

SOCIAL DESIRABILITY BIAS AND VOTER LIE DETECTION



Lauren Lawless
Kiel Brunner

Nyesher Serrant
Joe Hernandez

SOCIAL DESIRABILITY BIAS

“Social desirability bias is the tendency of some respondents to report an answer in a way they deem to be more socially acceptable than would be their ‘true’ answer” (Lavrakas, 2008).

Intent: project a more favorable image

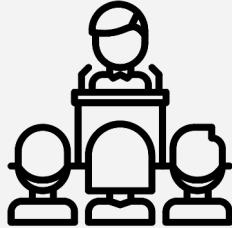
Outcome: overreporting of socially desirable behaviors and underreporting of socially undesirable behaviors

BACKGROUND INFORMATION

POLLING POST-2016

- **The situation:**
 - Voter polling remains integral to election analysis, despite the expanded use of other modeling and statistical techniques.
- **The catalyst:**
 - The 2016 election of Donald Trump and increased political divisiveness since have led to social desirability bias in self-reported data.
- **The result:**
 - Skewed voter preference predictions
 - Over-investment in unpersuadable voters

QUESTIONS



What demographic characteristics and political issue preferences influence a voter's decision to support Trump?

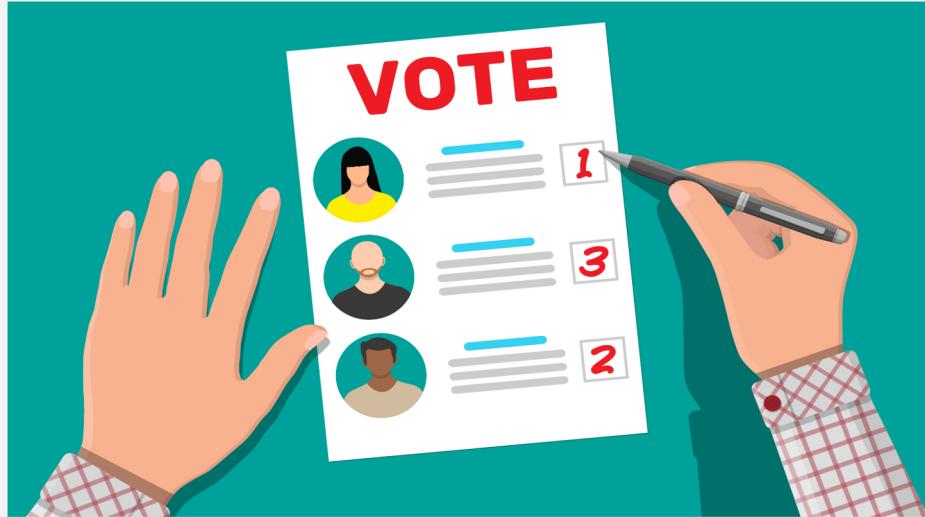
Who, then, should a Democratic candidate expend resources to persuade, and who is likely to support Trump despite not admitting to it previously?

Can a model based on these characteristics detect social desirability bias and help us determine the likelihood that a person would lie about having previously supported Trump?

SOLUTION STRATEGY

Using data from AP VoteCast 2018:

- Examine individuals' responses to Trump-support-equivalent questions:
 - 2016 candidate choice
 - 2020 anticipated candidate choice
 - Trump approval
 - Voting in 2018 to show support for Trump
 - Trump favorability
- Determine likelihood that the person supported Trump in 2016



- Compare to reported 2016 candidate choice to determine likelihood that social desirability bias influenced the response

This probability will help inform the campaign of the 2020 Democratic nominee about which voters are “gettable” and which would otherwise be an inefficient use of resources.

DATA ACQUISITION PROCESS

AP VoteCast was designed to improve polling accuracy by broadening traditional polling methods and providing unprecedented public access to the debut data set.

- Observations: 106,855 (7 train : 3 test)
- Variables: 53, including dummies



ESTIMATING SUPPORT

FOR TRUMP

Significant negative impacts:

- Issues (economy, healthcare, taxes, gun policy, environment, foreign policy)
- Immigration policy (path to citizenship)
- 2016 Russian collusion
- Age (18-24)
- Non-religious
- Urban, suburban, and small town

Significant positive impacts:

- Country track
- Trump economic impact
- Repeal ACA
- Sex (m)
- Age (40-49)
- Race/ethnicity (White, Asian, Hispanic)
- Education (“less is more”)
- Income (lowest for \$50k-\$75k)
- Party/ideology (scale: 0 Democrat/very liberal ... 4 Republican/very conservative)
- Religion (Protestant, Catholic)

Dep. Variable:	suppTrump	No. Observations:	74798			
Model:	Logit	Df Residuals:	74744			
Method:	MLE	Df Model:	53			
Date:	Tue, 17 Mar 2020	Pseudo R-squ.:	0.7316			
Time:	11:42:07	Log-Likelihood:	-13643			
converged:	TRUE	LL-Null:	-50824			
Covariance Type:	nonrobust	LLR p-value:	0			
mode	-0.139	0.081	-1.725	0.085	-0.298	0.19
electinterest	0.112	0.019	5.972	0.000	0.075	0.148
p_vote	-0.200	0.129	-1.546	0.122	-0.452	0.053
likelyvoter	-0.106	0.071	-1.492	0.136	-0.244	0.033
timevote	0.076	0.035	2.204	0.028	0.008	0.144
track	1.987	0.035	57.525	0.000	1.919	2.054
issues_economyandjobs	-0.217	0.066	-3.275	0.001	-0.346	-0.087
issues_healthcare	-0.447	0.065	-6.895	0.000	-0.573	-0.320
issues_immigration	0.058	0.070	0.837	0.402	-0.078	0.194
issues_taxes	-0.240	0.091	-2.626	0.009	-0.419	-0.061
issues_abortion	0.081	0.125	0.643	0.520	-0.165	0.326
issues_gunpolicy	-0.541	0.082	-6.610	0.000	-0.701	-0.380
issues_environment	-0.739	0.089	-8.279	0.000	-0.913	-0.564
issues_foreignpolicy	-0.674	0.088	-7.682	0.000	-0.846	-0.502
fanfinancial	0.065	0.029	2.209	0.027	0.007	0.122
trumpnatecon	0.668	0.023	28.511	0.000	0.622	0.714
trumplocalecon	0.452	0.027	16.663	0.000	0.399	0.506
supACA	0.332	0.017	19.546	0.000	0.299	0.365
imppol	-0.731	0.036	-20.068	0.000	-0.802	-0.660
russiacollude16	-0.912	0.038	-25.163	0.000	-0.983	-0.841
voted14	-0.127	0.041	-3.094	0.002	-0.207	-0.046
violencespeechD	0.831	0.036	23.181	0.000	0.760	0.901
violencespeechR	-0.786	0.037	-21.095	0.000	-0.859	-0.713
sex_m	0.129	0.035	3.684	0.000	0.060	0.198
age_18_24	-0.467	0.074	-6.329	0.000	-0.612	-0.322
age_25_29	-0.074	0.069	-1.069	0.285	-0.209	0.062
age_30_39	0.091	0.056	1.624	0.104	-0.019	0.200
age_40_49	0.122	0.058	2.078	0.038	0.007	0.236
age_50_64	-0.012	0.051	-0.234	0.815	-0.111	0.087
raceth_wht	0.250	0.081	3.083	0.002	0.091	0.409
raceth_blk	-0.118	0.096	-1.226	0.220	-0.306	0.071
raceth_hsp	0.204	0.094	2.169	0.030	0.020	0.388
raceth_asn	0.340	0.121	2.797	0.005	0.102	0.578
raceth_ntvam	-0.162	0.246	-0.658	0.511	-0.643	0.320
raceth_ipac	-0.274	0.346	-0.792	0.429	-0.951	0.404
edu_highschoolorless	0.520	0.062	8.410	0.000	0.399	0.641
edu_somewhereassocdegree	0.380	0.053	7.206	0.000	0.276	0.483
edu_collegegraduate	0.111	0.052	2.124	0.034	0.009	0.213
income_under25000	0.283	0.062	4.546	0.000	0.161	0.406
income_25000_49999	0.213	0.054	3.951	0.000	0.107	0.319
income_50000_74999	0.128	0.053	2.440	0.015	0.025	0.231
income_75000_99999	0.226	0.055	4.121	0.000	0.118	0.333
party	0.337	0.012	28.038	0.000	0.313	0.361
ideo	0.074	0.019	3.914	0.000	0.037	0.111
relig_prot_chrst	0.207	0.064	3.243	0.001	0.082	0.331
relig_catholic	0.187	0.068	2.739	0.006	0.053	0.321
relig_mormon	0.186	0.156	1.191	0.233	-0.120	0.493
relig_jewish	-0.010	0.120	-0.086	0.931	-0.245	0.225
relig_muslim	0.648	0.182	3.567	0.000	0.292	1.004
relig_none	-0.149	0.068	-2.179	0.029	-0.283	-0.015
sizeplace_urban	-0.231	0.054	-4.272	0.000	-0.337	-0.125
sizeplace_suburban	-0.346	0.047	-7.379	0.000	-0.438	-0.254
sizeplace_smalltown	-0.187	0.052	-3.578	0.000	-0.290	-0.085
intercept	-2.590	0.188	-13.786	0.000	-2.959	-0.222

MODEL COMPARISON

Model	Train Time (seconds)	Predict Time (train, test; seconds)	Accuracy (train, test)	Confusion Matrix (test)			
						Prediction	
Logistic	2.57	0.02 0.01	0.9301719 0.9317465			0	1
				Support	Trump	0	17,547
				1		1,172	12,322
Random Forest	7.51	1.76 0.64	0.9740100 0.9332751			Prediction	
						0	1
				Support	Trump	0	17,553
				1		1,129	12,365

QUANTIFYING INFLUENCE OF

SOCIAL DESIRABILITY BIAS

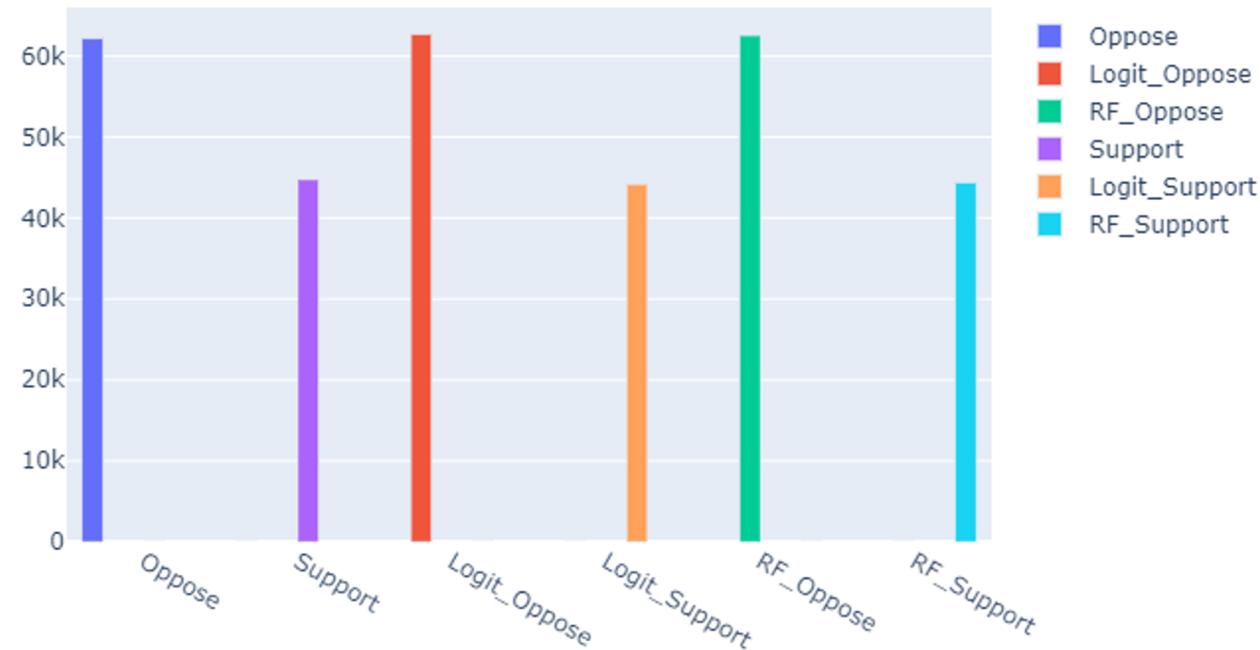
Significant positive variables indicate likely social desirability bias among Trump supporters*:

- Issues (abortion, environment, gun policy, healthcare)
- Age (25-39)
- Race/ethnicity (Black)
- Non-rural

Dep. Variable:	vote16_donaldtrump	No. Observations:	8108			
Model:	Logit	Df Residuals:	8054			
Method:	MLE	Df Model:	53			
Date:	Tue, 17 Mar 2020	Pseudo R-squ.:	0.2292			
Time:	13:42:52	Log-Likelihood:	-2173.9			
converged:	TRUE	LL-Null:	-2820.2			
Covariance Type:	nonrobust	LR p-value:	1.21E-235			
	coef	std err	z	P> z	[0.025	0.975]
mode	-0.001	0.184	-0.007	0.994	-0.361	0.359
electinterest	-0.204	0.047	-4.308	0.000	-0.297	-0.111
P_vote	0.138	0.480	0.287	0.774	-0.803	1.077
likelyvoter	0.313	0.260	1.203	0.229	-0.197	0.822
timevote	-0.141	0.086	-1.640	0.101	-0.309	0.027
track	0.021	0.101	0.212	0.832	-0.176	0.218
issues_economyandjobs	0.381	0.169	2.262	0.024	0.051	0.712
issues_healthcare	0.432	0.171	2.527	0.012	0.097	0.767
issues_immigration	0.017	0.173	0.099	0.921	-0.322	0.356
issues_taxes	0.610	0.230	2.650	0.008	0.159	1.061
issues_abortion	0.853	0.260	3.286	0.001	0.344	1.362
issues_gunpolicy	0.553	0.241	2.296	0.022	0.081	1.026
issues_environment	0.986	0.269	3.671	0.000	0.459	1.512
issues_foreignpolicy	0.132	0.257	0.515	0.607	-0.371	0.635
famfinacial	0.052	0.077	0.676	0.499	-0.099	0.203
trumpratecon	0.055	0.072	0.770	0.441	-0.085	0.196
trumplocalecon	-0.203	0.076	-2.671	0.008	-0.351	-0.054
supACA	-0.133	0.047	-2.839	0.005	-0.225	-0.041
impol	0.254	0.086	2.961	0.003	0.086	0.422
russiacollude16	0.132	0.118	1.121	0.262	-0.099	0.363
voted14	0.131	0.113	1.155	0.248	-0.091	0.353
violencespeechD	-0.419	0.094	-4.449	0.000	-0.604	-0.235
violencespeechR	0.166	0.102	1.632	0.103	-0.033	0.365
sex_m	0.124	0.086	1.452	0.147	-0.044	0.292
age_18_24	0.331	0.240	1.380	0.167	-0.139	0.802
age_25_29	0.412	0.179	2.296	0.022	0.060	0.764
age_30_39	0.452	0.137	3.298	0.001	0.183	0.720
age_40_49	0.261	0.137	1.902	0.057	-0.008	0.530
age_50_64	0.234	0.115	2.033	0.042	0.008	0.460
raceth_wht	-0.659	0.180	-3.664	0.000	-1.012	-0.307
raceth_blk	0.687	0.245	2.810	0.005	0.208	1.166
raceth_hsp	-0.017	0.221	-0.078	0.938	-0.450	0.416
raceth_asn	-0.272	0.310	-0.876	0.381	-0.879	0.336
raceth_ntvam	-2.300	1.062	-2.166	0.030	-4.382	-0.219
raceth_hipac	1.374	0.826	1.662	0.096	-0.246	2.993
edu_highschoolorless	-0.422	0.158	-2.661	0.008	-0.732	-0.111
edu_somewhereassocdegree	-0.281	0.129	-2.179	0.029	-0.534	-0.028
edu_collegegraduate	0.044	0.125	0.351	0.725	-0.201	0.289
income_under25000	0.046	0.162	0.285	0.776	-0.271	0.364
income_25000_49999	0.008	0.131	0.060	0.952	-0.249	0.265
income_50000_74999	-0.169	0.128	-1.320	0.187	-0.421	0.082
income_75000_99999	0.178	0.124	1.439	0.150	-0.065	0.421
party	-0.501	0.032	-15.843	0.000	-0.563	-0.439
ideo	-0.161	0.050	-3.239	0.001	-0.258	-0.064
relig_prot_chrst	-0.373	0.164	-2.279	0.023	-0.694	-0.052
relig_catholic	-0.322	0.174	-1.856	0.063	-0.663	0.018
relig_mormon	0.275	0.357	0.771	0.441	-0.424	0.974
relig_jewish	-0.204	0.317	-0.643	0.520	-0.825	0.417
relig_muslim	-0.670	0.568	-1.181	0.238	-1.783	0.443
relig_none	0.086	0.180	0.476	0.634	-0.267	0.438
sizeplace_urban	0.325	0.137	2.368	0.018	0.056	0.594
sizeplace_suburban	0.210	0.116	1.807	0.071	-0.018	0.438
sizeplace_smalltown	0.273	0.130	2.092	0.036	0.017	0.528
Intercept	0.587	0.490	1.199	0.231	-0.373	1.547

GRAPHS

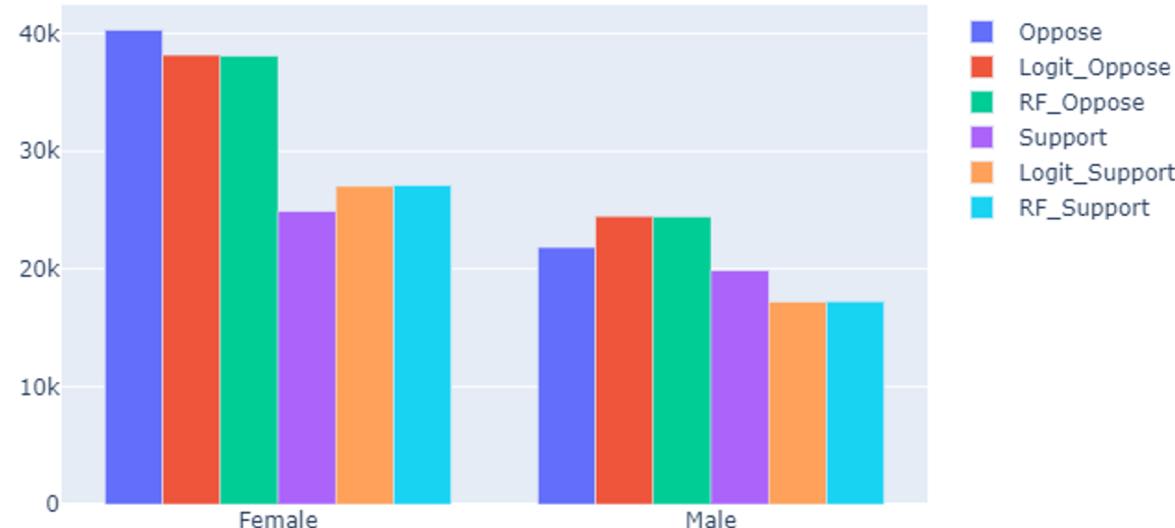
Trump Support
Both Models Appear to Match the Survey Questions



GRAPHS

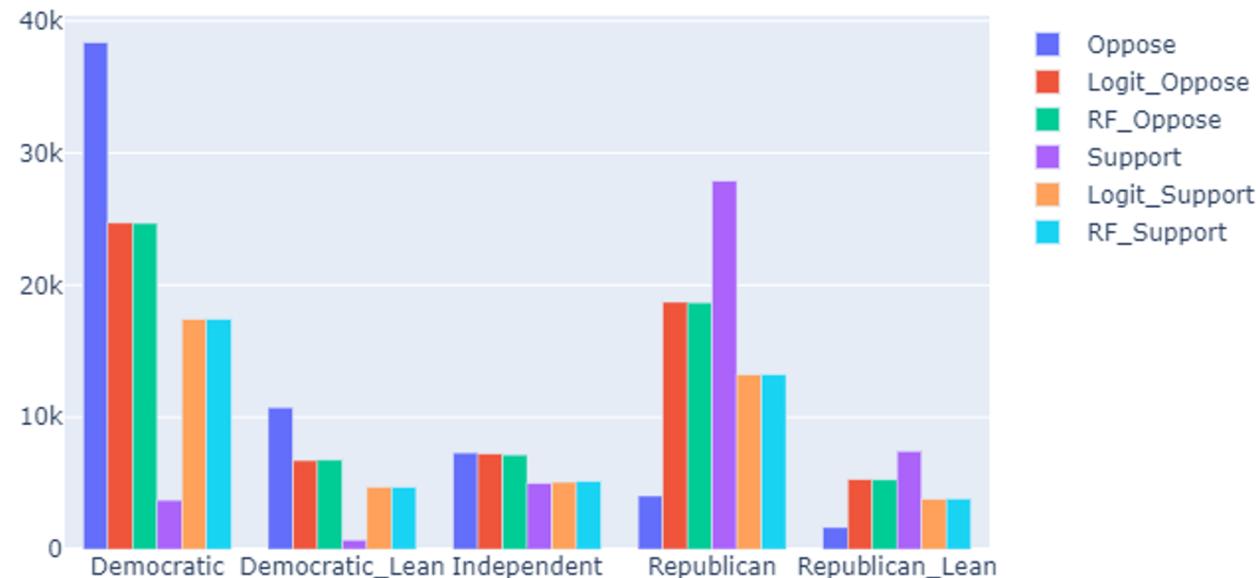
Trump Support by Sex

Both Models Made Accurate Predictions Across the Gender Divide, Although Some Differences Were Observed



GRAPHS

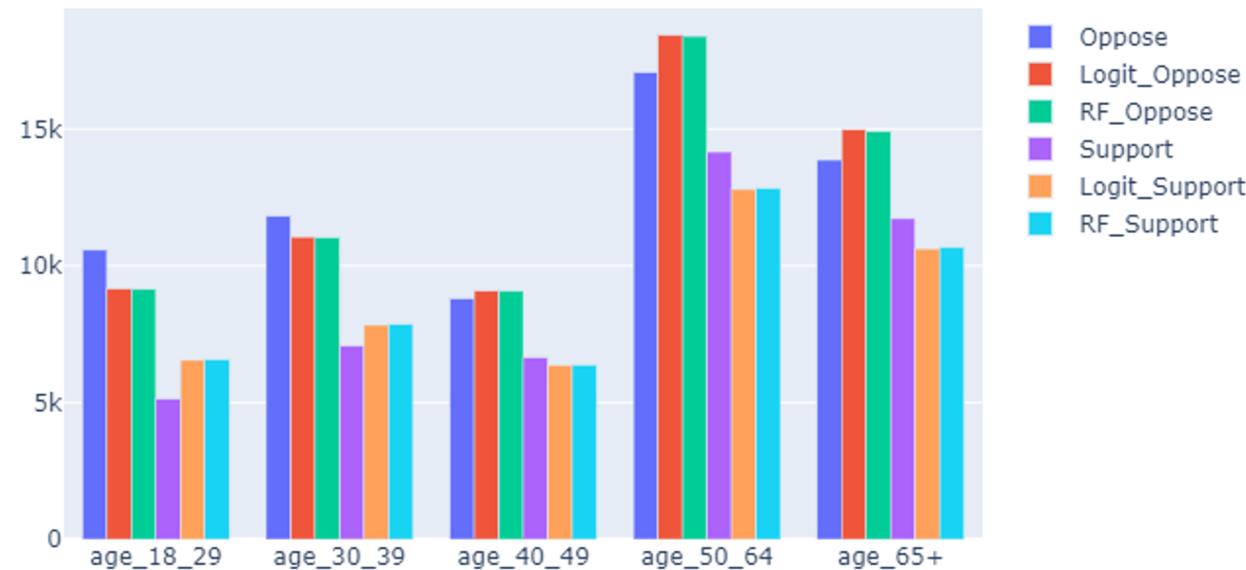
Trump Support and Models by Party:
Clear divide along party lines and Independents Slightly Opposed. Both Model



GRAPHS

Trump Support by Age

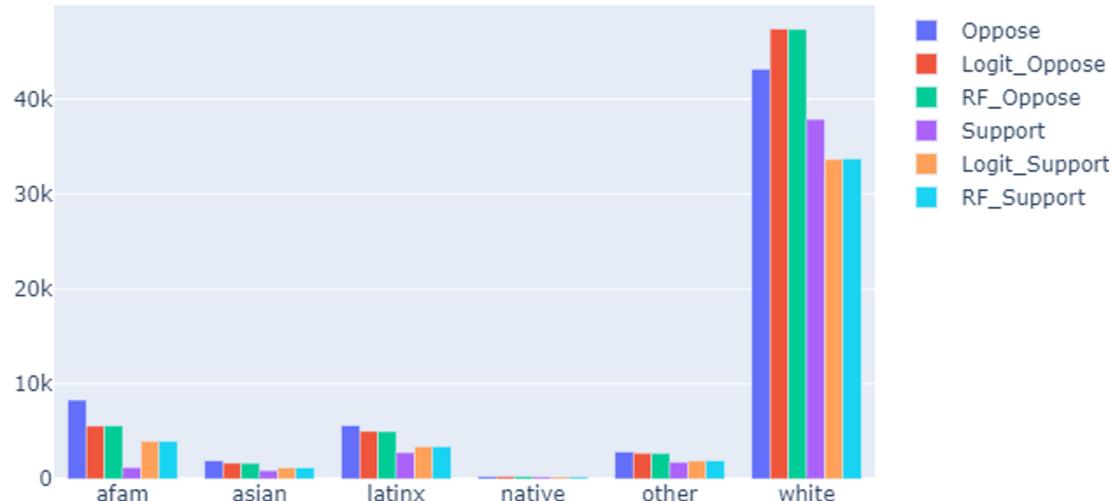
Both Models Tracked Well with the Survey Questions but Understated Support



GRAPHS

Trump Support by Ethnicity

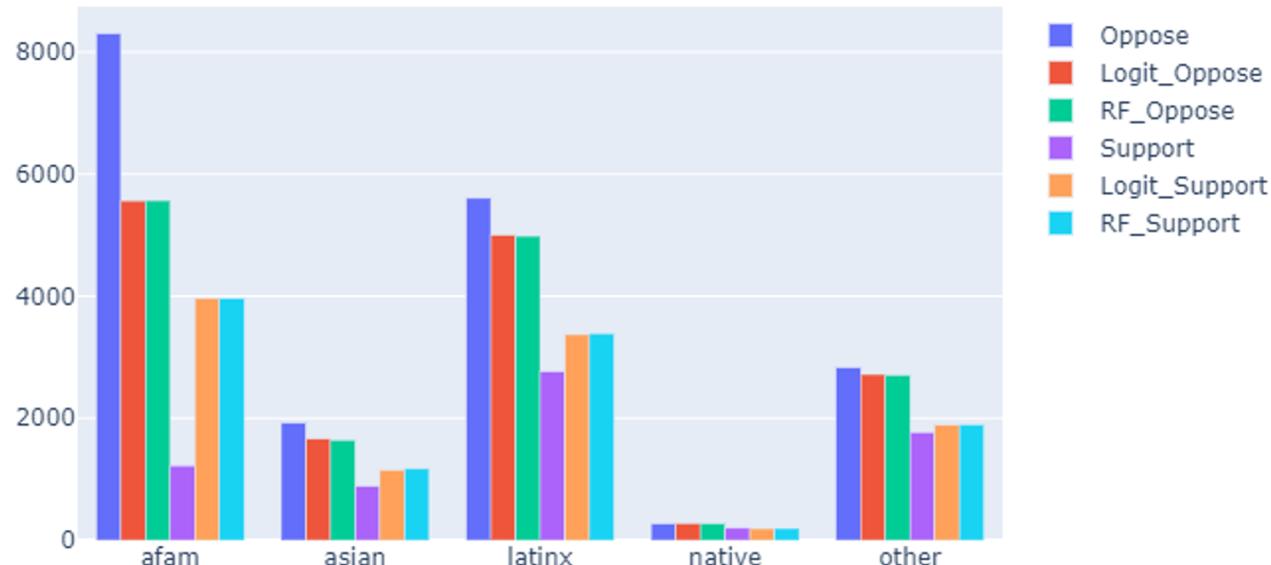
The Model Trends in the Direction Among White Voters, While Other Ethnicities



GRAPHS

Trump Support by Non-white Ethnicity

Among Non-white Ethnicities, the Models Follow the Survey Questions but Un





INTERACTIVE DASHBOARD

LIVE DEMO



USER INTERFACE

LIVE DEMO

CONCLUSION AND RECOMMENDATION

Voters in specific demographic or ideological groups are sometimes less willing to admit to supporting Donald Trump.

The campaign team of the 2020 Democratic nominee should estimate an individual's likelihood of supporting Trump based on information other than 2016 vote choice, since this is often underreported.

The best indicators include sex, age, education, religion, and specific issue opinions and beliefs (e.g., repeal/expand ACA, Russian collusion, deportation/naturalization of undocumented immigrants, direction of country).



LIMITATIONS

The data used for this project was collected leading up to the 2018 midterm, which preceded:

- Mexican border crisis, migrant detention, and family separation;
- Conclusion of Mueller investigation and report;
- Impeachment inquiry of Ukrainian phone call, impeachment, and acquittal; and
- COVID-19 outbreak and economic crisis.

As a result of these events, public opinion could have shifted immensely or not at all, so data should be consistently replenished and models updated to measure both the immediate and long-term impacts of current events, i.e., which controversies impact voter preference only for a short time, and which have more staying power.

REFERENCES

- Associated Press. (2020). AP VoteCast. Retrieved March 4, 2020, from Associated Press:
<https://www.ap.org/en-us/topics/politics/elections/ap-votecast/about>
- Klar, S., Weber, C., & Krupnikov, Y. (2017, February 22). Social Desirability Bias in the 2016 Presidential Election. *A Journal of Applied Research in Contemporary Politics*, 14(4).
- Lavrakas, P. J. (2008). Encyclopedia of survey research methods (Vols. 1-0). Thousand Oaks, CA: Sage Publications, Inc. doi: 10.4135/9781412963947
- Dashboard: <https://chart-studio.plot.ly/dashboard/kielrunner:144/view#/>