Supplementary Materials for "IRA Twitter activity predicted 2016 U.S. election polls".

Damian Ruck
University of Tennessee

Natalie Rice
University of Tennessee

Josh Borycz
University of Tennessee

Alex Bentley
University of Tennessee

April 24, 2019

Granger causality

We have shown that the success of Internet Research Agency (IRA) tweets R predicts a future increase in Donald Trump's polls T, but not Clinton's C. We also show that neither T nor C predicted future changes in R.

Here we show Granger Causality tests demonstrating that IRA Twitter success predicts future increases in Trump's opinion polls, but that statistical significance is not met in all other cases. Table 1 contains results where IRA tweet success Granger causes opinion polls and table 2 is where the opinion polls Granger cause IRA tweet success.

dep (C or T)	pred (S)	F-stat	P value	lag
adjpoll_clinton	retweet_count_mean	3.12	0.08	1
adjpoll_trump	retweet_count_mean	7.22	*0.01*	1
rawpoll_clinton	retweet_count_mean	0.2	0.65	1
rawpoll_trump	retweet_count_mean	4.5	*0.04*	1
adjpoll_clinton	like_count_mean	3.27	0.07	1
adjpoll_trump	like_count_mean	8.01	*0.01*	1
rawpoll_clinton	like_count_mean	0.17	0.68	1
rawpoll_trump	like_count_mean	4.8	*0.03*	1

Table 1: Granger causality tests for average IRA tweet success predicting election opinion polls. Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion.

1

dep (S)	pred (C or T)	F-stat	P value	lag
retweet_count_mean	adjpoll_clinton	0.65	0.42	1
retweet_count_mean	adjpoll_trump	0.38	0.54	1
retweet_count_mean	rawpoll_clinton	1.11	0.29	1
retweet_count_mean	rawpoll_trump	0.11	0.74	1
like_count_mean	adjpoll_clinton	0.76	0.39	1
like_count_mean	adjpoll_trump	0.94	0.33	1
like_count_mean	rawpoll_clinton	1.08	0.3	1
like_count_mean	rawpoll_trump	0	0.96	1

Table 2: Granger causality tests for election opinion polls predicting average IRA tweet success. Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion.

Top 25 most retweeted tweets from IRA accounts

retweets	tweet_text
97498	These emotions... You can hear the pain in her voice
41089	When Its slowly becoming illegal for black people to work #BlackTwitter
28568	!!!!!!DONT LET THIS GO UNNOTICED!!!!!!
25169	He didnt want a black nurse to help his dying child. And now his child is gone. Pathetic and ridiculous.
16670	#AnthonyCage was bloodied at St. Louis Trump rally. Spread this like wild fire #TrumpRallyChi #BlackToLive
16627	Normal??? Anti-racism poster gone horribly wrong...
16104	"His life mattered." #TerenceCrutcher's twin sister demands justice
15985	He really shot #AltonSterling while he was being held down... Imagine the reaction if this cop doesn't get indicted.
15548	BREAKING: #VoterFraud by counting tens of thousands of ineligible mail in Hillary votes being reported in Broward County, Florida Please, RT
15385	Every person on black twitter needs to rt this at least once, RIP queen, you wont be forgotten.
15376	Black TMZ staffer schools his white coworkers over The Weeknds hair. Today in black history. #blm
14664	The Angry Dark Skin Friend Ay this is very important
14433	Bus driver sexually assaulted teen girl. Black man defended her. Guess who was arrested.
13316	Wow. Hadn't thought of it that way but that's exactly what is happening. So true. #BlackLivesMatter
12716	OMG, this new Anti-Hillary ad is brilliant!It's fantastic!!!!!! Spread it far & wide!
12636	This really spoke to me #blm
12444	Im not sure Ive EVER enjoyed a Twitter story so much#BlackToLive
12171	Her Teacher cut off 1 of Lamya's beautiful braids as a punishment. Payed a small fine and kept her job #HATEIT
11838	Let this picture sink in. Maximize it. Zoom in. Stare at it. Take several moments. Now get angry. Be angry.
10848	Muhammad Ali, the only person whose Hollywood Walk of Fame Star is hanging on a wall, not for anyone to step on
10824	RT the hell out of it: Dem party operatives: 'We've been bussing people in.. for 50 yrs and we're not going to stop now' #EvangelicalTrump
10275	Some guy right in Hillary's face: HILLARY FOR PRISON Hillary: LET'S MAKE IT HAPPEN I almost feel sorry for her
10174	3 Black children carrying their daily water allowance. Flint, Michigan - 2016
9799	Javell Williams walked into a Columbus shooting range to be greeted by an African American teen as the target
9487	Spoiler alert: 2017 taken too

Table 3: 25 most retweeted IRA tweets

Top 25 National Pollsters

Pollster	# polls
Ipsos	975
USC Dornsife/LA Times	363
CVOTER International	345
The Times-Picayune/Lucid	282
Morning Consult	159
Rasmussen Reports/Pulse Opinion Research	153
SurveyMonkey	96
YouGov	84
IBD/TIPP	81
ABC News/Washington Post	72
Google Consumer Surveys	63
Fox News/Anderson Robbins Research/Shaw & Company Research	57
Gravis Marketing	54
NBC News/Wall Street Journal	45
CNN/Opinion Research Corp.	39
CBS News/New York Times	36
Quinnipiac University	36
Public Policy Polling	30
Monmouth University	24
Zogby Interactive/JZ Analytics	24
Public Religion Research Institute	21
Pew Research Center	21
RAND (American Life Panel)	18
Selzer & Company	18
Suffolk University	18

Table 4: 25 most frequent pollsters during 2016 presidential campaign

Robustness checks

In the main text, we showed that IRA tweet success predicted election opinion polls, whether we measure success as 'retweets per tweet' or 'likes per tweet'. Furthermore, it does not matter whether we use adjusted or raw polling data. Here we add more robustness checks and show we get the same results when we decrease the time resolution from seven to two days; use total number of retweets rather than average retweets and, finally, using the end date of the polls rather than the start date.

Two day time resolution

The finest time resolution we can attain for IRA Twitter activity and opinion polls is two days. Compared to the weekly time resolution analysis, the optimum VAR lag (determined by Akaike Information Criterion) has decreased slightly from 7 to between 4 and 6 days. As tables 5 and 6 show, the Granger causality structure is the same as the main analysis.

dep (C or T)	pred (S)	F-stat	P value	lag
adjpoll_clinton	retweet_count_mean	2.77	*0.04*	6
adjpoll_trump	retweet_count_mean	8.5	*0*	4
rawpoll_clinton	retweet_count_mean	0.29	0.75	4
rawpoll_trump	retweet_count_mean	8.91	*0*	4
adjpoll_clinton	like_count_mean	2.76	*0.04*	6
adjpoll_trump	like_count_mean	8.98	*0*	4
rawpoll_clinton	like_count_mean	0.31	0.73	4
rawpoll_trump	like_count_mean	9.22	*0*	4

Table 5: Two day time resolution: Granger causality tests for average IRA tweet success predicting election polls. Italic means statistical significance met at 5% level and lag is the optimum lag (in days) for VAR determined by Akaike Information Criterion

dep (S)	pred (C or T)	F-stat	P value	lag
retweet_count_mean	adjpoll_clinton	1.47	0.22	3
retweet_count_mean	adjpoll_trump	2.06	0.13	2
retweet_count_mean	rawpoll_clinton	1	0.37	2
retweet_count_mean	rawpoll_trump	0.95	0.39	2
like_count_mean	adjpoll_clinton	1.57	0.2	3
like_count_mean	adjpoll_trump	2.62	0.08	2
like_count_mean	rawpoll_clinton	0.79	0.45	2
like_count_mean	rawpoll_trump	1.39	0.25	2

Table 6: Two day time resolution: Granger causality tests for election polls predicting average IRA tweet success. Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion

Total number of retweets

In the main text, we assumed that the average retweets/likes per tweet was the appropriate measure of IRA twitter success. However, the operative quantity could feasibly be the total number of retweets/likes. Tables 7 and 8 show that the same Granger causal structure exists in both cases.

dep (C or T)	pred (S)	F-stat	P value	lag
adjpoll_clinton	retweet_count_sum	0.4	0.91	8
adjpoll_trump	retweet_count_sum	2.71	*0.01*	8
rawpoll_clinton	retweet_count_sum	0.96	0.48	8
rawpoll_trump	retweet_count_sum	1.94	0.07	10
adjpoll_clinton	like_count_sum	0.33	0.95	8
adjpoll_trump	like_count_sum	3.36	*0*	8
rawpoll_clinton	like_count_sum	1.26	0.28	8
rawpoll_trump	like_count_sum	2.52	*0.02*	8

Table 7: Total number of retweets: Granger causality tests for total IRA tweet success predicting election polls . Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion

dep (S)	pred (C or T)	F-stat	P value	lag
retweet_count_sum	adjpoll_clinton	0.23	0.98	8
retweet_count_sum	adjpoll_trump	0.4	0.91	8
retweet_count_sum	rawpoll_clinton	0.23	0.98	8
retweet_count_sum	rawpoll_trump	0.67	0.75	10
like_count_sum	adjpoll_clinton	0.4	0.91	8
like_count_sum	adjpoll_trump	0.26	0.98	8
like_count_sum	rawpoll_clinton	0.15	1	8
like_count_sum	rawpoll_trump	0.58	0.79	8

Table 8: Total number of retweets: Granger causality tests for election polls predicting total IRA tweet success. Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion

Polling end date (not start date)

Opinion polls take days/weeks to complete, so it is not clear whether we should use the start or the end date of the poll. In the main text we use polling start date, but tables 9 and 10 show the Granger causal structure is not changed when polling end dates are used instead.

dep (C or T)	pred (S)	F-stat	P value	lag
adjpoll_clinton	retweet_count_mean	4.35	*0.04*	1
adjpoll_trump	retweet_count_mean	10.09	*0*	1
rawpoll_clinton	retweet_count_mean	0.27	0.6	1
rawpoll_trump	retweet_count_mean	7.75	*0.01*	1
adjpoll_clinton	like_count_mean	2.65	*0.05*	3
adjpoll_trump	like_count_mean	11.05	*0*	1
rawpoll_clinton	like_count_mean	0.23	0.63	1
rawpoll_trump	like_count_mean	7.94	*0.01*	1

Table 9: Polling end date: Granger causality tests for average IRA tweet success predicting election polls. Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion

dep (S)	pred (C or T)	F-stat	P value	lag
retweet_count_mean	adjpoll_clinton	1.6	0.21	1
retweet_count_mean	adjpoll_trump	0.95	0.33	1
retweet_count_mean	rawpoll_clinton	0.79	0.38	1
retweet_count_mean	rawpoll_trump	0.05	0.82	1
like_count_mean	adjpoll_clinton	1.2	0.31	3
like_count_mean	adjpoll_trump	1.78	0.19	1
like_count_mean	rawpoll_clinton	0.78	0.38	1
like_count_mean	rawpoll_trump	0.03	0.87	1

Table 10: Polling end date: Granger causality tests for election polls predicting average IRA tweet success. Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion

Total number of tweets

Our hypothesis is that IRA Twitter *success* influenced U.S. election polls. Therefore, we should expect total number of unique tweets to be unrelated to the polls. To test this, we fitted an alternative VAR measuring IRA twitter activity as number of unique tweets per week.

We found no evidence that total number of IRA tweets predicted election polls (tables 11 and 12). If anything we see weak evidence for an effect in the opposite direction, suggesting the possibility that IRA twitter activity is increasing in response to Trump's polling.

dep (C or T)	pred (S)	F-stat	P value	lag
adjpoll_clinton	# of tweets	0.9	0.35	1
adjpoll_trump	# of tweets	0.12	0.73	1
rawpoll_clinton	# of tweets	0.21	0.65	1
rawpoll_trump	# of tweets	0.04	0.84	1

Table 11: Number of Tweets: Granger causality tests for total number of IRA tweets predicting election polls. Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion

dep (S)	pred (C or T)	F-stat	P value	lag
# tweets	adjpoll_clinton	0.61	0.44	1
# tweets	adjpoll_trump	3.33	0.07	1
# tweets	rawpoll_clinton	0.21	0.65	1
# tweets	rawpoll_trump	3.75	0.06	1

Table 12: Number of Tweets: Granger causality tests for election polls predicting total number of IRA tweets. Italic means statistical significance met at 5% level and lag is the optimum lag (in weeks) for VAR determined by Akaike Information Criterion

Controlling for success of Trump's personal Twitter account

Vector Auto-Regression (VAR) can only tell you if a changes in one time series (average IRA tweet success R) precede those in another (Donald Trump's opinion polls T). Therefore, to give us more confidence that the relationship is not spurious, we need to control for possible confounding variables.

One variable that could plausibly be causing changes in both Trump's opinion polls and IRA success is domestic Republican supporting U.S. media. Therefore, to control for this, we introduce a third time series for the average number of weekly re-tweets from Donald Trump's personal Twitter account P_t. We run the following revised VAR:

$$T_t \sim T_{t-1} + R_{t-1} + P_{t-1}$$
$$R_t \sim R_{t-1} + T_{t-1} + P_{t-1}$$
$$P_t \sim P_{t-1} + R_{t-1} + T_{t-1}$$

We see in figure 1 that introducing this control does not reduce the magnitude of the effect for average IRA retweets on Trump's polls (still around 10 re-tweets per tweet for every percentage point). However, the effect has become statistically insignificant because we now have a much larger error. This is because of collinearity caused by the very high correlation between the success of Trump and IRA tweets ($r = 0.8$). Figure 2 shows that Trump and IRA Twitter success are not merely correlated, but coevolving (both time series predict each other).

In conclusion. First, the effect of IRA retweets on Trump's polls is undiminished when controlling for the success of Trump's own personal Twitter account. Second, Trump and IRA Twitter successes are strongly correlated and coevolving.
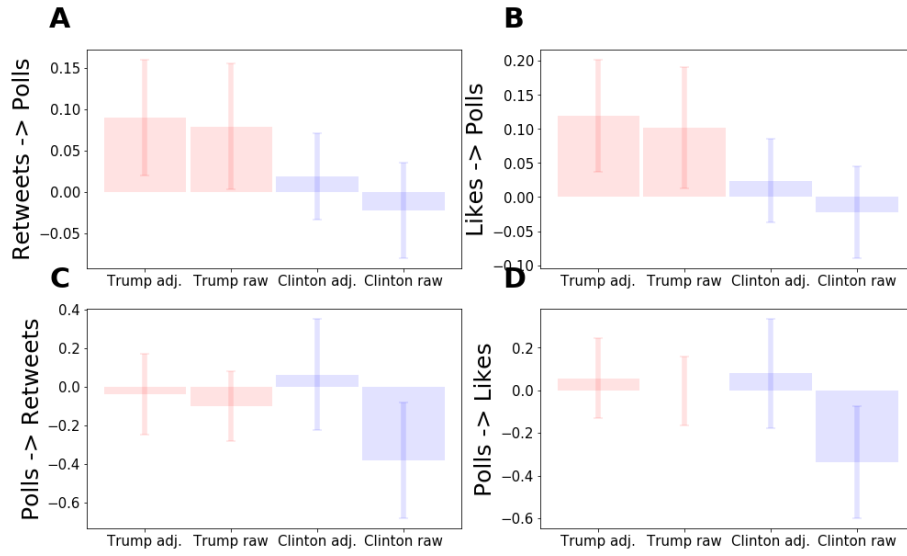
Figure 1: Vector Auto-Regression (VAR) for IRA twitter success and 2016 election polls, controlling for the number of retweets from Trump's personal Twitter account. (A) Trump and Clinton's opinion polls Granger caused by retweets per tweet and then (B) likes per tweet. Then (C) and (D) test for the reverse Granger causation. Height of bars are effect sizes, error bars are standard errors and bars are opaque if statistical significance was attained ($p < 0.05$).
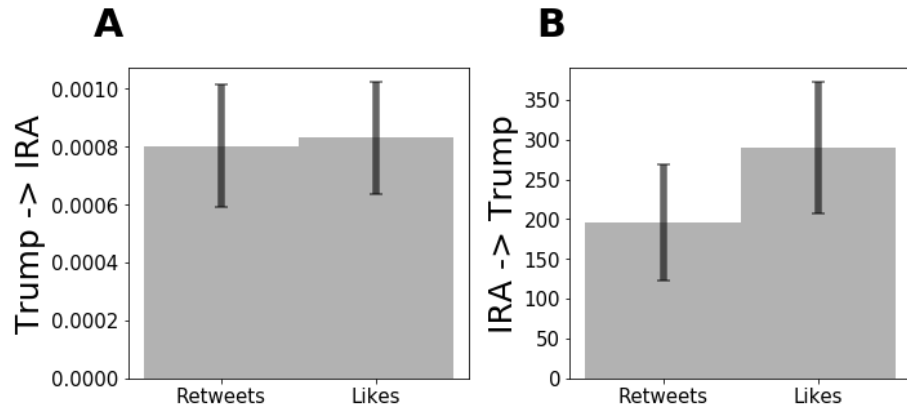

Figure 2: VAR results: (A) Trump tweet success predicts future IRA tweet success (likes and retweets) and (B) vice versa. Height of bars are effect sizes and error bars are standard errors. Statistical significance is met in all cases ($p < 0.05$).