Angry Birds: Measuring and Predicting Affective Polarization Online Using Survey-linked Twitter Data

Haohan Chen*

Postdoctoral Fellow, CSCC, University of Pennsylvania haohanch@sas.upenn.edu

Brian Guay*

Ph.D. Candidate, Dept. Political Science, Duke University brian.guay@duke.edu

*Duke Polarization Lab

Affective Polarization

• "The tendency of people identifying as Republicans or Democrats to view opposing partisans negatively and copartisans positively" (Iyengar & Westwood, 2015)



Measurement of Affective Polarization

• Self-reports on surveys as measures

- Feeling thermometers (lyengar et al. 2012)
- Trait ratings (Levendusky & Malhotra 2016)
 e.g., selfish, intelligent, open-minded
- Social distance (Iyengar et al. 2012, Druckman 2019)
 e.g., feelings about child marrying a Democrat
- Some alternative measures (Iyengar & Westwood 2015, Carlin & Love 2013)
 e.g., implicit association test, behavioral games

Measurement of Affective Polarization

• Self-reports on surveys as measures

- Feeling thermometers (lyengar et al. 2012)
- Trait ratings (Levendusky & Malhotra 2016)
 e.g., selfish, intelligent, open-minded
- Social distance (Iyengar et al. 2012, Druckman 2019)
 e.g., feelings about child marrying a Democrat
- Some alternative measures (Iyengar & Westwood 2015, Carlin & Love 2013)
 e.g., implicit association test, behavioral games
- Social media activities as measures?
 - An emerging literature on measuring *ideological* polarization on social media e.g., Barbera (2015)
 - Little systematic investigation on measuring affective polarization with social media data

Research Questions

- How do social media activities vary across levels of affective polarization?
- How does social media text serve as a measure of affective polarization?
- How does social media text predict affective polarization?

Data: Survey-linked Twitter Data

- Survey of 1,239 U.S. Democratic (56%) and Republican (44%) Twitter users
- Collected by YouGov in October 2017
- Two measures of affective polarization
 - Feeling Thermometers: "How positively/negatively do you feel about Democrats/Republicans"
 - Adjectives: "How close-minded, generous, intelligent, and selfish are Democrats and Republicans?"

Survey Measures of Affective Polarization



Republican Respondents



Data: Survey-linked Twitter Data

- Scraped respondents tweets over a 10 year period (N = ~2 million)
- Summary statistics of tweets from past 12 months
 - 877,258 tweets
 - 1,105 respondents with at least one tweet
 - Median # of tweets = 251 300 -
 - Mean # of tweets = 794
 - Max = 10,008



How do social media activities vary across levels of affective polarization?

Affectively Polarized Respondents Tweet More

Number of Tweets



Ideological "Echo Chamber" and Affective Polarization

Proportion of Network With Similar Political Views



Ideological "Echo Chamber" and Affective Polarization

Proportion of Network With Similar Political Views



Can we measure affective polarization with text of tweets?

Constructing Measures of Affective Polarization

- Step 1: Identifying political tweets: supervised learning
 - Human coders labelled 7,400 tweets
 - Supervised machine learning (Random Forest, F1 = 0.91)
- Step 2: Identifying partisan tweets: seed dictionary + word embedding
 - Started with a "seed dictionary" of partisan words
 - Trained word embeddings with all the tweet text
 - Augment the dictionary by looking up closest words and phrases in the embedded space
- Step 3: Coding sentiments in tweets: external dictionary
 - Crowd-sourced sentiment-emotion dictionary by Mohammad (2013, 2014)
 - Manual validation by the authors

Twitter Behavior and Affective Polarization

Proportion of Tweets About Politics



Twitter Behavior and Affective Polarization



16

Negative Tweets About Parties





Negative Tweets About In-Party

Positive Tweets about Parties



Positive Tweets About In-Party

100

Can we predict affective polarization with text of tweets?

Predicting Affective Polarization with Tweets

- We searched for the best predictors of affective polarization from tweets
 - Experimented with half of a million models with a variety of specifications
 - Selected 1,296 best models for evaluation and interpretation
- Models: Linear Model with L1 Regularization (LASSO)
 - Variable selection: N of respondents = 1,239, P up to 18,329
 - Interpretable coefficients: Focus on selected features with negative coefficients
- Why not more complex models (e.g., ensemble models, neural network)?
 - The best linear models perform reasonably well
 - Interpretability
 - Small dataset

Design of Experiment

All combinations of 5 types of specifications

- For each specification: Parameter grid search with 10-fold cross-validations
- Evaluated with RMSE of out-of-sample prediction

One/Two steps	One-step: One model for dems and reps Two-step: Predict party ID \rightarrow Separate models for Dems and Reps	
Outcomes	Feeling thermometer Adjectives Combination	
Targets	Democrats Republicans	
Features	Full text (trigram) @ only # only URLs only @ + # @ + # + URLs	
Time	~ 2 months to 2 years Tweets 60, 120, 180, 720 days before survey (60-day intervals)	

Our Best Models: Two-step models

Step 1: Predict Party ID (Regularized Logit, Out-of-sample F1 = 0.81)

Step 2: Use model 1,2 or 3,4 based on predicted party ID from Step 1

Model #	Predicted Party ID	Target of Affect	Out-of-sample RMSE (outcome scale 0 - 100)
1	Democrats	In-party	13.2
2	Democrats	Out-party	13.7
3	Republicans	In-party	16.1
4	Republicans	Out-party	17.0

Predictors of Negative Outparty Feeling in Text (Average negative coefficients of top 30 models)

bankrupt mike pence ieff sessions collinsImpeach equality hates aeper impac ment racism supremacist^{sir}hitchted cr straight, Snfl players boycottlands^{trump's} might RT^{conscience}

national anthem anti-american democrat wikileaks illegal immigrants instant winfake news media fake news illegals obama americans muslims liberals terrorists exposed destroy E a to tat sides E obamacare õ≡ ⊙maxine liberal illegal immigration impeached

Predictors of Negative Outparty Feeling in @ and # (Average negative coefficients of top 30 models)





Predictors of Negative Outparty Feeling in URLs (Average negative coefficients of top 30 models)

motherjones vanityfair thinkprogress.org klnk. bbcsalon newsweek theguardian m.youtube huffingtonpostyahoo shareblue slate theatlantic rawstory dailykos politicususa slingshot.rafflecopter thedailybeast

conservativetribune insider.foxnews today.yougov zpacreview nationalreview americanthinker dailymail.co.uk**TO** truthfeed IITezette ibotta dailywire foxbusiness a.msn_{reuters} telegraph.co.uk thegatewaypundit

Predicting Party ID Before Predicting Affect Boost Performance



26

Predict Democrats' Affect With Smaller Error Than Republicans'

60 subset 40 count Democrat Republican 20 0 16 20 24

RMSE

Models predicting democrats' affect perform better

Predict In-party Feelings with Smaller Error Than Out-party



Models predicting in-party affect perform better

Combining Feeling Thermometers and Adjectives as Outcome Boosts Performance



Models predicting the 'combo index' perform the best

Predictors: @, #, and URLs Shared Reveal Affective Polarization



Full text brings small gain in model performance

Time: More tweets = More Signals + Noise

Including tweets from a longer period not render better performance



31

Take-Aways

- Individuals with high levels of affective polarization...
 - tweet much more about politics and both political parties
 - have much more ideologically homogeneous Twitter networks
- Strong correlation between affective polarization and negative sentiment toward parties, particularly for out-party
- Models predicting affective polarization using Tweets, hashtags, URLs, and mentions perform reasonably well, even using a small set of text features and relatively few tweets
- Future work will incorporate more data (e.g., incorporate ideological networks, user profiles) into predictive models

Feedback and questions are welcome!

Haohan Chen*

Postdoctoral Fellow, CSCC, University of Pennsylvania haohanch@sas.upenn.edu

Brian Guay*

Ph.D. Candidate, Dept. Political Science, Duke University brian.guay@duke.edu

*Duke Polarization Lab