the campaign” (p. 488). They, too, find that political attributes like ideology and partisanship evolve of the course of the campaign and note that “vote preferences eventually evolved to coincide with… partisan identities” (p. 488).

Another way of looking at preference dynamics in 2016 is to correlate the results (the vote preferences of each demographic group) from each poll against the results from the other six polls. Overall, there is a strong
relationship between the preferences of different groups in each poll when compared with the preferences in all of the other polls in the dataset. The correlation values range from $r = .974 \ (p < .05)$ to $r = .996 \ (p < .05)$, indicating a high level of stability across the polls. Even the preferences among different groups in the earliest poll (June 2016) correlate at an extremely high level with the preferences of those groups measured on Election Day ($r = .983, \ p < .05$).

WHAT EXIT POLLS TELL US ABOUT CHANGES FROM 2012 TO 2016

Although the data presented above provide an interesting look at overall levels of support and preference dynamics (or lack thereof) among different demographic groups in the electorate, Figure 1 does not provide a sense of how the candidates in 2016 preformed among different groups relative to the past. Fortunately, there is high-quality exit poll data available for 2016 and for previous elections (with the same questions used over time).\textsuperscript{14} Given the concern expressed above about whether Hillary Clinton would be able to sustain or improve upon the support of groups that made up Obama’s coalition of voters, I focus on comparing the vote choices of different groups in 2016 to the vote choices of those groups in 2012.\textsuperscript{15}

In Figure 2, I provide a comparison between Clinton’s performance in 2016 to Obama’s performance in 2012 among different demographic groups in the electorate. The exit polls in 2016 and 2012 provide more demographic information than the Washington Post-ABC News tracking polls used above, so Figure 2 offers a more detailed examination of demographics and vote choice. The groups capture a number of different attributes, ranging from political traits (e.g., ideology and partisanship) to religious preferences to race and ethnicity. The groups in Figure 2 capture some of the important ways of categorizing the electorate. For each group displayed in Figure 2, I subtract the percentage of people saying they voted for Obama in 2012 from the percentage of people saying they voted for Clinton in 2016. Thus, bars to the left of the line indicating zero mean that Clinton underperformed relative to Obama in 2012 among the corresponding group. Bars to the right of the zero line indicate that Clinton did better in 2016 among the corresponding group than Obama did in 2012. It is important to note that just because Clinton underperformed Obama within a given demographic group does not mean that she lost among that group (among some groups, like Asian Americans, support for Clinton remained very high—she won 65% of the Asian vote—even though she lost ground relative to Obama in 2012). Similarly, just because Clinton did better than Obama among some groups in the electorate does not mean that she
won that group (Clinton won 8% of the Republican vote, which is an improvement on Obama’s 6%, but still represents a massive loss among that group). The colors of the bars in Figure 2 indicate which candidate in 2016 won each of the groups.

The general pattern in Figure 2 is fairly easy to identify: Clinton underperformed relative to Obama in 2012 among most demographic groups. Indeed, there are 35 different groups shown in Figure 2 and Clinton underperformed Obama in 29 of the groups (or about 83% of the groups). While some of the groups make up a small portion of the electorate (e.g., those in the “other religion” category only represent 7% of the electorate), some of the groups represent large and/or growing segments of the electorate (e.g., those who identify as Latino/Hispanic represent 11% of the electorate, which is an increase since the 2012 election).

A number of bars shown in Figure 2 stand out as places where Clinton’s underperformance relative to Obama in 2012 was quite substantial. For example, Clinton did significantly worse than Obama among Hispanic/Latino women (Obama got 76%; Clinton got 69%), union households (Obama got 58%; Clinton got 51%), 25–29 year olds (Obama got 60%; Clinton got 54%), Latinos (Obama got 71%; Clinton got 66%), 18–24 year olds (Obama got 60%; Clinton got 56%), Asian Americans (Obama got 73%;...
Clinton got 65%), and African Americans (Obama got 93%, Clinton got 89%). Interestingly, many of these groups are ones that political observers would point to as being part of Obama’s coalition in 2012 and 2008. As Teixeira, Halpin, and Griffin (2015) pointed out before the 2016 election, “If Democrats are to retain the presidency in 2016, they will need to successfully transfer the enthusiasm and support of the Obama coalition to a new candidate and overcome the wider belief that the party had its shot for eight years and that it is now time for a change.” They went on to note that “The party must also take seriously the need to knit together its more diverse coalition with a larger share of working class whites if it wants to be competitive…” (Teixeira, Halpin, and Griffin 2015).

It is also noteworthy that although Clinton did improve upon Obama’s 2012 performance among a number of groups (six of the 35 groups in Figure 2, she only won among two of the groups where she exceeded Obama (those who identify as LGBT and those who report being Jewish—both groups that have historically expressed high levels of support for Democrats to begin with). Importantly, many of the groups where Clinton improved upon Obama make up very small segments of the electorate. Indeed, according to the 2016 exit polls, just 3% of voters identified at Jewish, only 5% of voters identified as LGBT, and 16% of voters were in the 65+ age category. In addition, although Clinton improved upon Obama’s 2012 performance among a few groups shown in Figure 2, she still lost many of them by substantial amounts. For example, Clinton lost to Trump among Republicans by a margin of 80 percentage points, she lost to Trump among white women by a margin of 9 percentage points, and she lost to Trump among voters who are 65+ older by a margin of 7 percentage points.

DEMOGRAPHICS AND VOTE CHOICE IN THE AMERICAN STATES

Up until this point, I have used aggregate polling data to understand how demographics influenced vote choice during the 2016 presidential election. Despite the value of pre-election surveys and exit poll data for understanding elections, such information represents just one way of studying demographics and voter preferences. It is also possible to leverage data on the outcomes in the American states to get an even more comprehensive picture of how demographics mattered during the 2016 election. States are useful units of analysis for studying presidential elections since states are the basis for allocating Electoral College votes. In this section, I develop a model of candidate vote share using states as the units of analysis. In particular, I am interested in examining whether state-level demographic variables impacted candidate performance in the 50 states. I focus on many of the demographic attributes examined above, but also consider a number
of economic measures that received attention during the campaign (e.g.,
employment in manufacturing, state unemployment, etc.).

As a starting point for examining state-level outcomes, Figure 3
provides an overview of candidate performance in 2016 relative to 2012.16
For the sake of simplicity, I focus on Clinton’s share of the two-party vote.
The bars in Figure 3 indicate the extent to which Clinton improved upon
Obama’s performance (or lost ground) in the previous presidential election.
In short, I subtract Obama’s 2012 vote share in each state from Clinton’s
vote share in each state in 2016. Bars to the left of the zero line indicate
that Clinton did worse than Obama in a given state and bars to the right of
the zero line indicate that Clinton did better than Obama in a given state.
As was the case above, just because Clinton underperformed Obama does
not necessarily mean that she lost a given state. Of course, instances of
over performance do not necessary translate into victories either.

Overall, the state-level data displayed in Figure 3 indicate that in the
majority of states, Clinton underperformed relative to Obama in 2012. In
some states, differences in Clinton and Obama vote share are quite
substantial. In total, Clinton did worse than Obama in 38 states (or 76% of
the states). Although Clinton did still win some of the states where she
underperformed Obama, the inability to sustain or improve upon Obama’s
2012 performance in the American states was a severe problem for the
Clinton campaign. It is worth noting that the majority of the battleground

![Figure 3](image-url)

**FIGURE 3** Comparing Clinton 2016 performance to Obama 2012 in the American States.
states (e.g., Florida, Pennsylvania, Ohio, Iowa, Michigan, Nevada, New Hampshire, North Carolina, and Wisconsin) are located within the list of 38 states where Clinton underperformed relative to Obama.

Although the changes displayed in Figure 3 are interesting to examine, they lead to an important question: What explains Clinton’s performance across the states in 2016? In order to better understand the sources of Clinton’s performance, I constructed an original dataset containing information on candidate vote share, state demographic attributes, and a number of economic variables. The goal is not to examine every possible demographic group or economic indicator but instead to look at some of the key factors that previous research has identified as being important to candidate vote share or that received attention during the 2016 election (e.g., whites with low levels of education, states with high levels of manufacturing employment, etc.). I should note that I do not focus on campaign-related measures in this paper, but strongly encourage future research on the role of campaigns (e.g., field offices, campaign spending, candidate visits, etc.) in 2016. I focus on the following state-level measures: Obama’s vote share in 2012, state ideology (higher values indicate more conservative states), the Democratic advantage in partisanship (percent Democrat identifiers minus percent Republican identifiers in each state), state unemployment rates (relative to the national unemployment rate), the number of manufacturing jobs (per 1000 jobs) in each state, the Hispanic/Latino share of the CVAP (Citizen Voting Age Population), the percentage point change in the Hispanic/Latino CVAP, the percentage of whites with less than a high school degree or a high school degree, the percentage of white men with less than a high school degree or a high school degree, the percent of the state population living in a rural area, and the percent of the rural population that is white. I expect positive relationships between Clinton’s 2016 vote share and the following variables: Obama’s 2012 vote share, Democratic advantage in partisanship, the Hispanic/Latino share of the CVAP, and the change in the size of the Hispanic/Latino CVAP. I expect negative relationships between Clinton’s 2016 vote share and the following variables: state ideology, state unemployment, the number of manufacturing jobs (Trump heavily campaigned on bringing manufacturing jobs back to the United States), whites with low levels of education (Trump heavily campaigned on the idea that some people were getting overlooked by the system—a message likely to appeal to those, such as low-educated whites, facing economic stress/uncertainty), white men with low levels of education (Trump’s campaign featured populist and masculine messages), the size of the rural population (Trump’s campaign messages about some communities being left behind and not getting their fair share likely resonated with rural voters), and the percentage of the rural population that is white.
To get a preliminary sense of how these variables relate to electoral outcomes in the states, Table 1 provides a correlation matrix. In addition, in order to help visualize the relationships Figure 4 contains a series of scatterplots where each of the measures described above is plotted against Clinton’s share of the two-party vote in each state. In each graph, I include a line that indicates the linear fit for each relationship (in the first plot in Figure 4, the line is not the linear fit but instead represents the 45-degree mark in order to allow a look at Clinton’s under and over performance relative to Obama).

A look at Table 1 and Figure 4 indicates that many of the abovementioned variables are strongly related to Clinton’s performance in the states. It is worth noting that all of the relationships shown in Figure 4 (and in Table 1) are bivariate relationships. Predictably, Obama’s 2012 vote share \((r = .94, p < .05)\), state ideology \((r = -.91, p < .05)\), and state partisanship \((r = .89, p < .05)\), are highly correlated with Clinton’s 2016 vote share. The state unemployment rate, a possible measure of economic discontent, is weakly correlated with Clinton’s vote share \((r = .19, p = .19)\). Interestingly, the number of manufacturing jobs (per 1000 jobs) in the states is correlated with Clinton’s vote share at \(r = -.40 (p < .05)\), indicating that Clinton did worse in states with high numbers of manufacturing jobs. The race and ethnicity measures are all moderately correlated with Clinton vote share. The size of the Hispanic/Latino population correlates with Clinton vote share at \(r = .35 (p < .05)\), the change in the size of the Hispanic population from 2012 to 2015 correlates with Clinton vote share at \(r = .43 (p < .05)\), the percentage of whites with less than a high school degree or a high school degree correlates with Clinton vote share at \(r = -.44 (p < .05)\), and the percentage of white men with less than a high school degree or a high school degree correlates with Clinton vote share at \(r = -.52 (p < .05)\). In addition, the percent of the population that lives in a rural area correlates with Clinton vote share at \(r = -.53 (p < .05)\) and the percent of the rural population that is white correlates at \(r = -.21 (p = .14)\).

Although the bivariate relationships presented in Table 1 and Figure 4 provide an interesting first cut at understanding how demographic and economic factors relate to electoral outcomes in 2016, it is important to examine whether the relationships persist when the variables are considered simultaneously. In Table 2, I present the results from an OLS regression model where Clinton’s share of the two-party vote is used as the dependent variable. Importantly, I include Obama’s 2012 vote share as an independent variable in order to examine changes from 2012 to 2016. I use the other variables shown in Figure 4 as independent variables. I omit the percent of the rural population that is white and simply use the percent of the population that is rural. In addition, I omit the percent of white men with a low level of education and just use the overall measure of the percentage
### TABLE 1 Correlation Matrix of State-Level Measures

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Clinton 2016 vote share</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Obama 2012 vote share</td>
<td>0.94</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Ideology</td>
<td>−0.91</td>
<td>−0.92</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Democratic advantage</td>
<td>0.89</td>
<td>0.91</td>
<td>−0.88</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Unemployment rate</td>
<td>0.19</td>
<td>0.14</td>
<td>−0.18</td>
<td>0.30</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Manufacturing jobs per 1000</td>
<td>−0.40</td>
<td>−0.53</td>
<td>0.41</td>
<td>−0.26</td>
<td>−0.08</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) Hispanic population</td>
<td>0.35</td>
<td>0.22</td>
<td>−0.24</td>
<td>0.29</td>
<td>0.31</td>
<td>−0.49</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) Change in Hispanic population</td>
<td>0.43</td>
<td>0.29</td>
<td>−0.51</td>
<td>0.30</td>
<td>0.38</td>
<td>−0.57</td>
<td>0.82</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) Pct. White low education</td>
<td>−0.44</td>
<td>−0.32</td>
<td>0.35</td>
<td>−0.13</td>
<td>0.34</td>
<td>0.49</td>
<td>−0.21</td>
<td>−0.33</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10) Pct. Low education White men</td>
<td>−0.52</td>
<td>−0.37</td>
<td>0.39</td>
<td>−0.20</td>
<td>0.28</td>
<td>0.48</td>
<td>−0.27</td>
<td>−0.39</td>
<td>0.97</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11) Pct. rural population</td>
<td>−0.53</td>
<td>−0.37</td>
<td>0.38</td>
<td>−0.41</td>
<td>−0.26</td>
<td>0.42</td>
<td>−0.55</td>
<td>−0.73</td>
<td>0.41</td>
<td>0.56</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>(12) Pct. Rural Population White</td>
<td>−0.21</td>
<td>−0.10</td>
<td>0.01</td>
<td>−0.16</td>
<td>−0.32</td>
<td>0.36</td>
<td>−0.59</td>
<td>−0.50</td>
<td>0.24</td>
<td>0.25</td>
<td>0.32</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**Notes:** Relationships are bivariate. Pearson’s r reported in each cell.
of whites with low educational attainment. Descriptive statistics for all variables included in the model are provided in Table 3.

The results displayed in Table 2 indicate that a number of the bivariate relationships from Figure 4 diminish once other variables are considered. Even so, a number of state-level attributes have statistically significant effects on Clinton’s vote share. After controlling for other factors, state ideology and the measure of the Democratic advantage in partisanship impacted Clinton’s vote share in 2016 (both statistically significant at \( p < .10 \) level, one-tailed). Interestingly, the percentage of whites with less than a high school degree or a high school degree has a negative and statistically significant (\( p < .05 \), one-tailed) effect on Clinton’s vote share. In addition, after accounting for other factors, as the percentage of a state population living in a rural area increases, Clinton’s vote share declines (\( p < .05 \), one-tailed). Both of these variables exert fairly substantial effects on Clinton’s vote share, as Figure 5 illustrates. After controlling for the other variables in the model, Clinton’s predicted vote share is 52% when the percentage of whites with low levels of education is at lowest value, and Clinton’s predicted vote share is 41% when the variable takes on its highest value. After controlling for other factors in the model, Clinton’s predicted vote share is 49% when the rural variable takes on its lowest value, and Clinton’s predicted vote share is 44% when the rural measure takes on its highest value.
One thing worth noting is that both of these variables are more strongly correlated with the 2016 election results than the 2012 results. Indeed, in 2016, the variable measuring the percentage of whites with low levels of education correlates with Clinton vote share at $r = -0.44$ ($p < 0.05$); in 2012, the correlation between whites with low levels of education and Obama vote share was $r = -0.32$ ($p < 0.05$). The same pattern exists for the percent rural variable. In 2016, the correlation with Clinton vote share was $r = -0.53$. 

### TABLE 2
Regression Model of Clinton Vote Share in the American States

<table>
<thead>
<tr>
<th>Variables</th>
<th>b/se</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obama 2012 vote share</td>
<td>0.59*</td>
</tr>
<tr>
<td>State ideology</td>
<td>0.12</td>
</tr>
<tr>
<td>Democratic advantage in Partisanship</td>
<td>-9.17+</td>
</tr>
<tr>
<td>State unemployment rate relative to Nat’l</td>
<td>6.16</td>
</tr>
<tr>
<td>Manufacturing jobs per 1000 jobs</td>
<td>0.13+</td>
</tr>
<tr>
<td>% of CVAP Hispanic/Latino</td>
<td>0.09</td>
</tr>
<tr>
<td>Δ Hispanic/Latino population, 2012–2015</td>
<td>0.59</td>
</tr>
<tr>
<td>% of Whites with $&lt; HS$ or HS degree</td>
<td>1.84</td>
</tr>
<tr>
<td>% of Population rural</td>
<td>0.10</td>
</tr>
<tr>
<td>Constant</td>
<td>32.26*</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.94</td>
</tr>
</tbody>
</table>

*Note: *$p < 0.05$, +$p < 0.10$ (one-tailed tests).*

### TABLE 3
Descriptive Statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Min</th>
<th>Max</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton 2016 vote share</td>
<td>24.3</td>
<td>67.4</td>
<td>47</td>
</tr>
<tr>
<td>Obama 2012 vote share</td>
<td>25.4</td>
<td>71.7</td>
<td>49.2</td>
</tr>
<tr>
<td>State ideology</td>
<td>-0.39</td>
<td>0.33</td>
<td>0.04</td>
</tr>
<tr>
<td>Democratic advantage in Partisanship</td>
<td>-31.8</td>
<td>21.7</td>
<td>-2.1</td>
</tr>
<tr>
<td>State unemployment rate relative to Nat’l</td>
<td>-2.6</td>
<td>1.4</td>
<td>-0.33</td>
</tr>
<tr>
<td>Manufacturing jobs per 1000 jobs</td>
<td>25.2</td>
<td>128.1</td>
<td>67.3</td>
</tr>
<tr>
<td>% of CVAP Hispanic/Latino</td>
<td>0.98</td>
<td>44.7</td>
<td>7.6</td>
</tr>
<tr>
<td>Δ Hispanic/Latino population, 2012–2015</td>
<td>0.14</td>
<td>2.0</td>
<td>0.72</td>
</tr>
<tr>
<td>% of Whites with $&lt; HS$ or HS degree</td>
<td>24.5</td>
<td>56.9</td>
<td>39.5</td>
</tr>
<tr>
<td>% of population rural</td>
<td>5.05</td>
<td>61.3</td>
<td>26.4</td>
</tr>
</tbody>
</table>

*Note: $N = 50$ for all variables.*
(\(p < .05\)) and in 2012 the correlation with Obama vote share was \(r = -.37\) \((p < .05)\). The increasing correlation between these demographic variables and vote choice from 2012 to 2016 is consistent with Holbrook’s (2016) finding that over the past several decades many state-level demographic variables have become more strongly connected to vote choice in presidential elections. Interestingly, a look back at Table 1 indicates that nearly all of the independent variables considered in this paper correlate more strongly with state-level vote share in 2016 than in 2012.

**FIGURE 5** Effects of education and rural variables on Clinton vote share.

---

A NOTE ON TURNOUT, DEMOGRAPHICS, AND CLINTON VOTE SHARE

Although many accounts of the 2016 election have focused on state-level demographic attributes, it is worth noting that some analysts have also commented on the role of voter turnout in 2016.\(^{21}\) An article published by *FiveThirtyEight.com*, for example, noted that “Voter Turnout Fell, Especially in States That Clinton Won.”\(^{22}\) Interestingly, an article in *The New York Times* noted, “Since Donald Trump’s shocking victory, much of the political diagnosis has focused on white working-class swing voters, and for good reason. Across the industrial Midwest, white voters who
had supported Obama and previous Democrats abandoned the party for Trump.

The role that turnout played has been harder to figure out. In the initial days after the election, some people focused on the total number of votes cast, which appeared much lower than four years ago. That impression was wrong, though, because a few million absentee and mail ballots had not yet been counted out West. In the end, overall turnout in 2016 would not have changed much from 2012. Yet it's also becoming clear that turnout really was an important part of the 2016 story—and addressing it is crucial to the Democrats' comeback plans.”

An article in *The Washington Post* pointed out that “But part of it [Clinton’s loss] is simply a drop in turnout overall.” Although the overall turnout rate in 2016 (59.3%) was slightly higher than it was in 2012 (58%), it is clear that in some states turnout increased from 2012 to 2016 and in other states turnout rates declined over this period.

For decades now, political scientists have been discussing and studying the role of voter turnout in candidate performance (Radcliff 1994; Erikson 1995; Radcliff 1995; Tucker, Vedlitz, and DeNardo 1986; Hansford and Gomez 2010; Gomez, Hansford, and Krause 2007). Among the media, the conventional wisdom is that “higher turnout benefits Democrats,” although there has been mixed evidence from political scientists regarding this claim. The basic idea behind the argument that there is a connection between voter turnout and candidate vote share has to do with demographics. As Osborn, McClurg, and Knoll (2010) point out, “At the heart of the conventional wisdom is the plausibility of a key assumption—fluctuation in voter turnout is disproportionately a function of two decisions made by low socioeconomic (SES) status voters. Because there is a strong positive relationship between SES and voter turnout at the individual level of analysis, there are fewer untapped votes among high-status voters (who tend to vote Republican) than among the low-status voters (who tend to vote Democratic). Consequently, aggregate fluctuations are more likely to be driven by the decisions of potential voters to (a) turn out in the first place, and therefore, (b) vote Democratic. Or so the logic goes” (p. 213).

What role did turnout (or changes in turnout) play in shaping candidate performance in 2016?

In order to examine the relationship between overall rates of voter turnout and candidate performance, in Figure 6, I plot voter turnout rates in the states (and changes in turnout from 2012 to 2016) against Clinton’s share of the two-party vote. The first plot in Figure 6 shows that Clinton’s vote share increased as state voter turnout increased ($r = .28, p = .05$). This is identical to the relationship between Obama’s vote share and state turnout levels in 2012 ($r = .28, p = .05$). It is worth noting that Osborn, McClurg, and Knoll (2010) found a positive relationship between county voter
turnout and Obama vote share in the 2008 election. Interestingly, the other two plots in Figure 6 indicate that changes in turnout in the states had little to do with Clinton’s performance in 2016 (when measured relative to Obama’s performance in 2012 or as her share of the two-party vote in 2016). In both cases, the correlation is very weak ($r = .16$, $p = .26$ when change in Democratic vote share from 2012 to 2016 is the dependent variable; $r = .06$, $p = .66$ when Clinton share of the two-party vote in 2016 is the dependent variable). Thus, there is little evidence that changes in the overall turnout rates (at least in the 50 states) from 2012 to 2016 had a substantial impact on Clinton’s performance in 2016.27

This analysis obviously provides an admittedly limited look at the impact of turnout in 2016 since it makes use of overall turnout rates in the 50 states. I do not have access to turnout data broken down by different groups within states. Additional insights about the role of turnout in 2016 could be gained by examining the composition (and changes in the composition) of voters who turned out in each state (or in US counties). As additional data on the 2016 election become available, it may be possible to get a better sense of how the composition of those who turned out to vote in states or counties in 2016 compared with the composition of those who turned out in previous elections. A preliminary analysis (using data purchased from Catalist) of the role of changes in the composition of who
turned out in 2016 by Fraga, McElwee, Rhodes, and Shaffner (2017) noted that “Voter turnout among whites—the racial/ethnic group most strongly in Trump’s corner—increased by 2.4 percentage points in 2016 compared to 2012. In stark contrast, turnout among African Americans—the group most loyal to Hillary Clinton and the Democratic Party—fell by 4.7 percentage points nationally. Latinos and Asian Americans, both groups that are generally Clinton supporters, were more likely to vote—by 3.8 and 3.0 percentage points, respectively.” The researchers went on to note that “… the jump in white turnout in key swing states and drop in black turnout may well have handed the presidency to Trump.” Future studies would be well served by investigating the connection between turnout, demographics, and vote choice in 2016 and by comparing the 2016 election to other recent presidential elections.

DISCUSSION AND CONCLUSION

In this paper, I examined the relationship between demographics and vote choice in 2016 in a number of ways. I used data from pre-election tracking polls to study the dynamics of vote preferences among different demographic groups. Consistent with previous research, I found that there was a fair amount of preference stability among different groups in the electorate in 2016. Importantly, some groups, like conservatives, exhibited movement in their vote preferences as the campaign unfolded. I also used national exit poll data to compare the relationships between key demographic variables and vote choice in 2016 to the relationships in the 2012 presidential election. The key finding was that Clinton underperformed Obama in 83% of the demographic groups examined in this paper. Before and during the 2016 election, numerous political observers warned that Clinton would need to sustain and build upon the support of many groups that made up the “Obama coalition” in 2008 and 2012. Relative to Obama in 2012, Clinton lost ground among African Americans, Hispanic/Latino voters, young people (18–24 year olds and 25–29 year olds), moderates, Independents, liberals, and a variety of other groups in the electorate.

In addition to using polling data, I analyzed the relationship between demographic variables and vote share in the American states. I found that a number of demographic attributes, including the percentage of low educated whites in each state and the size of each state’s rural population, had important effects on electoral performance in 2016. The relationships persisted even after controlling for a range of well-known determinants of candidate performance including previous electoral performance, ideology, and partisanship. It is worth noting that the percentage of low educated whites and the size of the rural population were more strongly correlated with Democratic vote share in 2016 than in 2012. Thus, these variables
became more important in 2016 than they were in the previous presidential election.

As a final look at vote choice in 2016, I examined whether voter turnout levels and changes in voter turnout from 2012 to 2016 impacted candidate performance in the states. Many analysts have noted that changes in turnout can impact how candidates perform—primarily through the mobilization or demobilization of voters in certain demographic groups (e.g., low SES voters). I found that overall state turnout rates were moderately correlated with Clinton’s vote share (just as they were in 2012), but uncovered little evidence to support the idea that changes in overall turnout rates from 2012 to 2016 in the 50 states had a substantial impact on Clinton’s performance.

Ultimately, it is clear that certain demographic variables played an important role during the 2016 presidential election. Interestingly, even though the outcome of the 2016 election was unexpected, from the standpoint of demographics 2016 does not appear to look that much different than previous elections. Indeed, many demographic variables were related to vote choice in 2016 in the same ways that they have been in the past. In addition, partisanship played a similar role in structuring vote choice as it has in the past. It is clear that most Republican voters supported their party’s candidate and most Democratic voters supported their candidate. Thus, despite the historic unpopularity of both major party candidates, partisans did not behave in an unusual manner. Perhaps the key difference between 2016 and recent presidential elections is the underperformance of Clinton compared the previous Democratic candidate. As I mentioned above, Obama was a rare political candidate who was able to persuade, mobilize, and energize voters. Being able to get the demographic groups that are most supportive to turn out on Election Day is critically important to winning elections at all levels of government. Although numerous political observers warned the Clinton campaign not to take certain groups of voters for granted during the 2016 election and to work on mobilizing important voting blocs in key battleground states, it appears that—for one reason or another—that advice was not fully heeded. Perhaps political campaigns will learn from the Clinton campaign in future elections.

As a final note, I should mention that although demographic variables are clearly important in presidential elections, demographics alone certainly do not tell the entire story of the 2016 election. As I noted at the outset, elections are complicated and multifaceted events. It is only by looking at different facets of elections and by studying them using different data sources and different levels of analysis that we can develop a comprehensive understanding of electoral processes and outcomes.
NOTES

1. Most forecasts from political scientists predicted a Clinton victory, though one notable exception was Helmut Norpoth who predicted a Trump victory. Details on his forecasting model can be found here: http://primarymodel.com/2016-forecast-full/


5. See this report for details on demographic changes from 2012 to 2016: http://white.pewresearch.org/fact-tank/2016/02/03/2016-electorate-will-be-the-most-diverse-in-u-s-history/


9. Data were gathered by the author from the following website: https://www.washingtonpost.com/wp-stat/polls/postpollarchive.html

10. The Washington Post-ABC News decided to conduct two separate polls in the month of September. This is the only month when they conducted two tracking polls. Note that the dates for the September polls are different (they represent two different samples).

11. I am limited to the demographic breakdowns provided by The Washington Post-ABC News. Thus, the specific categories (e.g., White versus Non-White) that I plot were determined by available data.


13. This quote is from the article “Donald Trump ‘hold your nose’ Presidency” by Chris Cillizza.

14. Exit polls, like any other types of polls, are imperfect. For an overview of how exit polls are conducted and potential issues with exit polling, see the following websites: http://www.edisonresearch.com/behind-numbers-2016-national-election-exit-poll/ and http://www.pewresearch.org/fact-tank/2016/11/02/just-how-does-the-general-election-exit-poll-work-anyway/


16. Data taken from Dave Leip’s Atlas of U.S. Presidential Elections: https://docs.google.com/spreadsheets/d/1D1edaVTHnZNHVU840EPUtilz3CgdGJm39Urx7HM88Pq6Pus/edit#gid=29622862

17. Data on candidate vote share from Dave Leip’s Atlas of U.S. Presidential Elections (link provided above). Ideology is captured using the state-level measure from the American Ideology Project (http://www.americanideologyproject.com), Democratic partisan advantage is measured using 2015 Gallup data (https://www.gallup.com/poll/188969/red-states-outnumber-blue-first-time-gallup-tracking.aspx), state unemployment measure is from 2015 Bureau of Labor Statistics estimates, number of manufacturing jobs measure is from the 2015 Bureau of Labor Statistics estimates, Hispanic/Latino share of the CVAP (Citizen Voting Age Population) is calculated using the American Community Survey 2015 estimates, the percentage point change in the Hispanic/Latino CVAP calculated using the American Community Survey 2012 and 2015 estimates, the percentage of whites with less than a high school degree or a high school degree is from the American Community Survey 2015 estimates, as is the percentage of white men with less than a high school degree or a high school degree, the percent of the state population living in a rural area, and the percent of the rural population that is white.

18. When the unemployment measure is restricted to whites (unemployment rate among whites in 2015), there is still a weak relationship ($r = .23, p = .11$).

19. Note that the results are identical to this specification if one uses the change from 2012 to 2016 as the dependent variable and includes Obama 2012 vote share as an independent variable to capture the baseline from which change is occurring.
20. Although some of the independent variables are highly correlated (e.g., ideology, Obama vote share, and Democratic partisan advantage), the model results are quite similar when the ideology and Democratic partisanship advantage measures are omitted from the model (and the Obama 2012 vote share measure and the other independent variables are left in the model). When the full model (shown in Table 1) is estimated, the mean VIF is 5.50. When the alternative specification is used (ideology and Democratic partisanship omitted), the mean VIF drops to 2.57 and the Obama measure remains significant at $p < .05$ (one-tailed), as does the percent of Whites with low levels of education measure and the rural population measure. The only changes are that the unemployment measure becomes significant at $p < .05$ (one-tailed) and the Hispanic population measure becomes significant at $p < .10$ (one-tailed). The other variables are not statistically significant at conventional levels. Full model results available from the author on request.


22. The article was published only a few days after the election (using preliminary data) and was updated on Tuesday, November 15, 2016.


27. Osborn, McClurg, and Knoll (2010), using county level data to compare the impact of changes in turnout from 2004 and 2008 on Democratic vote share, note that “Consistent with the hypothesis that the Obama campaign benefited in those places where turnout increased, we see a positive and statistically significant coefficient. The substantive effects are impressive, especially considering the range of the dependent variable. For every unit increase in turnout, there is approximately a 0.2-unit increase in voting Democratic. Thus, if turnout in 2008 was 10 points higher in a county than in 2004, Barack Obama’s campaign could expect to outgain John Kerry in the county by a little more than 2%” (p. 223). It would be useful to replicate Osborn et al.’s analysis to see how county turnout (and changes in county turnout) influenced changes in Democratic vote share from 2012 to 2016. Interestingly, as a robustness check of their county-level results, Osborn et al. estimated the effect of turnout change (from 2004 to 2008) on the change in Democratic vote share (also from 2004 to 2008) in the states. They note, “we estimate our models using state-level data. Though the small $n$ in this data source renders the produced coefficients imprecise, the same pattern emerges in these results: Obama received more votes [than Kerry in 2004] but not enough votes to influence the outcome” (p. 224).

28. Only 40 percent of U.S. adults had a favorable assessment of Clinton, and a mere 35 percent viewed Trump favorably (Gallup 2016).


REFERENCES


