Electoral Studies xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

Bad characters or just more polarization? The rise of extremely negative feelings for presidential candidates☆

Dino P. Christenson[a,*], Herbert F. Weisberg[b]

[a] Department of Political Science, Boston University, 232 Bay State Rd, Boston, MA, 02215, USA
[b] Department of Political Science, Ohio State University, 2140 Derby Hall, 154 N. Oval Mall, Columbus, OH, 43210, USA

A B S T R A C T

What explains the recent rise in extremely negative feelings towards presidential candidates? Using the American National Election Studies survey data from 1984 to 2016, we show that negative feelings towards presidential candidates have grown steadily in recent elections, with unusually large numbers of zero ratings on candidate thermometers in 2004, 2012, and, especially, 2016. Such evaluations are primarily reserved for candidates of the other party and shown to be strongly related to partisan polarization. Importantly, however, candidate traits have long played and continue to play major roles in candidate evaluations, though their effects vary by year. Indeed, the unprecedented number of the most negative scores in 2016 appears due more to increases in negative perceptions of candidate leadership, competence and empathy than to polarization. Clinton and Trump are just as much to blame for the public's animosity as the rising tide of polarization.

Partisan polarization has been increasing steadily over recent presidential elections. Originally, there was controversy in the literature as to whether polarization was limited to elites or also extended to the public (e.g., Abramowitz and Saunders, 2008), with Fiorina et al. (2005) arguing that the movement was due to partisan sorting and that most voters were still centrists even though they had to react to more extreme candidate choices. However, there is now general recognition that public reactions to the parties have become polarized (e.g., Greenberg, 2004; Jacobson, 2007; Bafumi and Shapiro, 2009; Iyengar et al., 2012; Mason, 2015; Huddy et al., 2015). The polarization in the public seemingly increased further in the early 2000s, with extreme animosity developing as more partisans gave the opposite party's candidate the most negative rating possible. By 2016, Pew (2016) found that 59% of Republicans put Clinton at the lowest score, with 68% of Democrats doing so for Trump. As further evidence of the extent of partisan dislike in modern times, they reported an April 2016 survey in which just over 40% of the identifiers with each party considered the other party's policies "so misguided that they threaten the nation's well-being."

However, it is important to distinguish between increased partisan polarization and strong opposition to the opposite party's candidate. The 2016 U.S. presidential election was marked by two of the most unpopular major party candidates in recent history. Donald Trump was controversial even within his own party, while the decades of investigations of the Clintons made Hillary Clinton's candidacy controversial, even though she was relatively popular during the early years of her service as Secretary of State. Pew (2016) found that Republicans gave Clinton an average thermometer rating of 12, while Democrats gave Trump an average of 11. Of course, there have been unpopular candidates before, so understanding popular reactions to the 2016 candidates requires placing them in the context of public evaluations of the presidential nominees over recent decades. In particular, both parties' candidates also received significant numbers of the most negative ratings in 2012 and 2004, permitting comparisons with the 2016 race.

Thus, two clear trends—extremely negative ratings of candidates and partisan polarization—have emerged in recent elections and spiked in 2016 with little understanding of the relationship between the two, if any. We ask whether the increase in negative feelings for the candidates are better explained by the increase in partisan polarization or the characteristics of the particular candidates. That is, do candidate evaluations have anything to do with the candidates themselves or are they merely products of the broader trend in partisan polarization?

To understand the extent to which low evaluations of the 2016 presidential candidates were due to assessments of their characteristics versus increased polarization—given differences across the public's political, economic, and social characteristics—we make use of the American National Election Studies (ANES) surveys from 1984 through 2016. The earlier elections are analyzed as a basis against which 2016 is compared. An important part of this study is a detailed examination of the factors that affect the public's rating of presidential nominees on the feeling thermometer. While the thermometer has become a common

☆ We would like to thank Douglas Kriner and Corwin Smidt for helpful discussions and the anonymous reviewers for insightful comments.
* Corresponding author.
E-mail addresses: dinopc@bu.edu (D.P. Christenson), weisberg.1@osu.edu (H.F. Weisberg).

https://doi.org/10.1016/j.electstud.2019.03.008
Received 29 June 2018; Received in revised form 12 December 2018; Accepted 19 March 2019

measure of assessing candidate popularity in both academic writing and in media surveys, there has not been a systematic study of the effects of candidate characteristics alongside the longer trend in polarization. Such a baseline is necessary in order to understand the seemingly drastic changes in these ratings in recent elections. Furthermore, the particular nature of extremely negative ratings suggest that there are different dynamics for own and other party candidates. An additional contribution of this study is therefore taking seriously the asymmetry in candidate evaluations.

1. Keeping the candidate in candidate evaluations

The early voting behavior literature recognized the importance of candidate characteristics in understanding vote change across presidential elections (Stokes, 1966). Subsequent research by Kinder (1986) brought attention to the role of perceptions of candidate traits, of which four trait factors are usually distinguished: leadership, integrity, competence, and empathy. Since then, a number of studies have compared the perceived traits of the major party candidates in a host of elections and found varying but overall strong effects on vote choice (e.g., Kinder, 1986; Miller and Miller, 1976; Miller and Shanks, 1996; Funk, 1999; Weisberg and Mockabee, 1999; Weisberg and Hill, 2004; Weisberg and Christenson, 2007; Abramson et al., 2007; Holian and Prysby, 2014). As such, they have become a fairly standard component of vote choice models, and even basic elements in early structural models of electoral behavior and candidate evaluations (Markus and Converse, 1979; Markus, 1982; Page and Jones, 1979; Kenney and Rice, 1988; Funk, 1999). In line with these studies, we expect a positive relationship between the standard four traits and overall candidate evaluations.[1]

Of course, there may be some patterns over the years as to which of these four types of traits generally matter more. However, Funk (1999) disproved early expectations that the same traits would be relevant for all candidates, even for candidates of the same party (see also Miller et al., 1986). Instead, different traits are evoked by election campaigns in different years: for example, leadership might be stressed in one election and competence in another. Usually own party candidates will evoke more positive trait ratings than other party candidates, a phenomenon that appears to have increased in recent years due to racial and moral attitude changes (Hetherington et al., 2016). In terms of the unusually extreme scores in candidate thermometers in the more recent elections, we expect those when a candidate is rated low on some or all of these traits, though it is not clear a priori which will have mattered more in these three elections.

There is good reason, however, to expect a generally diminishing role for candidate traits over time. First, recent work looking at over 60 years of open-ended question comments about candidates finds a substantial decrease in the public's spontaneous use of candidate traits, and, moreover, that such perceptions are increasingly tied to party identification (Wattenberg, 2016). Second, and relatedly, general findings of increased partisan polarization imply less explanatory power for candidate specific factors. Thus, while there is an abundance of historical evidence suggesting strong, albeit variable, relationships between candidate traits and overall evaluations as well as vote choice, the recent trends in partisan polarization may be supplanting this link—a possibility we address more directly in the next section.

2. The all powerful partisan polarization?

The role of partisanship in candidate evaluations can only be thought to have increased in recent years, given recognition of the vast effects of partisan polarization, predominantly perceived at the individual level as differences between one's own party's and the other party's thermometer evaluations (e.g., Bafumi and Shapiro, 2009; Greenberg, 2004; Jacobson, 2007; Iyengar et al., 2012; Mason, 2015; Abramowitz and Webster, 2016; Huddy et al., 2015). Unsurprisingly, evaluations of one's own party's candidate have been found to be positive, though some partisans may be less satisfied with their party's nominee in years when the race for that party's nomination was heavily contested. In any case, evaluations of the other party's candidate are generally more negative.

The discussion of polarization began in the voting behavior literature as a debate on whether polarization was limited to elites with most of the public being centrist (Fiorina et al., 2005) or whether there was also deep polarization among the public (Abramowitz and Saunders, 2008). A common ground was found, with recognition that partisan-ideological sorting had occurred, such that party identification and ideological self-identification became more consistent with one another (Levendusky, 2009). As the sorting argument became widely accepted, more nuanced analyses of the relationship of issues, identity, and polarization have been conducted.

Recent work in this vein emphasizes identity over issues. Attitudes of partisans toward the opposite party have become increasingly negative, whereas the effects of economic policy preferences on attitudes toward the out-party have not changed (Iyengar et al., 2012). Similarly, partisan and ideological identification have become more aligned, with this sorting leading to stronger partisan identity, greater differentiation between attitudes toward the two parties, and greater anger toward the other party's presidential nominee, even though issue positions have not become more polarized (Mason, 2015). Of course, issue positions are not necessarily fundamental to people's ideology. Converse (1964) demonstrates that many people do not have tightly constrained issue-based ideology. More recently, Ellis and Stimson (2012) show that ideological self-identification (symbolic, or identity-based) does not necessarily correspond to operational (issue-based) ideology, while Mason (2018) emphasizes that identity-based ideology is associated with increased polarization of ideological groups, even when issue-based ideology is weak.

However, issue-based ideology may not be weak in every contest. If the public perceives increased policy divergence between candidates, the effect may be greater affective polarization, though biographical information about the candidates—commonplace in presidential campaigns—has been shown to mitigate the effect (Rogowski and Sutherland, 2016). Bougher (2017) offers that perceptions of issue disagreement underlie political dislike, and thus the alignment of issues, contemporaneous with partisan polarization, helps explain recent political hostility. In all, the current scholarly landscape paints a complex picture for the relationship of ideology to polarization.

More widely agreed upon in the literature is that the movement in party polarization has largely been in terms of increasingly negative evaluations of the other party, while the public's feelings toward their own party has remained fairly stable—though, as we show below, 2016 broke slightly from this trend. Such dynamics suggest an increased role of negative partisanship. Stemming from multiparty voting behavior studies (Rose and Mishler, 1998; Caruana et al., 2015; Medeiros and Noël, 2014), the concept posits that negative feelings about the other party—as opposed to positive feelings about one's own, which we also expect from affective polarization—are dominant in political judgments and behavior. It has been shown in the US context to increase party loyalty and straight-ticket voting, among other effects (Abramowitz and Webster, 2016). In line with the conclusion we might draw from evidence of negative partisanship, the increase in those identifying as independents also exposes a limitation of affective polarization theory. Though independents increasingly vote as partisans due to polarization (Smidt, 2017), that they can have negative associations of partisanship more generally is not easily explained by simple group attachment (Klar and Krupnikov, 2016; Klar et al., 2018; Bougher, 2017). Indeed, Klar

[1] Additional experimental work by Funk (1996) confirms the causal direction as from specific candidate traits to overall candidate evaluations (for a contrary argument, see Bartels, 2002).

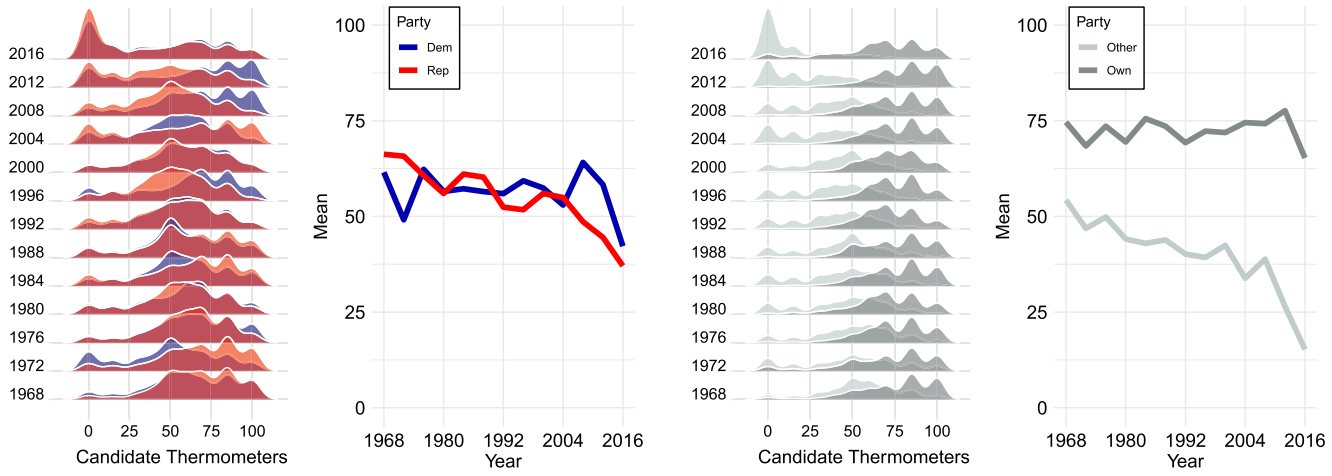


Fig. 1. Candidate thermometer distributions & means by year.

et al. (2018) find that affective polarization is limited to strong partisans in 2016, with many preferring to simply avoid politics.

A potentially big implication from the literature is then that different dynamics may pertain to own and other party evaluations. However, much of the literature that directly addresses the factors underlying candidate thermometer ratings examines the ratings of single candidates (e.g., Funk, 1999) or models the *difference* between own and other party candidates (e.g., Abramson et al., 2007; Jacobson, 2010; Rogowski and Sutherland, 2016; Enders and Armaly, 2018). While studies of this nature provide invaluable insights into candidate evaluations, they do not allow one to determine whether different factors might be driving own and other party candidate evaluations more generally—a consideration that drives our approach in the analyses below.

Conventional wisdom as well as some clear trends over time suggest a powerful role for partisan polarization in candidate evaluations, though the extent to which this is related to group attachments, symbolic identity, or policy positions is still a topic of debate.[2] Regardless, the question remains whether partisan polarization and/or contemporaneous changes in ideology are primarily responsible for the increasingly negative evaluations of candidates—or are the perceived characteristics of the candidates still in play? Moreover, which explains the recent rise in extreme animosity towards the candidates?

3. Electoral trends

To explore our expectations above, we make use of the American National Election Studies (ANES) datasets, both the 2016 Time Series and the Time Series Cumulative Data File, the latter of which combines the common questions from surveys conducted since 1948. The presidential candidate thermometers have been offered since 1968. The left side of Fig. 1 plots the density of respondents' thermometer ratings for the Democratic and Republican Party candidates in each year. Partisan leaners are coded with their respective parties. Evident from the plot is the growing left tail over time. Though not without exception—1972 for the Democrats, in particular—recent years have seen a marked increase in extremely low rating values. By 2016, for both parties the distribution of values have entirely transformed, moving from more moderate and warm peaks (the largest numbers of scores in the center to right) to increasingly larger number of lower scores (on the left).

In terms of the most extreme values, 2016 is a clear outlier. 24% of the ANES sample gave Hillary a 0 as did 32% for Trump, versus previous highs of 14% for McGovern and 13% for W in 2004. Obviously partisan polarization is part of that. 45% of Republican identifiers (not including leaners) gave Hillary a 0. The previous high was 26% for McGovern and for Obama in 2012. While just 4% of Democrats gave Hillary a 0 (versus 9% for McGovern), 60% of Democrats gave Trump a 0 (previous high was 28% for Romney). While only 5% of Republicans gave Trump a 0, this was the first time that more than 1% of them gave their nominee a 0.

Given the strongly partisan nature of these evaluations, our subsequent analyses look specifically at how the public evaluates their own versus the other party's candidates.[3] The right side of Fig. 1 plots the density of respondents' thermometer ratings for their own party's and the other party's candidate for each year. The distribution of ratings for own party candidates had been remarkably similar over most of these elections, with the exception of 1972 when Democrats were more negative toward McGovern. The 2016 distribution also evidences unusually low enthusiasm for own party candidates. Ratings of the other party candidates, however, follow a trend of becoming more negative over the years, especially in 2004 and then even more so in 2012 and 2016. However, the patterns for 2004, 2012, and 2016 are somewhat different. The zero ratings in 2004 were primarily of George W. Bush, though there were also a small number of zero ratings of John Kerry. By 2012, the percentage of zero ratings for the two candidates were more similar. The percentage of zero ratings for both candidates hit records in 2016, with, for the first time since 1972, a perceptible percentage of zero ratings given by partisans to the nominee of their own party. In addition, the 2016 ratings of the other party candidates were drastically more negative than in any of the previous years.

Our primary expectations concern the degree to which candidate traits and ideological distance to the candidates correlate with their overall feelings towards them, particularly with respect to the expectedly large effect of partisan polarization. Before turning to multivariate tests of that nature, we look at their univariate dynamics. Fig. 2 provides density plots and mean values of each of these covariates for each of the years that they are available.[4] As we did with the candidate thermometers above, the plots are constructed for respondents' own

[2] While some literature suggests that the causal direction implied here—from party to candidate—is appropriate (e.g., Campbell et al., 1960; Whiteley, 1988; Green et al., 2004; Achen and Bartels, 2017), it is entirely possible that the relationship is reinforcing (Barber and McCarty, 2015) or that it works, at least partially, in the other direction, given strong evidence of elite polarization (e.g., Jacobson, 2000; Fiorina et al., 2005).

[3] Like the nominal partisanship measure, dividing evaluations into own and other party measures includes the respective partisan leaners.

[4] While the measures comprising the leadership, integrity and competence indices were available by 1980, the empathy measures were not available until 1984. Thus, the models that follow begin in 1984.

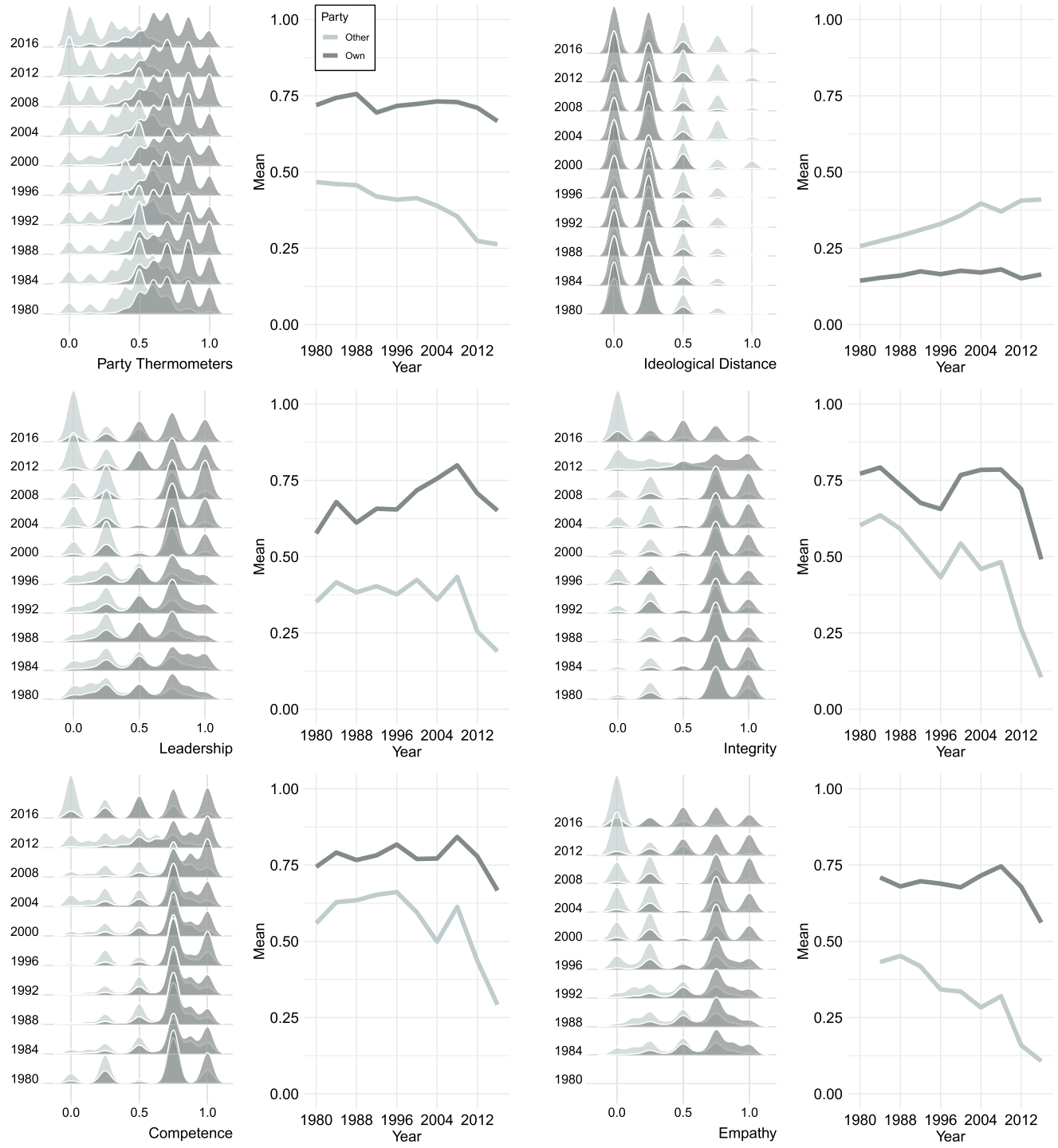


Fig. 2. Key covariate distributions & means by year.

party versus the other party.

The top-left graphs in Fig. 2 illustrate the over time changes in the thermometer ratings for one's own party and the other party (rescaled from 0 to 1). Two major trends are evident. First, partisan polarization—frequently measured as the difference between these scores—has been increasing fairly steadily over time since 1992, with the high level in 2016 only second to 2012. Second, this is primarily due to the growth in low ratings for the other party. Ratings of the own party have been fairly stable over time, though 2016 was an all time low here. In short, polarization has been on the rise across a broad swath of the public, though it may be leveling off. To capture *partisan polarization* in the analyses below, we follow standard practice and take the difference in respondents' own party thermometer less that of the other party.[5] The range of the data then is from −100 to 100, with zero implying that a respondent rates both parties the same. Like all independent variables in our analyses, we have rescaled the measure from 0 to 1.

The trend in *ideological distance* (top-right graphs) provides similar

[5] Enders and Armaly (2018) call this "perceived" partisan polarization, in contrast to "actual," which they operationalize as the difference between the individual rating and that of the sample average.

support for negative partisanship. This variable is created by taking the absolute perceived difference between the respondent and each candidate on a five point strength of ideology scale. Again, we see over time change largely confined to the other party, with the own party distance remaining flat. While the difference in the two was only about .12 in 1980 it steadily rose to almost 0.20 by its current high in 2016. The public perceives the other party's candidate as having moved away from them ideologically, while their own have remained constant. Beginning in the 2000s, the density plot shows that a substantial number of folks see the other candidate as far as possible from them ideologically.

We also explore the standard candidate traits of leadership, integrity, competence, and empathy over this time period. To measure these concepts, we create a *competence* index from subjective survey questions on a candidate's "intelligence" and "knowledge," a *leadership* index on being "inspiring" and a "strong leader," an *integrity* index on "morality" and "honesty," and, finally, an *empathy* index on "compassion" and "care" for "people like me" (Kinder, 1986; Funk, 1996, 1999). All four of the candidate traits saw substantial increases in the difference between own and other party candidate evaluations over time, though the divergence for some occurs more recently than for others. Today, the largest difference between own and other candidates is found on leadership evaluations, followed by empathy. Still, notable differences have developed on the integrity and competence indices as well. In all cases the middle of the distribution has largely moved to the tails for the other candidate, with the general effect on the distribution looking evenly distributed around the mean or nearly flat in the 80s to more bimodal in recent years.

In addition, 2016 stands out as an outlier on most of the evaluations. Evaluations of traits of own party candidates hit record highs in 2008, and even those of other party candidates spiked up that year, but there have been sharp falls since. From 2012 to 2016, the largest drops occur in integrity. Interestingly, the drops here are a bit larger for own party candidates, though nontrivial for other party candidates as well. Other party competence and own party empathy also drop considerably in this four year span. The traits in 2016 for both own and other party candidates show large negative movements that are fairly far from 2012 as well as the cumulative central tendency. In short, the mean perceptions of the leadership, empathy, competence, and, especially, integrity of one's own party candidate and the other party's candidate were much lower than in any of the elections for which we have data.

To emphasize this point further, the public was critical of both major party candidates on the candidate trait questions in 2016. Scoring each item as 1 for saying a trait described the candidate "extremely well" to 0 for "not well at all," the only trait in which either candidate averaged above 0.50 was Hillary Clinton's 0.64 when asked about her knowledge (competence). There was a sharp gap between the candidates on that trait, with Donald Trump's average only being .36. The smallest gap between the candidates was on being a strong leader, with Clinton's .44 rating just slightly above Trump's 0.40. Trump's weakest rating was a 0.28 on "really cares about people like you," with Clinton's 0.38 being several points higher but still well below a neutral 0.5 value. Clinton's weakest rating was on being perceived as honest, though her 0.27 was just a few points below Trump's 0.33. Even Democrats were, on average, negative about Clinton's honesty (0.44), and Republicans gave Trump an average above 0.5 only on leadership (0.60). Partisans gave particularly low trait ratings in 2016 to the other party's nominee: the only average above 0.25 given by partisans was a nearly neutral 0.48 that Republicans gave Clinton on being knowledgeable.

Overall, the trends are consistent with negative partisanship. Across all the covariates, the gap between own and other party evaluations have increased over the more than 30 years observed here, and the bulk of the movement was in more negative evaluations of the other party and other party candidate. However—unlike polarization and ideology, where the own party evaluations were fairly flat—the own and other party traits better reflect one another. That is, many of the own and other dynamics for each evaluation have similar bumps and drops depending on the year, suggesting that own and other party candidate evaluations are not entirely independent from each other. Indeed, the trends are not perfectly linear over time, exhibiting substantial variation both year-to-year and across each of the trait indices. Moreover, 2012, 2016 saw decreases in traits for both own and other party candidates, diverging from the negative partisanship trend. The heterogeneity suggests that while the trend may be one of increased partisan division, not all candidates are equal nor are their evaluations solely products of polarization. We turn to multivariate hypothesis tests of this nature in the next section.

4. Explaining candidate evaluations

In what follows, we test whether perceptions of candidate traits and ideological distance to candidates affect general feelings towards them, as measured by their thermometer ratings. The strongly partisan nature of these evaluations demand an analytical approach that controls for such asymmetry. As in the longitudinal univariate analyses above, our hypothesis tests look specifically at how the public evaluates their own versus the other party's candidates, both in terms of the dependent and independent variables. In this way, we directly control for the strongest correlation with candidate evaluations, partisanship, allowing us to perceive more general trends and relationships.

For purposes of comparison, we do so for the last nine presidential elections, beginning with Ronald Reagan's reelection over Walter Mondale and ending with Donald Trump's election over Hillary Clinton, as well as a pooled model with year fixed effects.[6] Beyond being the most recent elections and therefore the most comparable, a benefit of this selection is that they are nearly balanced in terms of partisanship, in so far as it includes five elections in which Republicans won and four in which Democrats won. While the candidate thermometers span 13 presidential elections, the availability of other covariates are limited to more recent elections. In particular, 1984 was the first year all measures for the four trait indices were available. Due to the significantly larger sample sizes in 2012 and 2016 afforded by internet samples, the pooled model uses only the face-to-face interviews for 2012 and 2016, so as to be more similar to and not overpower the earlier year sample sizes. The single election year analyses use the full sample (i.e., including the internet samples), with a dummy variable to control for potential online mode effects.[7]

In the multivariate analyses, we group the values of the candidate thermometers into seven categories to better represent the underlying distribution of the data. As illustrated in Fig. 1, the distributions are not smooth across the full set of values. That is, the respondents' placements of candidates occurs primarily in a limited number of categories across the full 101 point range, and thus the measure may be better considered as ordinal in practice even though its theoretical motivation is a continuous variable (Rabinowitz, 1976; Alwin, 1992, 1997; Lupton and Jacoby, 2016). Importantly, and despite this lumping into common values, candidate thermometer ratings appear to be appropriate for candidate comparisons (Weisberg and Miller, 1979) and reliable measures (Lupton and Jacoby, 2016). For our variable, the zero category remains zero, while values 1–25 are coded to 1, 26–49 to 2, 50 to 3, 51–74 to 4, 75–96 to 5, and 97–100 to 6.[8]

With seven point ordinal scales on the dependent variables, we use

[6] We include pooled models with linear time trends, instead of year dummies, in the Appendix. The substantive results are unchanged.

[7] The Appendix includes a pooled model with internet samples and a control for online mode effects. The substantive conclusions are unchanged.

[8] The ANES grouped the 97–100 values together in a single category for some early elections—due to missing value conventions at the time—making them inseparable.

ordinal probit to model the expected relationships. However, as discussed above and evident in the right-side of Fig. 1, there is substantial zero inflation in a few of the years for the other party candidates: 2004, 2012 and 2016. Moreover, we are explicitly interested in understanding the (potentially unique) factors that lead to the most negative evaluation for the other party candidates and whether these differ from those that predict higher evaluations. To these ends, we rely on zero inflated ordinal probit models for these years, which use a double-hurdle combination of a split probit and ordered probit (Harris and Zhao, 2007). That is, we account for the high number of zeros in the other party evaluations by assuming that they come from both a probit and ordered probit model. We use the same set of covariates to test whether the factors that lead to higher evaluations also explain the lowest one.

The literature motivates a test of the relationships of partisan polarization, ideology and candidate traits with candidate evaluations. Importantly, comparing separate year and pooled analyses means that our models allow for evidence of polarization to show up in multiple ways. For example, partisan polarization might be at play if ideology is shown to become more important over time, particularly if ideology is responsible for the increase in zero ratings, rather than candidate specific factors. Another indication of polarization would be if year effects in a pooled model were greater in recent elections after controlling for the other factors; i.e., that polarization is essentially or not completely observed by our partisan polarization or ideological distance measures, and thus captured by the increasing explanatory power of the year dummies.

Beyond the candidate specific factors of primary interest, there may also be general patterns of particular types of people being more likely to rate candidates higher or lower, though it would probably be difficult for social-demographic factors to have much influence beyond those of partisan identity and candidate perceptions. Nonetheless, it is worth checking for gender and race effects as well as income, marriage, and religiosity, but the most likely effects would be ones related to age since older citizens are likely to be stronger partisans and hence rate their party's candidate more positively. There may also be systematic effects associated with education since people with more education might be less likely to react with extremely negative responses.

We proceed by discussing the multivariate results in a series of focused sections. We begin by addressing our expectations on the relationships of polarization, ideology and traits both in 2016 and across the previous eight presidential elections as well as a pooled model of all them for context. We do so by first examining the relationships for own party candidates before turning to the other party candidates, the former serving as a baseline for the latter.

4.1. Own party candidates

Table 1 presents the results of the ordinal probit models for the own party candidate evaluations. Looking first at the predictors of thermometer ratings of one's own party's candidate provides a point of comparison for our more central expectations on the ratings of the other party's candidate. Here, the four candidate traits look important overall, but there are differences across elections. In particular, integrity has a significant effect on ratings in only six of the eight preceding elections. Competence, while consistently significant, appears to have a weaker effect, overall, than leadership and empathy. Leadership had the largest effect in most of the early years, with empathy sometimes coming close or even larger, as in 2000, 2008 and 2016. However, in 2016, integrity had the largest effect of these four candidate trait perceptions, so that partisans who were least positive about the integrity of their party's candidate were giving that candidate a lower rating. Also notable in 2016, competence mattered the least in own party candidate evaluations.

Three demographic predictors have significant effects when all the election years are pooled—age, race, and religious guidance—but none is significant in every election. Whites usually give lower ratings to their party's candidate, though they did not do so in 2004 and the effects were substantively marginal in many years. Older people typically give more positive ratings to their party's candidate. Marriage effects are never significant, though religious guidance was consistently positive and significant in both 2000 and in the pooled model. Women were only more likely to support their party's candidate in Bill Clinton's 1996 reelection, but not in Hillary Clinton's 2016 run. The sign of the coefficient on education varies over the years, but there is a significant negative effect in 2012 and 2016 with the more educated respondents being more likely to give their party's candidate a lower rating, perhaps suggesting greater objectivity among the most educated—a possibility we return to in the final results section below. The sign on income is also unstable over the years, but in 2016 those with higher incomes gave their party's candidate a lower rating, though this just fails the conventional cutoff for statistical significance. Indeed, it is interesting that the identity predictors (white and religious) were not effectual for one's own party candidate in 2016, while age and education (and income, nearly) were.

The party thermometer difference capturing partisan polarization is always significant, positive and relatively large: the more highly the person rates her own party versus the other party, the higher she rates her party's candidate on the thermometer. However, there is not much trend to the coefficients. Indeed, rather than an increase in recent elections, the coefficient for 2016 is smaller than it had been in any of the preceding elections. This implies that increased partisan polarization over the years has not had a comparable increasing effect on own party candidate evaluations. Likewise, the year dummies do not show a steady increase over time, though the dummy for 2008—one of the most extremely negative rating years—was considerably larger than the other years. Finally, note too that ideological distance from one's own party's candidate was negative, as expected, but it was only inconsistently significant: partisan considerations trumped ideological factors in evaluating one's own party's candidate.

4.2. Other party candidates

Table 2 presents the results of the ordinal probit and zero inflated ordinal probit models for the other party candidate evaluations. The top half of the table provides coefficients and standard errors for the ordinal probit equation, while the bottom half does so for the inflation equations of the zero inflated models for 2004, 2012 and 2016. Once again, we do not include party identification in the equation because it is handled through the construction of the dependent variable. In general, if thermometer ratings are just thermometer ratings, one would expect similar patterns for ratings of the other party's candidate as we saw with the own party's candidate. Here, however, we find strikingly different relationships.

Looking first at the ordinal equations, we find that whereas integrity mattered least over the years of the traits in rating own party's candidate, competence holds that position for the other party's candidate; that is, it seems easier to discount the competence of the candidate of the other party. As in the previous years, that was also the case in 2016. However, perceptions of integrity mattered substantially more to positive evaluations of other party's candidate in 2012 and, especially, in 2016 than it did in previous election years. Equivalently, perceptions of lack of integrity led to more negative evaluations of the other party's candidate in 2016 than in previous elections. Leadership and empathy had consistently strong and significant effects over the years in evaluating the other party's candidate, and that remained evident in 2016, though the empathy coefficient is clearly on the historically low side.

Some demographic effects are again evident. Whites gave lower ratings to the other party's candidate in nearly every year. Age also has a significant coefficient overall, but only in two of the years, 2012 and 2016. Gender, this time, was not significant in any of the years, including in the Bill Clinton election of 1992. While insignificant, its sign actually reversed in 1988, 2008 and 2016 with women giving lower

Table 1
Own party candidate evaluations.

	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)	
	1984		1988		1992		1996		2000		2004		2008		2012		2016		All	
	b	se	b	se	b	se	b	se	b	se	b	se	b	se	b	se	b	se	b	se
Age	0.703***	(0.160)	0.215	(0.149)	0.418**	(0.133)	0.551***	(0.152)	0.607***	(0.160)	0.660***	(0.187)	0.197	(0.195)	0.110	(0.087)	0.370***	(0.092)	0.362***	(0.052)
Education	0.236*	(0.117)	0.147	(0.108)	0.012	(0.099)	−0.010	(0.116)	0.064	(0.115)	−0.113	(0.141)	0.244	(0.149)	−0.123*	(0.056)	−0.181*	(0.071)	0.001	(0.037)
Female	0.014	(0.063)	−0.069	(0.061)	0.021	(0.055)	0.174**	(0.063)	0.024	(0.062)	−0.103	(0.075)	0.063	(0.077)	0.058	(0.033)	−0.050	(0.038)	0.006	(0.021)
White	−0.103	(0.085)	−0.009	(0.075)	−0.040	(0.068)	−0.394***	(0.079)	−0.124	(0.073)	0.092	(0.084)	−0.350***	(0.085)	−0.364***	(0.037)	−0.085	(0.045)	−0.159***	(0.025)
Income	−0.087	(0.131)	0.177	(0.131)	0.060	(0.118)	−0.183	(0.138)	0.161	(0.130)	0.109	(0.143)	0.174	(0.166)	−0.093	(0.066)	−0.146	(0.076)	0.029	(0.044)
Married	0.106	(0.068)	−0.044	(0.066)	0.000	(0.060)	0.089	(0.070)	0.060	(0.066)	0.104	(0.080)	−0.106	(0.085)	−0.023	(0.036)	0.043	(0.042)	0.016	(0.023)
Religious Guidance	0.146	(0.084)	−0.070	(0.079)	0.083	(0.073)	0.010	(0.083)	0.189*	(0.079)	0.118	(0.096)	0.038	(0.101)	0.033	(0.041)	0.015	(0.047)	0.075**	(0.027)
Ideological Distance	−0.396*	(0.183)	−0.280	(0.172)	−0.490**	(0.167)	−0.064	(0.188)	−0.141	(0.149)	−0.465*	(0.225)	−0.092	(0.209)	−0.256**	(0.096)	−0.274*	(0.109)	−0.166**	(0.059)
Polarization	3.243***	(0.239)	2.378***	(0.221)	2.881***	(0.220)	2.604***	(0.244)	2.123***	(0.233)	1.789***	(0.280)	2.772***	(0.267)	2.973***	(0.128)	1.769***	(0.144)	2.440***	(0.078)
Leadership	2.262***	(0.152)	1.488***	(0.159)	1.485***	(0.160)	1.957***	(0.174)	0.965***	(0.154)	2.122***	(0.199)	1.212***	(0.209)	1.185***	(0.098)	1.192***	(0.101)	1.497***	(0.055)
Integrity	0.591**	(0.185)	0.510**	(0.166)	0.672***	(0.105)	0.136	(0.112)	0.819***	(0.156)	0.397	(0.205)	0.675**	(0.217)	1.219***	(0.118)	1.393***	(0.093)	0.652***	(0.049)
Competence	0.900***	(0.203)	0.845***	(0.213)	0.614**	(0.194)	0.580**	(0.200)	0.508**	(0.178)	0.702***	(0.205)	1.231***	(0.285)	0.469***	(0.122)	0.746***	(0.086)	0.696***	(0.064)
Empathy	1.419***	(0.177)	1.116***	(0.186)	1.465***	(0.161)	1.600***	(0.161)	1.162***	(0.137)	1.173***	(0.182)	1.231***	(0.188)	1.273***	(0.094)	1.175***	(0.099)	1.327***	(0.054)
Online Sample															−0.111**	(0.037)	−0.016	(0.042)		
1988																			0.024	(0.042)
1992																			−0.392***	(0.040)
1996																			−0.170***	(0.043)
2000																			−0.341***	(0.043)
2004																			−0.300***	(0.047)
2008																			−0.459***	(0.049)
2012																			0.123**	(0.043)
2016																			−0.304***	(0.050)
cut 1	2.031***	(0.260)	1.197***	(0.225)	1.568***	(0.211)	1.118***	(0.233)	0.988***	(0.236)	1.114***	(0.286)	1.921***	(0.287)	0.846***	(0.117)	0.639***	(0.121)	1.038***	(0.080)
cut 2	2.587***	(0.241)	1.477***	(0.216)	2.029***	(0.199)	1.591***	(0.223)	1.600***	(0.215)	1.655***	(0.258)	2.268***	(0.277)	1.354***	(0.109)	1.335***	(0.118)	1.505***	(0.076)
cut 3	3.521***	(0.233)	2.189***	(0.207)	2.995***	(0.194)	2.327***	(0.219)	2.216***	(0.211)	2.392***	(0.248)	3.204***	(0.270)	2.239***	(0.105)	2.266***	(0.119)	2.321***	(0.074)
cut 4	4.276***	(0.236)	3.170***	(0.210)	3.585***	(0.196)	2.850***	(0.221)	2.961***	(0.214)	3.070***	(0.250)	3.940***	(0.276)	2.664***	(0.105)	2.568***	(0.120)	2.961***	(0.074)
cut 5	6.007***	(0.257)	4.261***	(0.220)	5.419***	(0.213)	4.788***	(0.241)	4.359***	(0.226)	4.845***	(0.272)	5.423***	(0.298)	4.312***	(0.114)	4.237***	(0.130)	4.518***	(0.079)
cut 6	7.516***	(0.275)	5.518***	(0.232)	6.781***	(0.228)	6.347***	(0.257)	6.006***	(0.245)	6.119***	(0.288)	6.648***	(0.314)	5.907***	(0.124)	5.628***	(0.140)	5.869***	(0.084)
N	1383		1388		1771		1394		1305		938		864		4845		3396		11663	

Note: Ordinal probit model coefficients with standard errors in parentheses. ***$p < 0.001$, **$p < 0.01$, *.$p < 0.05$

Table 2
Other party candidate evaluations.

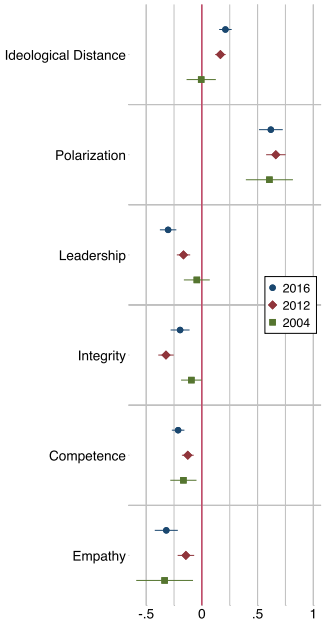
	(1)		(2)		(3)		(4)		(5)	
	1984		1988		1992		1996		2000	
	b	se	b	se	b	se	b	se	b	se
Age	0.292	(0.153)	0.196	(0.148)	0.111	(0.128)	0.023	(0.143)	0.066	(0.154)
Education	−0.107	(0.113)	−0.134	(0.104)	−0.269**	(0.096)	−0.191	(0.110)	−0.227*	(0.112)
Female	0.074	(0.060)	−0.011	(0.060)	0.081	(0.052)	0.014	(0.060)	0.026	(0.060)
White	−0.055	(0.081)	−0.279***	(0.074)	−0.260***	(0.066)	−0.474***	(0.074)	−0.148*	(0.071)
Income	0.112	(0.126)	0.222	(0.129)	−0.008	(0.113)	−0.102	(0.131)	0.036	(0.125)
Married	0.069	(0.065)	−0.143*	(0.065)	−0.017	(0.057)	0.037	(0.066)	−0.076	(0.064)
Religious Guidance	0.097	(0.081)	−0.030	(0.078)	0.034	(0.071)	−0.059	(0.079)	0.137	(0.077)
Ideological Distance	−0.977***	(0.137)	−0.831***	(0.126)	−0.460***	(0.111)	−0.836***	(0.130)	−0.467***	(0.102)
Polarization	−3.402***	(0.246)	−2.843***	(0.222)	−3.273***	(0.216)	−2.601***	(0.234)	−2.817***	(0.236)
Leadership	1.672***	(0.131)	1.410***	(0.144)	1.362***	(0.134)	1.313***	(0.153)	0.751***	(0.117)
Integrity	0.407***	(0.120)	0.629***	(0.123)	0.570***	(0.089)	0.475***	(0.095)	0.905***	(0.113)
Competence	0.153	(0.132)	0.216	(0.148)	−0.024	(0.130)	0.386**	(0.130)	0.438***	(0.128)
Empathy	0.889***	(0.131)	0.790***	(0.141)	1.260***	(0.122)	1.170***	(0.131)	0.832***	(0.115)
Online Sample										
1988										
1992										
1996										
2000										
2004										
2008										
2012										
2016										
Inflation Equation										
Age										
Education										
Female										
White										
Income										
Married										
Religious Guidance										
Ideological Distance										
Polarization										
Leadership										
Integrity										
Competence										
Empathy										
Online Sample										
Constant										
cut 1	−2.563***	(0.237)	−2.534***	(0.214)	−2.839***	(0.204)	−2.872***	(0.223)	−2.791***	(0.228)
cut 2	−1.925***	(0.233)	−1.997***	(0.210)	−2.137***	(0.200)	−2.112***	(0.219)	−1.912***	(0.222)
cut 3	−0.849***	(0.229)	−1.017***	(0.205)	−0.979***	(0.196)	−0.917***	(0.213)	−0.791***	(0.219)
cut 4	−0.225	(0.228)	−0.152	(0.204)	−0.279	(0.195)	−0.241	(0.213)	0.039	(0.219)
cut 5	1.395***	(0.237)	0.892***	(0.209)	1.175***	(0.203)	1.220***	(0.222)	1.151***	(0.221)
cut 6	2.443***	(0.257)	1.897***	(0.225)	2.137***	(0.225)	2.208***	(0.246)	2.256***	(0.244)
N	1383		1386		1771		1396		1304	

(*continued on next page*)

Table 2 (*continued*)

	(6)		(7)		(8)		(9)		(10)	
	2004		2008		2012		2016		All	
	b	se	b	se	b	se	b	se	b	se
Age	0.259	(0.215)	0.040	(0.192)	0.211*	(0.101)	0.329*	(0.139)	0.115*	(0.051)
Education	−0.497**	(0.171)	0.050	(0.144)	−0.278***	(0.067)	−0.429***	(0.105)	−0.207***	(0.037)
Female	0.004	(0.085)	−0.053	(0.075)	0.015	(0.039)	−0.077	(0.059)	−0.015	(0.020)
White	−0.222*	(0.098)	−0.228**	(0.080)	−0.352***	(0.044)	−0.186**	(0.067)	−0.244***	(0.024)
Income	−0.332*	(0.165)	−0.239	(0.160)	−0.267***	(0.080)	−0.115	(0.114)	−0.026	(0.043)
Married	0.031	(0.093)	0.042	(0.081)	−0.031	(0.043)	−0.036	(0.063)	−0.014	(0.022)
Religious Guidance	0.271*	(0.117)	0.035	(0.098)	0.181***	(0.050)	0.051	(0.072)	0.072**	(0.027)
Ideological Distance	−0.808***	(0.183)	−0.514***	(0.152)	−0.872***	(0.086)	−0.906***	(0.124)	−0.671***	(0.041)
Polarization	−2.931***	(0.397)	−2.287***	(0.272)	−2.658***	(0.165)	−2.210***	(0.225)	−2.845***	(0.080)
Leadership	0.992***	(0.160)	0.676***	(0.144)	0.795***	(0.098)	0.794***	(0.134)	1.099***	(0.045)
Integrity	0.692***	(0.156)	0.691***	(0.146)	0.955***	(0.113)	1.314***	(0.132)	0.661***	(0.039)
Competence	0.382*	(0.192)	0.257	(0.167)	0.074	(0.106)	0.387***	(0.115)	0.314***	(0.045)
Empathy	0.830***	(0.165)	1.004***	(0.149)	1.023***	(0.098)	0.841***	(0.143)	1.044***	(0.045)
Online Sample					−0.089*	(0.043)	−0.112	(0.063)		
1988									0.104**	(0.040)
1992									−0.031	(0.039)
1996									0.153***	(0.041)
2000									0.246***	(0.042)
2004									−0.004	(0.047)
2008									0.079	(0.048)
2012									0.097*	(0.042)
2016									−0.460***	(0.052)
Inflation Equation										
Age	0.141	(0.394)			0.026	(0.122)	0.070	(0.119)		
Education	−0.268	(0.341)			0.179*	(0.081)	−0.201*	(0.093)		
Female	−0.219	(0.178)			−0.052	(0.048)	−0.034	(0.050)		
White	−0.170	(0.186)			−0.181**	(0.056)	0.034	(0.059)		
Income	0.553	(0.357)			0.303**	(0.097)	0.015	(0.099)		
Married	0.196	(0.179)			−0.040	(0.052)	−0.017	(0.054)		
Religious Guidance	0.120	(0.229)			−0.024	(0.061)	0.055	(0.061)		
Ideological Distance	0.040	(0.442)			−0.692***	(0.101)	−0.711***	(0.102)		
Polarization	−3.964***	(0.646)			−2.758***	(0.196)	−2.083***	(0.202)		
Leadership	0.296	(0.369)			0.686***	(0.128)	1.020***	(0.131)		
Integrity	0.617*	(0.285)			1.338***	(0.150)	0.657***	(0.147)		
Competence	1.085**	(0.357)			0.525***	(0.113)	0.717***	(0.100)		
Empathy	2.187*	(1.051)			0.596***	(0.157)	1.075***	(0.179)		
Online Sample					−0.000	(0.053)	0.454***	(0.057)		
Constant	2.664***	(0.594)			2.186***	(0.174)	1.094***	(0.177)		
cut 1	−3.621***	(0.508)	−2.304***	(0.260)	−4.601***	(0.417)	−5.422	(3.396)	−2.487***	(0.080)
cut 2	−2.330***	(0.336)	−1.685***	(0.256)	−2.433***	(0.142)	−1.510***	(0.193)	−1.814***	(0.079)
cut 3	−1.037**	(0.321)	−0.581*	(0.250)	−1.080***	(0.138)	−0.500**	(0.190)	−0.747***	(0.078)
cut 4	−0.409	(0.319)	0.164	(0.250)	−0.426**	(0.138)	−0.071	(0.190)	−0.029	(0.078)
cut 5	0.935**	(0.326)	1.425***	(0.261)	0.920***	(0.144)	1.126***	(0.199)	1.284***	(0.080)
cut 6	1.776***	(0.349)	2.236***	(0.286)	1.795***	(0.161)	1.886***	(0.215)	2.238***	(0.088)
N	936		862		4845		3396		11658	

Note: Ordinal probit and zero inflated ordinal probit model coefficients with standard errors in parentheses. ***$p < 0.001$, **$p < 0.01$, *.$p < 0.05$



Note: The point estimates are averages of the marginal effects at every observed value of the covariates. The confidence intervals are based on standard errors calculated via the delta method.

Fig. 3. Average marginal effects of key covariates on excess zero by year.
Note: The point estimates are averages of the marginal effects at every observed value of the covariates. The confidence intervals are based on standard errors calculated via the delta method.

ratings to the other party's candidate. Marriage was largely insignificant, though in 1988, with HW Bush versus Dukakis, it was negative and significant. Overall, the more religious were more likely to give higher evaluations to the other party's candidate, with single year effects in 2004 and 2012. However, education has become significantly more important in rating the other candidate, with the more educated giving lower scores especially in 2004 and 2016. Similarly, income is consistently negative after 2000, though significant only in 2004 and 2012, where the wealthier gave lower ratings. Thus, the identity predictors (race and religiosity, but not gender) matter in other party's candidate evaluations, along with education and age.

Again, here, we find strong effects for partisan polarization. The party thermometer difference is always significant, but, again, there is no trend across the years; that is, no sign of increasing polarization effects, which is corroborated by the inconsistency in the year dummies.[9] In stark contrast to the own party candidate, the relationship to ideology was significant in every election. Ideological distance from the other party's candidate decreases evaluations of that candidate on the feeling thermometer, and, as Fig. 2 shows, that distance has been steadily increasing over the years.

4.3. Zero inflation

We use a zero inflated ordinal probit model in Table 2 to examine separately the factors that lead more people to give the most extremely negative thermometer score to the other party's candidate in 2004, 2012 and 2016. To revisit the arguments above, the recent and unprecedented increase in the number of these scores are coterminous with increases in partisan polarization and ideological sorting, leading to expectations that candidate specific qualities have become less relevant to their overall evaluations, and, in particular, the extremity of negative feelings toward the candidates. Thus, the focused examination here on zero ratings is the most direct and difficult test for our hypotheses.

Before turning to the covariates of primary interest, we note that demographic factors had surprisingly minimal effects on zero inflation. Notably, education was significant in both 2012 and 2016, but in 2012 having more of it led to an increase in rating scores, while in 2016 it led to an increase in zero ratings for the other party candidate. The only other significant effects were for race and income in 2012. Then, whites were more likely to give zero ratings and the wealthy less likely to do so. In all, and particularly in 2004 and 2016, the demographic variables, including the identity predictors, were relatively ineffectual on predicting an extremely negative rating of the other party's candidate. The backgrounds and considerations that lead to zero ratings are quite different from those that lead to the higher categories of ratings.

Turning to our major concern on the relationships of polarization, ideology, and traits with candidate evaluations, we again note that the excess zero ratings of the other party candidate diverge both from the own party candidate (in Table 1) and the higher ratings of other party candidate (in the top-half of Table 2). However, apart from 2004, where leadership and ideology are ineffectual and polarization unusually strong, the differences do not appear to be quite as stark. Greater partisan polarization consistently leads to more zero ratings, while perceiving positive traits for the other party candidate leads to higher ratings, though the effect sizes vary by year.

To more precisely compare the magnitude of these effects, Fig. 3 illustrates estimates of the average marginal effect of the ideology, trait, and polarization variables on giving excess zero ratings. Here, we compute the marginal effects for each variable for each observation using the observed covariates and then average across all the observations. Standard errors are calculated via the delta method.[10]

Overall, party polarization has the strongest relationship with zero ratings. The greater difference in one's evaluation of the two parties, the more zeros are given to the other party's candidate. In each year, polarization increases the predicted probability of a zero rating to more than .5. Notably, while the effects are large, they are also consistent over time. That is, there is no evidence that they are playing a greater role in the increasing number of zero ratings over time. Even in 2016, the magnitude of the effect is about the same as in the two previous years with high zero ratings.

By contrast, the effects for ideological distance show an increasing trend over time. While substantively ineffectual and statistically insignificant in 2004, the effect grew in 2012 and again in 2016, though only marginally so. In 2012, 2016, the greater one's ideological distance from the other party's candidate, the more likely one is to give a zero rating to that candidate. Still, the effects are fairly small relative to polarization.

We find a similar, though oppositely signed, trend in one of the candidate traits, leadership. Seeing the other party's candidate as a strong leader in 2004 had little to no effect on zero ratings. However, eight years later, it pushed respondents away from giving the other party's candidate the most negative score. It made a similarly sized jump in only four years, making it even more powerful than ideology in 2016. In contrast to the other traits, leadership appears to be playing an increasing role in preventing zeros.

Competence and empathy are consistently significant, though the magnitude is generally larger for the latter. In addition, for both cases we find a similar pattern, in so far as 2016 looked more like 2004 than 2012. Specifically, perceptions of the other party's candidate as competent and empathetic were more powerful in moving that candidate's evaluation away from zero in 2004 and 2016 than in 2012. As such, the

[9] Despite the lack of trend, the year dummy coefficient is largest in 2016, which may suggest that ideological distance and partisan polarization did not do as good a job of capturing the effect of polarization on candidate evaluations in that year, though this could also be due to campaign effects.

[10] Discrete changes in predicted probabilities for these are illustrated in the Appendix.

over time pattern for integrity was the most unusual of the bunch. The other party candidates were most protected from the zero ratings by positive perceptions of integrity in 2012.

4.4. The 2016 election

Looking across these results, 2016 stands out in several respects. The effect of ideological difference from the candidate was on the high end both for own and other party candidates. And while party polarization was substantively and statistically significant, by historical standards the effect size was relatively average to low. Integrity, for once, predominated as a consideration for trait effects on the higher rating categories of both own and other party candidates—and that was the only trait on which the public rated Clinton below Trump. By contrast, Clinton had a strong lead over Trump overall on perceived competence, but, unfortunately for Clinton, that mattered least for own and other party candidate evaluations.

As to zero ratings of the other party's candidate, empathy and leadership mattered the most of the traits in 2016, and the two candidates obtained fairly similar ratings on those (Clinton's 0.23 from Republicans on leadership versus Trump's 0.22 from Democrats; Clinton's 0.17 from Republicans on empathy versus Trump's 0.10 from Democrats). Relative to previous years with many zeros, the zero inflation story is that of the highest marginal effects for ideological distance as well as for at least half of the candidate traits. Leadership and competence had their greatest effects, with empathy approximately the same as in 2004. Only integrity had a lesser effect in 2016 than in previous years, and even that effect was fairly large. The effect of polarization, however, is typical for high inflation years. Thus the relationship of traits to the lowest candidate evaluation appears unusually strong in 2016. In sum, candidates' images—i.e., their perceived characteristics—appear to be as responsible as the rising tide of polarization, if not more so, for the public's extreme hostility in 2016.

5. Conclusion

Both the descriptive statistics and multivariate analyses of candidate evaluations suggest support for the recognized trend of partisan polarization, and negative partisanship in particular (e.g., Iyengar et al., 2012; Abramowitz and Webster, 2016). The thermometers for party candidates have dropped substantially over the last couple of decades, and the bulk of the movement has been in the evaluations of the other party candidates, with own party candidates remaining fairly consistent. Moreover, three of the last four elections set records in the number of zero ratings for candidates, the absolute lowest score available on the thermometer, and, again, these were predominantly for other party candidates. What, if anything beyond partisan polarization, explains these recent trends?

Using ANES data on the last eight presidential elections we show that partisan polarization was a consistent and powerful factor in the own and other party candidate ratings, though the effect appears to have diminished rather than grown in recent years. Still, its effect on predicting the most negative ratings in 2016, 2012 and 2004 were second to none. Furthermore, its effect is beyond that of ideological distance, suggesting that partisan polarization is more tribalism than ideological.

However, the strong polarization effects did not crowd out the use of candidate traits. The public relied on these in their evaluations of both own and other party candidates, though to somewhat different degrees depending on the candidates, confirming Funk (1999) on variability in the importance of the different traits across election years. In explaining the most extreme negative ratings, leadership and competence appear to be on the rise, with empathy and integrity also quite powerful, but not as much as in past years. Given the rather static effects of polarization then, the large number of zero ratings in 2016 appear primarily due to big drops in key candidate traits. Clinton had an unusually long and controversial history, while Trump exhibited a unique ability to stir up controversy that rebounded on himself. The dismal showing of both Clinton and Trump on the standard candidate trait items is directly responsible for their low thermometer scores and the intense dislike that many Americans felt towards them. Trait perceptions matter, and depictions by campaigns of their opponents as "crooked" or "lying" lead to intensely negative thermometer scores.

The results here should not be interpreted without context. While looking at own and other party candidates allows us strong control for partisanship, it comes with a price. Specifically, the analysis does not permit us to say anything about the small number of pure independents (recall that leaners are included in their respective parties). Indeed, candidate traits may be particularly important to independents (see, e.g., Holian and Prysby, 2014), and subsequent work may shed light on whether the relationships uncovered here exist for them as well.

Additionally, our analyses are purely correlational, and should not be interpreted as causal without caution. As we mentioned above, there is unexplored potential for reverse causality or reinforcement in the relationship between party and candidate evaluations. Neither the data nor the statistics applied here allow us to test the causal directions between the various measures in our models, suggesting that future research is necessary to clarify them.

Finally, we have examined feelings toward the presidential candidates rather than actual vote. People nearly always vote for the candidate they like more, so the difference between the candidate thermometers is, of course, highly correlated with vote choice. Still, our evidence that the public relies heavily on candidate specific factors when evaluating candidates may not translate perfectly to votes, where the calculus may be slightly different and even more closely tied to partisanship (Miller and Shanks, 1996; Bartels, 2002). Regardless, the fact that candidate traits have varied in their relative influence but generally remained substantively significant factors of both own and other party candidate evaluations across nine elections—and particularly in 2016—suggests that claims of the end of candidate considerations in the age of polarization are premature.

Appendix A. Online Appendix

Supplementary Appendix can be found online at https://doi.org/10.1016/j.electstud.2019.03.008.

References

Abramowitz, Alan I., Saunders, Kyle L., 2008. Is polarization a myth? J. Politics 70 (2), 542–555.
Abramowitz, Alan I., Webster, Steven, 2016. The rise of negative partisanship and the nationalization of US elections in the 21st century. Elect. Stud. 41, 12–22.
Abramson, Paul R., Aldrich, John H., Rickershauser, Jill, Rohde, David W., 2007. Fear in the voting booth: the 2004 presidential election. Polit. Behav. 29 (2), 197–220.
Achen, Christopher H., Bartels, Larry M., 2017. second ed. Democracy for Realists: Why Elections Do Not Produce Responsive Government Princeton University Press.
Alwin, Duane F., 1992. Information transmission in the survey interview: number of response categories and the reliability of attitude measurement. Socio. Methodol. 83–118.
Alwin, Duane F., 1997. Feeling thermometers versus 7-point scales: which are better? Socio. Methods Res. 25 (3), 318–340.
Bafumi, Joseph, Shapiro, Robert Y., 2009. A new partisan voter. J. Politics 71 (1), 1–24.
Barber, Michael, McCarty, Nolan, 2015. Causes and consequences of polarization. Political Negotiation: Hand 37, 39–43.
Bartels, Larry M., 2002. Beyond the running tally: partisan bias in political perceptions. Polit. Behav. 24 (2), 117–150.
Bougher, Lori D., 2017. The correlates of discord: identity, issue alignment, and political hostility in polarized America. Polit. Behav. 39 (3), 731–762.
Campbell, Angus, Converse, Philip E., Miller, Warren E., Stokes, Donald E., 1960. The American Voter. University of Chicago Press, New York.
Caruana, Nicholas J., McGregor, R Michael, Stephenson, Laura B., 2015. The power of the dark side: negative partisanship and political behaviour in Canada. Canadian Journal of Political Science/Revue canadienne de science politique 48 (4), 771–789.
Converse, Philip E., 1964. The nature of belief systems in mass publics. Crit. Rev. 18 (1–3), 1–74.
Ellis, Christopher, Stimson, James A., 2012. Ideology in America. Cambridge University Press.

Enders, Adam M., Armaly, Miles T., 2018. The differential effects of actual and perceived polarization. Polit. Behav. 1–25.
Fiorina, Morris, P., Abrams, Samuel J., Pope, Jeremy, 2005. Culture War?: the Myth of a Polarized America. Pearson Longman, New York.
Funk, Carolyn L., 1996. The impact of scandal on candidate evaluations: an experimental test of the role of candidate traits. Polit. Behav. 18 (1), 1–24.
Funk, Carolyn L., 1999. Bringing the candidate into models of candidate evaluation. J. Politics 61 (3), 700–720.
Green, Donald P., Bradley, Palmquist, Schickler, Eric, 2004. Partisan Hearts and Minds: Political Parties and the Social Identities of Voters. Yale University Press.
Greenberg, Stanley B., 2004. The Two Americas: Our Current Political Deadlock and How to Break it. Macmillan.
Harris, Mark N., Zhao, Xueyan, 2007. A zero-inflated ordered probit model, with an application to modelling tobacco consumption. J. Econom. 141 (2), 1073–1099.
Hetherington, Marc J., Long, Meri T., Rudolph, Thomas J., 2016. Revisiting the myth: new evidence of a polarized electorate. Publ. Opin. Q. 80 (S1), 321–350.
Holian, David B., Prysby, Charles, 2014. Candidate character traits in the 2012 presidential election. Pres. Stud. Q. 44 (3), 484–505.
Huddy, Leonie, Mason, Lilliana, Aarøe, Lene, 2015. Expressive partisanship: campaign involvement, political emotion, and partisan identity. Am. Pol. Sci. Rev. 109 (1), 1–17.
Iyengar, Shanto, Sood, Gaurav, Lelkes, Yphtach, 2012. Affect, not IdeologyA social identity perspective on polarization. Publ. Opin. Q. 76 (3), 405–431.
Jacobson, Gary C., 2000. Party polarization in national politics: the electoral connection. In: Cohen, Jeffrey, Fleisher, Richard, Kantor, Paul (Eds.), Polarized Politics: Congress and the President in a Partisan Era. CQ Press, pp. 9–30.
Jacobson, Gary C., 2007. A Uniter, Not a Divider. George W. Bush and the American People.
Jacobson, Gary C., 2010. George W. Bush, the Iraq war, and the election of barack Obama. Pres. Stud. Q. 40 (2), 207–224.
Kenney, Patrick J., Rice, Tom W., 1988. The contextual determinants of presidential greatness. Pres. Stud. Q. 161–169.
Kinder, Donald R., 1986. Presidential character revisited. Political cognition 19.
Klar, Samara, Krupnikov, Yanna, 2016. Independent Politics. Cambridge University Press.
Klar, Samara, Krupnikov, Yanna, Ryan, John Barry, 2018. Affective polarization or partisan disdain? Untangling a dislike for the opposing party from a dislike of partisanship. Publ. Opin. Q. 82 (2), 379–390.
Levendusky, Matthew, 2009. The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans. University of Chicago Press.
Lupton, Robert, Jacoby, William G., 2016. The Reliability of the ANES Feeling Thermometers: an Optimistic Assessment. Working Paper. http://polisci.msu.edu/jacoby/research/therms/2016%20SPSA/Lupton-Jacoby,%202016%20SPSA%20Paper.pdf.
Markus, Gregory B., 1982. Political attitudes during an election year: a report on the 1980 NES panel study. Am. Pol. Sci. Rev. 76 (3), 538–560.
Markus, Gregory B., Converse, Philip E., 1979. A dynamic simultaneous equation model of electoral choice. Am. Pol. Sci. Rev. 73 (4), 1055–1070.
Mason, Lilliana, 2015. I disrespectfully agree: the differential effects of partisan sorting on social and issue polarization. Am. J. Pol. Sci. 59 (1), 128–145.
Mason, Lilliana, 2018. Ideologues without issues: the polarizing consequences of ideological identities. Publ. Opin. Q. 82 (S1), 280–301.
Medeiros, Mike, Noël, Alain, 2014. The forgotten side of partisanship: negative party identification in four Anglo-American democracies. Comp. Pol Stud. 47 (7), 1022–1046.
Miller, Arthur H., Wattenberg, Martin P., Malanchuk, Oksana, 1986. Schematic assessments of presidential candidates. Am. Pol. Sci. Rev. 80 (2), 521–540.
Miller, Arthur H., Miller, Warren E., 1976. Ideology in the 1972 election: myth or realitya rejoinder. Am. Pol. Sci. Rev. 70 (3), 832–849.
Miller, Warren Edward, Shanks, J Merrill, 1996. The New American Voter. Harvard University Press Cambridge, MA.
Page, Benjamin I., Jones, Calvin C., 1979. Reciprocal effects of policy preferences, party loyalties and the vote. Am. Pol. Sci. Rev. 73 (4), 1071–1089.
Pew, 2016. Partisanship and Political Animosity in 2016. Pew Research Center. http://www.people-press.org/2016/06/22/partisanship-and-political-animosity-in-2016/, Accessed date: 25 November 2017.
Rabinowitz, George, 1976. A procedure for ordering object pairs consistent with the multidimensional unfolding model. Psychometrika 41 (3), 349–373.
Rogowski, Jon C., Sutherland, Joseph L., 2016. How ideology fuels affective polarization. Polit. Behav. 38 (2), 485–508.
Rose, Richard, Mishler, William, 1998. Negative and positive party identification in post-communist countries. Elect. Stud. 17 (2), 217–234.
Smidt, Corwin D., 2017. Polarization and the decline of the American floating voter. Am. J. Pol. Sci. 61 (2), 365–381.
Stokes, Donald E., 1966. Some dynamic elements of contests for the presidency. Am. Pol. Sci. Rev. 60 (1), 19–28.
Wattenberg, Martin P., 2016. The declining relevance of candidate personal attributes in presidential elections. Pres. Stud. Q. 46 (1), 125–139.
Weisberg, Herbert F., Miller, Arthur H., 1979. Evaluation of the Feeling Thermometer: A Report to the ANES Board Based on Data from the 1979 Pilot Study Survey.
Weisberg, Herbert F., Christenson, Dino P., 2007. Changing horses in wartime? The 2004 presidential election. Polit. Behav. 29 (2), 279–304.
Weisberg, Herbert F., Mockabee, Stephen T., 1999. Reelection 1996: How Americans Voted. Chatham House chapter Attitudinal correlates of the 1996 presidential vote The people reelect a president, Chappaqua, NY, pp. 45–69.
Weisberg, Herbert F., Hill, Timothy G., 2004. The Succession Presidential Election of 2000: the Battle of the Legacies. Models of voting in Presidential Elections, pp. 27–48.
Whiteley, Paul F., 1988. The causal relationships between issues, candidate evaluations, party identification, and vote choice–the view from "rolling thunder". J. Politics 50 (4), 961–984. http://www.jstor.org/stable/2131387.