QAC Data Science Practicum:
Examining the Relationship Between Candidate Stance on Trump and Ultimate Electoral Performance in Congressional Republicans during the 2016 Election

Abstract

This study uses logistical regression, genetic matching and ordinary least square regression models to examine the relationship between Trump stance and electoral performance during the 2016 election cycle in regards to Republican House of Representative candidates. After analysis, it was found that partisan makeup of the district was the only predictor of both Trump stance and electoral performance. Stance on Trump had no significant role in determining electoral performance. This paper fits into prior literature as it affirms findings generated after the 2016 election, while utilizing a different methodology.

Casey Jordan '17
Wesleyan University

The 2016 election was vastly different than preceding elections throughout history. In an era of high partisanship and polarization, the Republican party nominated a presidential candidate, Donald Trump, so controversial that members of their own party condemned the nominee. To endorse or not to endorse became the question. Some candidates for United States Congress endorsed Trump early and never looked back, some stated their opposition loudly and repeatedly throughout the election, and some could not seem to make up their mind. Take Jason Chafetz (R-UT) for example, an original endorser of Trump, on October 8th he released this statement, "I'm out. I can no longer in good conscience endorse this person for president. It is some of the most abhorrent and offensive comments that you can possibly imagine."[1] However, it took him just two weeks to change his mind once again, when he tweeted "I will not defend or endorse Donald Trump, but I am voting for him. HRC is that bad. HRC is bad for the USA."[2] Using genetic matching fit by logistic regression, and ultimately ordinary least squares models, this paper will examine the association between these public declarations of support, or condemnation, and the electoral success of Republican candidates for the United States Congress in the 2016 election cycle.

A few prior studies have examined this relationship in the context of the 2016 election, and found mixed results. In determining the predictors of taking a stance on Donald Trump, one study found the only significant predictor to be the partisan makeup of the congressional district (Binder 2016). Prior literature into the effect of an endorsement of Trump on electoral performance varied. One study found that "abandoning Trump *mid-campaign* — a strategy Speaker Paul Ryan seemed to adopt, possibly in an effort to preserve a Republican majority in

[1] http://www.politico.com/story/2016/10/rep-chaffetz-withdraws-his-endorsement-of-trump-229335
[2] https://twitter.com/jasoninthehouse/status/791445788656226304?lang=en

the House — probably *cost* Republicans at least one Senate seat" (Wasserman 2017). This study looked at only five candidates for office when comparing election results, and did not analyze the entire field, tempering its findings of a Trump condemnation's moderately negative effect on ultimate vote share. A study utilizing more robust statistical analysis determined that "strategic choices had little general impact in an election almost completely dominated by partisanship" (Liu Jacobson 2017, 1), and "in no equation does support or non-support of Trump have anything close to a significant effect on the vote either way. District partisanship dominates" (Liu Jacobson 2017, 24). The few studies that have been done, though disagreeing slightly on outcomes, highlight the key role district partisanship played in the 2016 election, the most determining factor of both which stance was taken on Trump, and ultimate electoral performance.

This paper will base its core analytical methodology off a study from 2010 examining the effect of certain roll call votes on electoral performance (Nyhan et al 2010). A basic logistical regression will be used to determine the significant predictors of taking a stance on Trump, measured in a binary variable with one corresponding to ultimately explicitly endorsing Trump, and zero meaning no endorsement, or and issuance of a statement of opposition. The results of this logistical regression will be used in a genetic matching algorithm. Genetic matching is a "method of multivariate matching that uses an evolutionary search algorithm to determine the weight each covariate is given" (Diamond Sekhon 2013, 932). The goal is to create balance in the treatment and control groups by finding both treated and controlled variables with similar covariates, simulating two groups in an experimental design. Covariate balance is achieved after a new dataset has been created from the search algorithm, and "the treatment and control groups have the same joint distribution of observed covariates" (Diamond Sekhon 2013, 932). The independent variables to match observations on must be the predictors of the binary treatment variable, in this case stance on Trump. These predictors will be used to match Trump supporters to Trump opponents, against whom the effect of the treatment variable can be analyzed. Once the matched dataset has been created, an OLS model will be run to determine the association of Trump stance and vote share, controlling for partisanship of both the district and candidate.

The data collected for the logistical regression analysis of predictors of Trump stance included every Republican candidate for United States Congress in the 2016 election cycle that was not running unopposed and who had made a publically discernable stance on Trump at one point. This included both incumbents and challengers where applicable. For incumbents, Trump stance was collected from a dataset put together by Daniel Nichanian, a post doc at the University of Chicago[3]. It was aggregated after the release of the Access Hollywood tapes. Challengers stances were found online through local publications when available. District demographic information was collected from the American Community Survey[4], and variables included percentage of congressional district that is Hispanic, Black, Asian or Pacific Islander, college educated, employed in construction jobs, manufacturing jobs and the median household income. Partisanship of the district was gathered from DailyKos[5], measured by the share of the two party vote Trump received on election day. Partisanship of the candidate was measured by DW-NOMINATE scores (Carroll et al 2008), "the acronym DW-NOMINATE stands for dynamic, weighted, nominal three-step estimation. It is a FORTRAN program that estimates a

[3] The list is available at: https://docs.google.com/spreadsheets/d/15MudzswjU45efqvus21goEqtknbsfy9w2CLaWvOa_z0/edit#gid=0
[4] https://www.census.gov/programs-surveys/acs/
[5] http://www.dailykos.com/story/2017/1/30/1627319/-Daily-Kos-Elections-presents-the-2016-presidential-election-results-by-congressional-district

probabilistic model of binary choice of legislators in a Parliamentary setting over time" (Carroll et al, 1). Applied to the United States Congress it outputs a positive or negative value depending on the partisanship of a member of congress, depending on certain roll call votes the member makes. Negative values point to a more liberal candidate, and positive values point to a more conservative candidate. The candidate's gender and whether they were white or non-white was recorded as well.

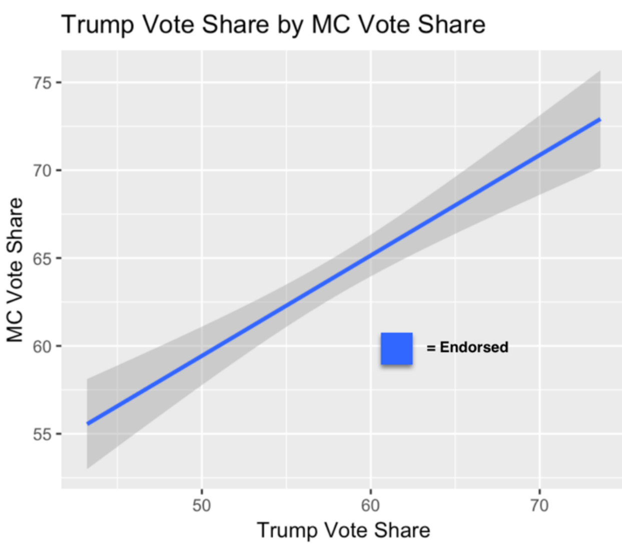
The results for the preliminary logistic regression are shown in Appendix A. The most significant, and only predictor at the $\alpha = .05$ level, of Trump support is Trump vote. A higher share of the two party vote for Trump, corresponding to a more conservative congressional district, increases the likelihood of the candidate endorsing Trump. This is in line with previous literature (Binder 2016). When controlling for potential confounders, the only significant predictor of endorsing Trump is how conservative the congressional district is.

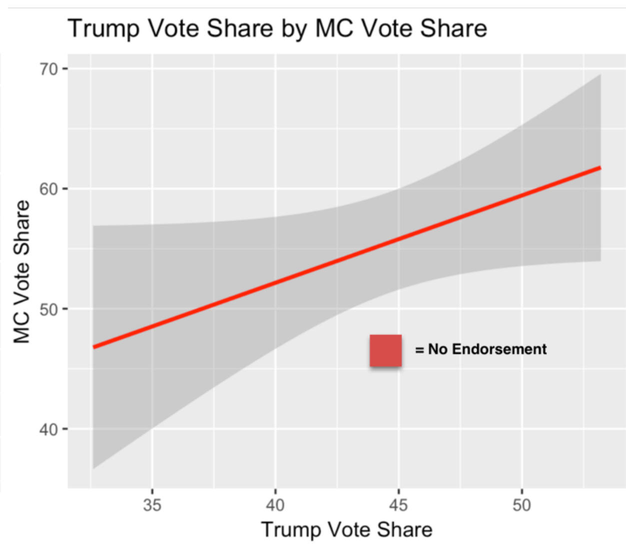
Working off the findings of the logistic regression, a genetic matching algorithm was run matching the most similar Trump Supporters and Opponents based on Trump vote of their district. The matched dataset included 204 matched observations for inference. An OLS regression was run on the matched dataset, the results are shown below.

	Dependent variable:	
	MC Vote Share	
	All	Matched
Trump Stance	-0.024+	-0.016
	(0.012)	(0.012)
TrumpVote	0.683***	0.679***
	(0.064)	(0.064)
DWNom	0.007	-0.008
	(0.034)	(0.034)
Constant	0.258***	0.261***
	(0.034)	(0.036)
Observations	213	204
R2	0.375	0.372
Adjusted R2	0.366	0.362
Residual Std. Error	0.071 (df = 209)	0.068 (df = 200)
F Statistic	41.786*** (df = 3; 209)	39.473*** (df = 3; 200)
Note:	+p<0.1; **p<0.05; ***p<0.01	

The dependent variable is candidate vote share, and the independent variables are Trump stance, DW-NOMINATE and Trump vote. The results are shown in comparison to an OLS regression run on the unmatched data. After controlling for the partisanship of the district, as well as the member of congress, only Trump vote share is significant when predicting member of congress vote share. An increase in Trump vote share corresponds to an increase in Republican candidate's vote share, regardless of whether the candidate endorsed Trump or not. In other words, the more conservative the district, the better the Republican candidate performed. Partisanship of the district is the only significant factor in the 2016 election, agreeing with prior studies (Liu Jacobson 2017). When compared to the OLS model on unmatched data, the findings are similar. Trump support showed an almost interesting negative relationship at the $\alpha = .1$ level, noteworthy, but not significant. This lends credence to the benefit of using a genetically matched dataset. Causation can never be 100% proven on observational data, but genetic matching allows for stronger analysis of relationships and associations, as it transforms the data into mimicking quasi-experimental data, allowing for more confident inference.

The relationship between Trump vote share and Republican candidate vote share is shown below. Regardless of endorsement, or lack thereof, there is a strong positive relationship.





two graphs is where Trump vote share is roughly equal to 50% of the two party vote. On each graph that corresponds with just under 60% of the two party vote for the member of congress, visually depicting the finding that Trump stance is not a factor that influenced vote share.

After conducting analysis using a range of statistical techniques, it can be concluded that Trump stance has no association with ultimate vote share in the 2016 election. This study agrees with prior literature when concluding that the partisan makeup of a congressional district is the most significant predictor of Trump support or opposition (Binder 2016), and that Trump stance had no association with electoral performance when controlling for partisanship (Liu Jacobson 2017). This study, while affirming prior findings, contributes to the literature by introducing a different technique than what has been used, genetic matching. Further analysis into the 2016 election might examine the influence of the Trump candidacy on the public, and their motivations for supporting Trump in an effort to determine similarities between members of congress and the general public.

Works Cited

Binder, Sarah. 2016. "These are the House Republicans Who Have Abandoned Donald Trump," Oct. 9: https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/09/these-are-the-house-republicans-who-have-abandoned-donald-trump/?utm_term=.d6d274bffc2c

Carroll, Royce and Lewis, Jeffrey B. and Lo, James and Poole, Keith T. and Rosenthal, Howard, Measuring Bias and Uncertainty in DW-NOMINATE Ideal Point Estimates via the Parametric Bootstrap (September 4, 2008).

Diamond, Alexis and Jasjeet S. Sekhon. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *The Review of Economics and Statistics* 95(3): 932-945.

Liu, Huchen and Gary C. Jacobson. 2017. "Ride with the Top of the Ticket? Republican Congressional Candidates' Positions on Donald Trump and Their Impact on the 2016 Elections." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago, IL.

Nyhan, Brendan, Eric McGhee, John Sides, Seth Masket, and Steven Greene. 2012. "One Vote Out of Step? The Effects of Salient Roll Call Votes in the 2010 Election." *American Politics Research* 40(5): 844-879.

Wasserman, David. 2017. "Republicans Could Pay a Price for Ditching Trump—Even if He's Unpopular," Fivethirtyeight.com, March 22: https://fivethirtyeight.com/features/republicans-could-pay-a-price-for-ditching-trump-even-if-hes-unpopular/

Appendix A.

	Dependent variable:
	Trump Support
TrumpVote	13.240***
	(4.315)
MedianHHIncome	-0.00004
	(0.00004)
% Hispanic	1.712
	(2.286)
% Black	-3.358
	(3.629)
% AsiaPI	2.229
	(8.732)
% CollegeEducated	9.075
	(6.314)
% Manufacturing	2.408
	(6.820)
% Construction	54.302
	(36.018)
NonWhite	-1.986+
	(1.175)
Sex	-0.940
	(0.670)
DWNom	3.692+
	(1.960)
Constant	-10.593***
	(3.829)
N	214
Log Likelihood	-58.596

Note: +p<0.1; **p<0.05; ***p<0.01