Does Fake News Affect Voting Behavior?

Labor Lunch, Fall 2020

Cheng Yizhuang Alden

September 8, 2020

Growing Attention to the Influence of Fake News

- Fake news has become an increasingly hotly debated topic since the 2016 election.
- Numerous news outlets have claimed that fake news propagated via social media played a role in the election of Donald Trump.
  - "Yes, Russian Trolls Helped Elect Trump." Opinion piece by Michelle Goldberg in the New York Times.
  - "A new study suggests fake news might have won Donald Trump the 2016 election." The Washington Post.

Consequences for Social Media

- The debate over the spread of fake news has led to changes:
  - Facebook and other social media platforms have altered the way they present news information.
    - No more newsfeed sidebars on Facebook.
    - Stories on Facebook can be labelled as "disputed".
- Some policymakers think that these do not go far enough:
  - Some lawmakers are proposing that Facebook gets broken up.

But How Much Does Fake News Actually Affect Voting Behavior?

- A priori, it is not obvious fake news will have much of an effect on the presidential elections.
- If much of the voting population are as polarized as the media suggests, a few pieces of fake news may not be enough to sway many voters.
  - Partisans are likely to vote for their party regardless of whether they see fake news.
  - Independents may be more sophisticated than partisans, so they may be less likely to be fooled by fake news.

Studies on the Effect of Fake News on Voting

- There have been a few studies on whether fake news resulted in Trump winning the 2016 election.
- Boxell, Gentzkow, and Shapiro (2018) find suggestive evidence that fake news is unlikely to have played a major role in electing Trump.
  - In particular, they find that Trump did disproportionately well among voters who are least likely to use the Internet, compared to previous Republican candidates.
  - However, this does not draw a direct link between fake news consumption and voting behavior.
- Gunther, Beck, and Nisbet (2018) reach the opposite conclusion, based on correlational evidence from a survey of Obama voters.
  - Despite the headlines this study made, there is a very real possibility of omitted variables bias.

Threats to Internal Validity

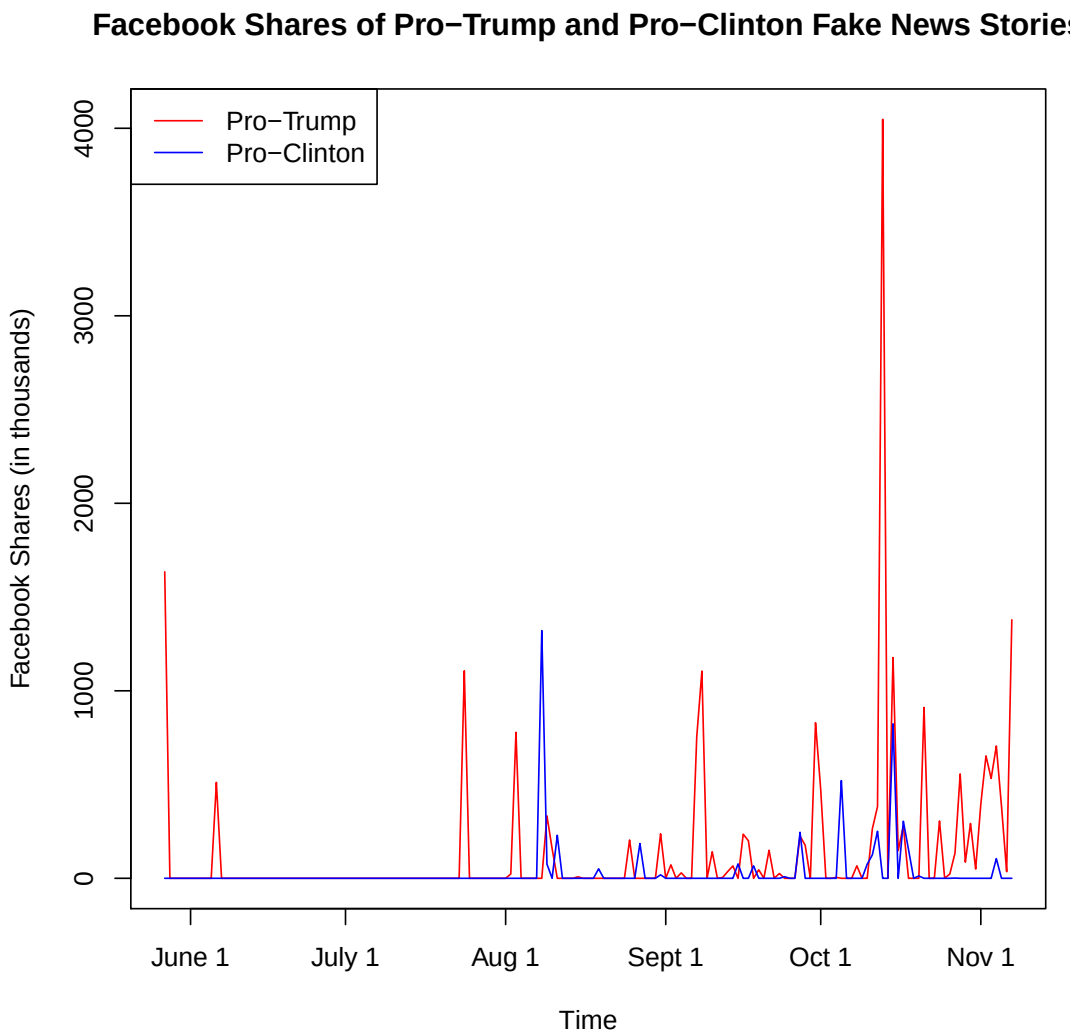
- Selection bias is a first-order threat to internal validity.
  - If voters who are more inclined to vote for Trump selectively choose to consume more fake news, the OLS coefficient will be biased upwards.
- In addition, most data on fake news consumption may also suffer from substantial measurement error.
  - For example, Allcott and Gentzkow (2017) find a high false recall rate for fake news stories in a survey they conducted.
  - While 15 percent of survey respondents recalled seeing a fake news story that surfaced around the election period, 14 percent also recalled seeing a fake news story that was never circulated (i.e. was invented by the authors).
  - Measurement error may bias the OLS coefficient towards zero.
- Solution to both these problems: IV!

Big College Football Games

What Does Football Have to do with Fake News Consumption?

- Consider voters living close to where a big college football game has just been played, or is about to be played.
  - Their social media feed may be dominated by sports-related news, crowding out concurrent fake news stories.
    - This is similar in spirit to a finding in Eisensee and Stromberg (2007) that when natural disasters occured as the Olympics or World Series is ongoing, they were covered less by the US news, and as a result the US sent less aid.
  - On the other hand, if they spend more time online due to the football game, it is also possible that they are exposed to more fake news.
- It may be unclear a priori which of these effects will dominate, but this is an empirical question that the first stage will shed light on.

Timeline of Fake News Outbreaks, Using Data from Allcott and Gentzkow (2017)



Time Window of Interest

- As a first pass, I consider the period from the start of October to October 27.
  - Comey announced the FBI was reopening its investigation of Clinton's emails on October 28, so I exclude the period from then until election day, so as to not conflate its effect with that of fake news.
  - There are many Facebook shares of both pro-Trump, and pro-Clinton fake news stories, but many more of the former.

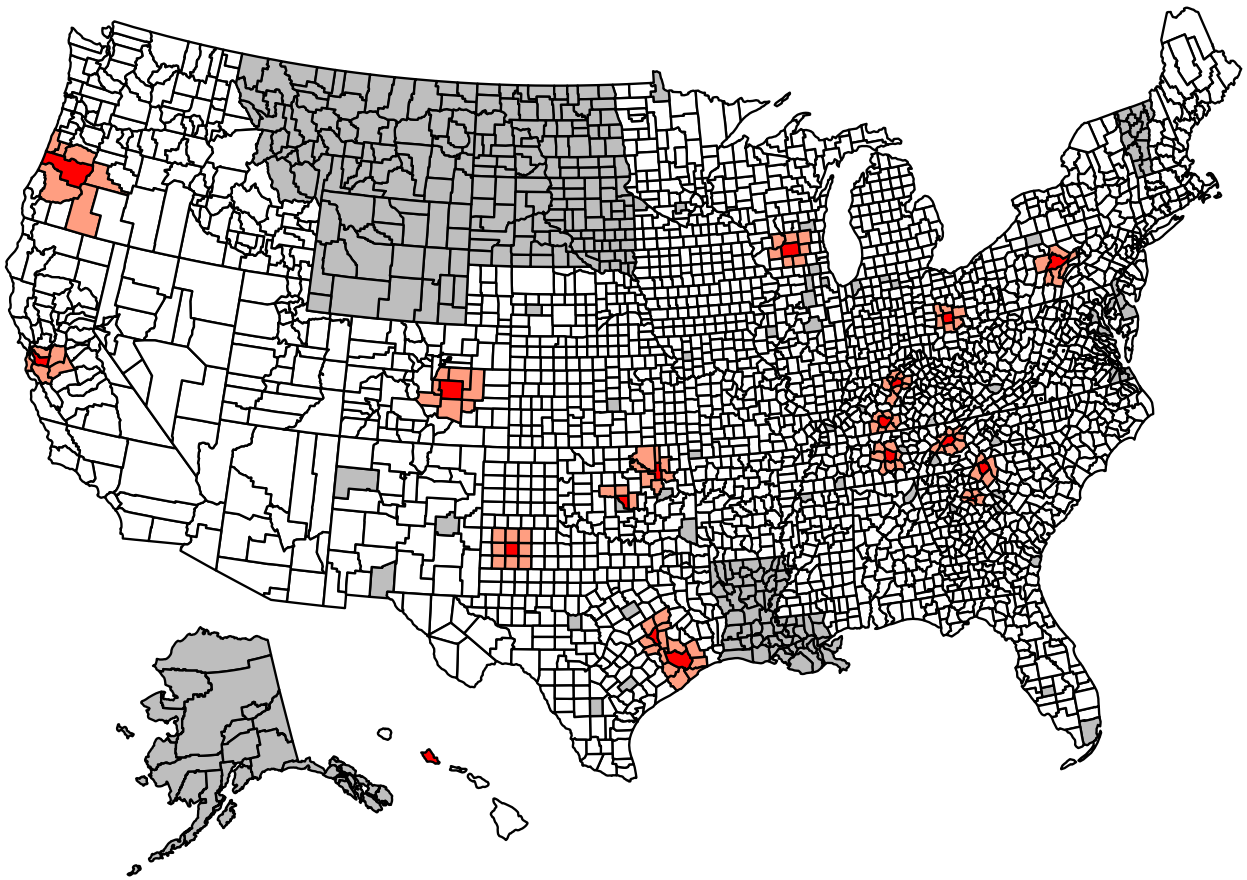
Election and Football Data

- I conduct my analysis at the county level.
  - Voting returns for the 2016 presidential elections is available at the county level.
- I collect football data by hand.
  - I start with 25 big football games according to an article by Sports Illustrated (https://www.si.com/college/2019/08/12/best-rivalries-college-football-history), and then limit this to games during the time period mentioned before.
  - I search for the county that each football team is located in, and merge this to the voting data.

Election and Football Data (continued)

- The procedure I just described returns 18 unique counties that were home to teams which played a big football game during this time window.
- We might worry that the sample with the instrument turned on is too small, so I consider an alternative definition.
  - In particular, I also include counties that are adjacent to one of the 18 counties in the earlier definition, yielding 125 counties.
  - For example, the county where Cal Football is located in is Alameda, but it seems likely that many residents of San Francisco city (less than half an hour's drive away, and classified as being in San Francisco county) also follow the team.

Counties Associated with Big Football Games



Data on Fake News Consumption

- As a proxy for fake news consumption, I use search data from Google Trends.
  - In particular, it seems quite unlikely that a user searches for terms associated with various fake news stories, unless she has already seen it somewhere.
  - I am able to obtain relative search volume data by DMA for certain terms related to fake news.
- Problems with Google Trends data:
  - The search volume data is based on sample of all searches, and this sample changes. So, I may get different results when I conduct the same search on different days.
  - There is a privacy threshold, so if there are not enough searches for the term in some DMA, Google Trends returns a missing value.
    - In practice, this happens quite often for various fake news terms, since there are relatively few searches for these terms (even during the sample period).

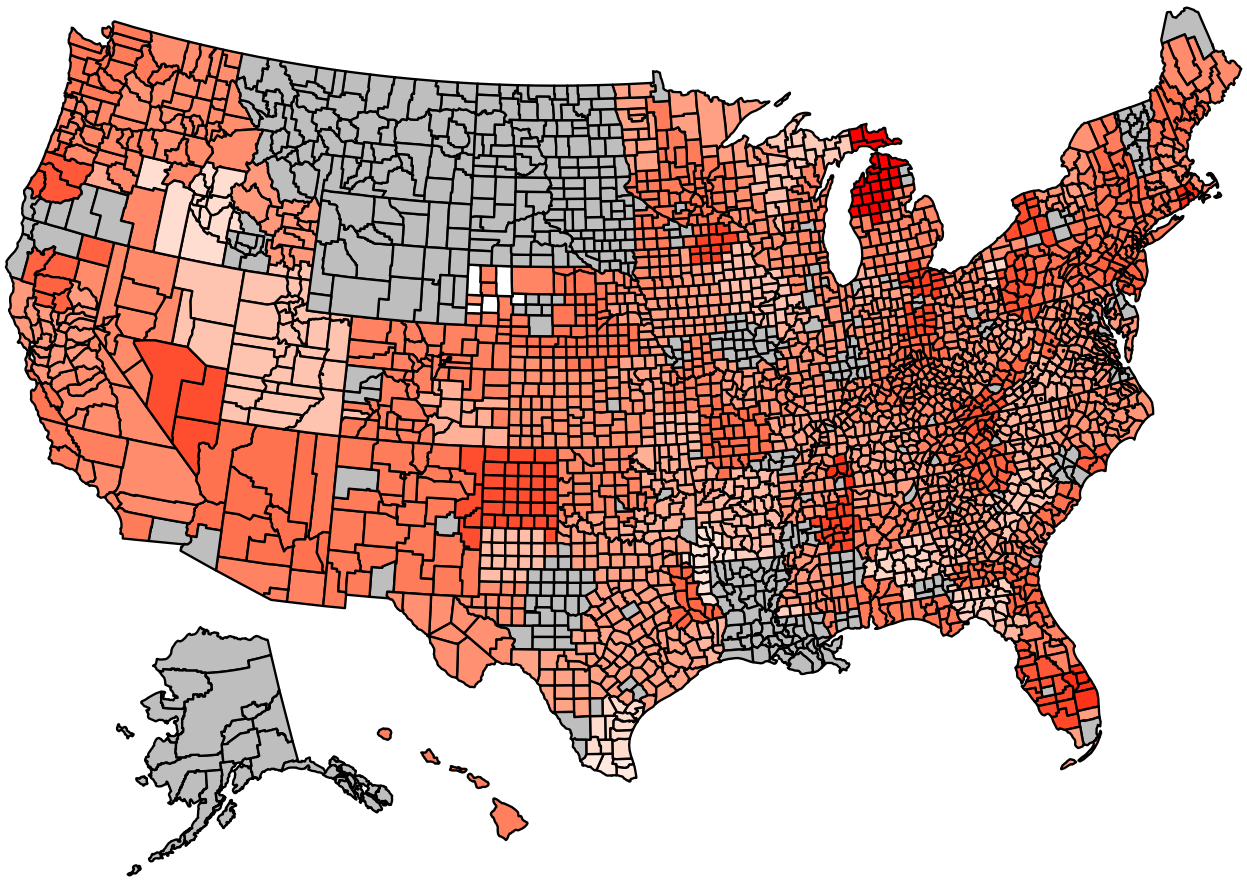
Data on Fake News Consumption (continued)

- To deal with the fact that search results vary, I conduct multiple searches over different days.
- In order to mitigate the missing values from the privacy threshold, I use a procedure from Stephens-Davidowitz (2014) and Oster (2018).
  1. Obtain the (relative) number of searches for a common term (call this term 1), e.g. chair.
  2. Get the number of searches for the term of interest (call this term 2) or the common term.
  3. Take the difference between the two:

$$[\textit{Searches for terms } 1 \textit{ or } 2] - [\textit{Searches for term } 1].$$

- I follow this procedure using several different common search terms, then average and standardize the results.

Search Intensity for Fake News Terms



Presidential Election (Reduced Form)

TABLE — Reduced Form Evidence of the Effect of Fake News on Voting Behavior

	Difference Between Trump and Clinton Vote Shares					
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy for Big Football Game in County	-0.114	-0.040	-0.029			
	(0.077)	(0.029)	(0.032)			
Dummy for Big Football Game in County or Adjacent Counties				-0.047	-0.040**	-0.031*
				(0.063)	(0.016)	(0.016)
Controls for Vote Shares in Past 4 Elections		X	X		X	X
Log Unemp. Rate (2015)			X			X
Number of Observations	2,747	2,747	2,747	2,747	2,747	2,747

Notes: Observations are weighted by the total population of the county. Robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Presidential Election (First Stage)

TABLE — First Stage Evidence on the Effect of Football Game on Fake News Consumption

	Standardized Searches for Fake News Terms					
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy for Big Football Game in County	-0.410**	-0.361**	-0.346**			
	(0.175)	(0.162)	(0.165)			
Dummy for Big Football Game in County or Adjacent Counties				-0.377**	-0.352**	-0.343**
				(0.153)	(0.150)	(0.151)
Controls for Vote Shares in Past 4 Elections		X	X		X	X
Log Unemp. Rate (2015)			X			X
F-Statistic	5.5	4.93	4.39	6.08	5.51	5.12
Number of Observations	2,747	2,747	2,747	2,747	2,747	2,747

Notes: Observations are weighted by the total population of the county. Robust standard errors clustered at the DMA level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Presidential Election (2SLS)

TABLE — OLS and Two-Stage Least Squares Estimates of the Effect of Fake News on Voting Behavior

	Difference Between Trump and Clinton Vote Shares								
	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Standardized Searches for Fake News Terms	-0.024	0.018**	0.017**	0.279	0.112*	0.083	0.125	0.114**	0.090*
	(0.025)	(0.009)	(0.008)	(0.240)	(0.058)	(0.069)	(0.318)	(0.048)	(0.051)
Instrument Based on Football County Only				X	X	X			
Instrument Includes Adjacent Counties							X	X	X
Controls for Vote Shares in Past 4 Elections		X	X		X	X		X	X
Log Unemp. Rate (2015)			X			X			X
Number of Observations	2,747	2,747	2,747	2,747	2,747	2,747	2,747	2,747	2,747

Notes: Observations are weighted by the total population of the county. Robust standard errors clustered at the DMA level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Possible Channels

- Assuming that these effects are causal, there are two likely channels through which fake news may increase Trump's vote share.
  1. Voters exposed to fake news may become disenchanted with the political system and decide not to turn out and vote.
  2. Consumption of fake news may change voters' decision of who to vote for.
- To test the first channel, I look at the effect of fake news consumption on voter turnout, obtained from David Leip's Atlas of U.S. Presidential Elections.
  - Unfortunately, party-specific voter turnout is unavailable at the county level, so I look at overall turnout.

Turnout (Reduced Form)

TABLE — Reduced Form Evidence of the Effect of Fake News on Voter Turnout

	Voter Turnout (Percent)					
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy for Big Football Game in County	-3.603	-1.693**	-1.969***			
	(2.631)	(0.690)	(0.715)			
Dummy for Big Football Game in County or Adjacent Counties				0.198	-0.390	-0.560
				(1.982)	(0.597)	(0.599)
Turnout in 2012 Election		X	X		X	X
Log Unemp. Rate (2015)			X			X
Number of Observations	2,612	2,612	2,612	2,612	2,612	2,612

Notes: Observations are weighted by the total population of the county. Robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Turnout (First Stage)

TABLE — First Stage Evidence on the Effect of Football Game on Fake News Consumption

	Standardized Searches for Fake News Terms					
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy for Big Football Game in County	-0.410**	-0.396**	-0.322*			
	(0.175)	(0.169)	(0.175)			
Dummy for Big Football Game in County or Adjacent Counties				-0.377**	-0.389***	-0.340**
				(0.153)	(0.147)	(0.150)
Turnout in 2012 Election		X	X		X	X
Log Unemp. Rate (2015)			X			X
F-Statistic	4.2	4.38	2.68	4.69	5.14	3.82
Number of Observations	2,747	2,678	2,678	2,747	2,678	2,678

Notes: Observations are weighted by the total population of the county. Robust standard errors clustered at the DMA level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Turnout (2SLS)

TABLE — OLS and Two-Stage Least Squares Estimates of the Effect of Fake News on Voter Turnout

	Voter Turnout (Percent)								
	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Standardized Searches for Fake News Terms	0.630	0.103	0.144	9.800	4.765	6.711	-0.592	1.152	1.881
	(0.808)	(0.216)	(0.213)	(6.432)	(2.985)	(4.704)	(9.003)	(2.115)	(2.508)
Instrument Based on Football County Only				X	X	X			
Instrument Includes Adjacent Counties							X	X	X
Turnout in 2012 Election		X	X		X	X		X	X
Log Unemp. Rate (2015)			X			X			X
Number of Observations	2,612	2,612	2,612	2,612	2,612	2,612	2,612	2,612	2,612

Notes: Observations are weighted by the total population of the county. Robust standard errors clustered at the DMA level are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Effect on House Elections

- There are two reasons why we might want to look at the effect of fake news on House elections:
  1. There may be down-ballot effects (e.g. a voter who believes anti-Clinton fake news may also view the entire Democratic party to be corrupt as a whole, and decide to vote against Democratic House candidates).
  2. On the other hand, one may think that down-ballot effects are unlikely, given how most of the fake news is targeted specifically at Clinton and Trump (rather than at their parties). In this case, the exercise below can be seen as a placebo check.
- Results of the House elections are only available at the congressional district (CD) level.
  - Hence, I use a county-CD crosswalk from the Census to aggregate up to the CD level (using population weights).
  - In cases where a county straddles multiple CDs, I divide it evenly between the CDs (dividing the population weights accordingly).

House Elections (Reduced Form)

TABLE — Reduced Form Evidence of the Effect of Fake News on Voting Behavior

	Difference Between Trump and Clinton Vote Shares					
	(1)	(2)	(3)	(4)	(5)	(6)
Big Football Game	0.037	0.028	0.031			
	(0.158)	(0.060)	(0.060)			
Big Football Game (Including Adjacent Counties)				0.023	0.044	0.047
				(0.109)	(0.036)	(0.036)
Controls for Vote Shares in Past 4 Elections		X	X		X	X
Log Unemp. Rate (2015)			X			X
Number of Observations	419	408	408	419	408	408

Notes: Observations are weighted by the total population of the county. Robust standard errors are shown in parentheses in italic font. Robust standard errors clustered at the DMA level are shown in parentheses with regular font. *** p<0.01, **

House Elections (First Stage)

TABLE — First Stage Evidence on the Effect of Football Game on Fake News Consumption

	Standardized Searches for Fake News Terms					
	(1)	(2)	(3)	(4)	(5)	(6)
Big Football Game Variable	-0.698***	-0.643***	-0.526***			
	(0.160)	(0.162)	(0.170)			
Big Football Game (Including Adjacent Counties)				-0.538***	-0.520***	-0.452***
				(0.101)	(0.105)	(0.112)
Controls for Vote Shares in Past 4 Elections		X	X		X	X
Log Unemp. Rate (2015)			X			X
F-Statistic	11.9	10.61	6.9	16.23	16.17	11.88
Number of Observations	420	409	409	420	409	409

Notes: Observations are weighted by the total population of the district. Robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

House Elections (2SLS)

TABLE — OLS and 2SLS Estimates of the Effect of Fake News on Voting Behavior in House Elections

	Difference Between House Vote Shares of Republican and Democratic Candidates								
	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Standardized Searches for Fake News Terms	-0.065**	-0.012	-0.013	-0.053	-0.043	-0.058	-0.042	-0.083	-0.102
	(0.027)	(0.014)	(0.014)	(0.177)	(0.090)	(0.113)	(0.153)	(0.075)	(0.089)
Instrument Based on Football County Only				X	X	X			
Instrument Includes Adjacent Counties							X	X	X
Controls for Vote Shares in Past 4 Elections		X	X		X	X		X	X
Log Unemp. Rate (2015)			X			X			X
Number of Observations	419	408	408	419	408	408	419	408	408

Notes: Observations are weighted by the total population of the district. Robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Conclusion

- Today, I presented a project on the effect of fake news consumption on voting behavior in the 2016 US elections.
- Any feedback is greatly appreciated!

Thank you!

References

- Alcott, Hunt, and Matthew Gentzkow, 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives*, 31(2): 211-236.
- Boxell, Levi, Matthew Gentzkow, and Jesse M. Shapiro. 2018. "A note on internet use and the 2016 US presidential election outcome." *PLoS One*, 13(7): e0199571.
- Eisensee, Thomas, and David Strömberg, 2-17. "News droughts, news floods, and US disaster relief." *The Quarterly Journal of Economics*, 122(2): 693-728.

References (continued)

- Oster, Emily. "Does disease cause vaccination? Disease outbreaks and vaccination response." *Journal of Health Economics*, 57: 90-101.
- Stephens-Davidowitz, Seth, 2014. "The cost of racial animus on a black candidate: Evidence using Google search data." *Journal of Public Economics*, 118: 26-40.