ipl immigration policy lab

Working Paper Series

Working Paper No. 22-01

Tracking Legislators' Expressed Policy Agendas in Real Time

Alexandra A. Siegel, David D. Laitin, Duncan Lawrence, Jeremy M. Weinstein, and Jens Hainmueller

Tracking Legislators' Expressed Policy Agendas in Real Time[*]

Alexandra A. Siegel[a,b,c,1], David D. Laitin[b,d], Duncan Lawrence[b], Jeremy M. Weinstein[b,d], and Jens Hainmueller[b,d,e]

[a]Department of Political Science, University of Colorado at Boulder
[b]Immigration Policy Lab, Stanford University
[c]Center for Social Media and Politics, New York University
[d]Department of Political Science, Stanford University
[e]Graduate School of Business, Stanford University
[1]Corresponding author: alexandra.siegel@colorado.edu

January 18, 2022

Abstract

We develop a real-time scalable method to analyze strategic communication by political actors on salient policy issues through their tweets. Using word embeddings and supervised machine learning models, we classify legislators' tweets according to whether or not they reference policy issues as well as what positions they promote. This allows us to measure the microdynamics of elite communication with a high level of temporal granularity in a manner that is scalable across diverse issue areas and legislatures. As a proof of concept, here we use this method to track the multi-year evolution of positions that members of Congress express on immigration and climate change. Validation with issue-specific vote scores suggests that our method performs with a satisfactory level of accuracy and enables us to identify legislators whose online rhetoric differs substantively from their voting behavior. We also display the immigration analysis on an automatically updated publicly available interactive website, enabling researchers, journalists, and policy makers alike to explore legislators' shifting rhetoric on immigration in real-time.

1   Introduction

Despite charges of "cheap talk" (Farrell 1995), communication plays a critical role in elected officials' ability to remain in office, to shape policy debates, and to pass legislation (Sellers 2009).[1] Indeed, legislators devote substantial resources to public communication with constituents, other elected officials, interest groups, and members of the news media (Grimmer 2010). Strategic communication—or shaping and targeting "messages to maximize their desired impact while minimizing undesired collateral effects" (Manheim 1994) has become an integral part of legislators' day to day activities. From press releases and impromptu speeches to tweets and Facebook posts, elected officials use diverse forms of strategic communication to shape policy debates both within the legislature and in the public sphere more broadly. This enables them to gain electoral advantage and acheive their preferred policy outcomes (Jerit 2006; Riker 1986; Rojecki 2008). Legislators' daily strategic communication thereby offers a window into how parties and individual politicians take ownership of particular policy issues, as well as how their framing of these issues shifts over time.

Research spanning decades has highlighted the importance of legislators' public communication for understanding political representation and legislative behavior (Fenno 1978; Mayhew 1974; Franklin 1991; Vavreck 2009). Advances in communication technologies, data collection and text analysis methods have fostered a growing body of literature systematically analyzing strategic communication by legislators in the U.S. and other democratic contexts. Using manual human coding as well as supervised and unsupervised machine learning methods, these studies analyze legis-

---

[1] Replication materials will be posted at: Harvard Dataverse

lators' stump speeches (Fenno 1978), campaign mail (Golbeck, Grimes and Rogers 2010), television advertising (Lau, Sigelman and Rovner 2007), floor speeches (Martin and Vanberg 2008; Martin 2011; Quinn et al. 2010), press releases (Grimmer 2010, 2013b; Grimmer, Westwood and Messing 2014; Grimmer 2016, 2013a; Catalinac 2018; Klüver and Sagarzazu 2016), websites (Adler, Gent and Overmeyer 1998; Anstead and Chadwick 2008; Wilson 2009; Druckman, Kifer and Parkin 2009), RSS feeds (Cormack 2013), and social media posts (Gulati and Williams 2010; Barberá et al. 2018; Radford and Sinclair 2016; Kousser 2019; Shapiro et al. 2014; Lilleker and Koc-Michalska 2013).

Given the consistent finding in the political science literature that citizens tend to support the positions their elected officials take—whether through elites responding to citizens' preferences (Downs et al. 1957; Loewen and Rubenson 2011), citizens being persuaded by elites (Chong and Druckman 2007a,b; Grose, Malhotra and Parks Van Houweling 2015; Jacoby 2000; Kinder and Sanders 1990; Tesler 2015; Zaller et al. 1992), or citizens deferring to elite positions (Broockman and Butler 2017; Achen and Bartels 2006; Bartels 2005; Leeper 2013; Lenz 2009; Mackie and Cooper 1984)—the policy positions that elected officials express offer insights into the policy making process and democratic politics more generally.

To facilitate the study of elite strategic communication, we develop a real-time scalable method for tracking elected officials' expressed policy positions[2] through their tweets. Using a combination of word embeddings and supervised machine learning models, we classify legislators' tweets according to whether or not they dis-

---

[2]Depending on the legislator, this includes either direct communications or those produced by designated staff.

cuss policy issues as well as what positions they promote. We use this method to create an automatically updated publicly available interactive website, which updates weekly, enabling researchers, journalists, and policy makers alike to explore legislators' shifting rhetoric over time from 2013 to present.

Legislators across the globe are increasingly using social media to communicate with their constituents, other legislators, and the news media. For example, as of 2018, every member of Congress has an active Twitter account (Golbeck et al. 2018). The majority of elected officials at the national level in Europe and in the EU parliament similarly have Twitter accounts (Scherpereel, Wohlgemuth and Lievens 2018; Kat Devlin and Cha 2019), and the platform is also widely used by legislators in Latin America (Munger et al. 2018), and Asia (Yoon and Park 2014).

Elected officials primarily use Twitter to advertise their policy positions, to provide information about their activities (Golbeck, Grimes and Rogers 2010), to set political agendas, to interact publicly with other elected officials (Evans, Cordova and Sipole 2014), to communicate with constituents (Hemphill, Otterbacher and Shapiro 2013), and to increase media coverage of their messages and policy agendas to be covered by journalists and media outlets (Paulussen and Harder 2014; Graham et al. 2013).

Our approach of leveraging Twitter data to study legislators' strategic communication offers several advantages. First, it is highly scalable. It can be used to track shifts in legislators' rhetoric on any topic of interest, by any political actor with a Twitter account, in any country around the world, from the past decade or into the future. Such scalability is difficult to achieve with more conventional methods such

as analyzing floor speeches or press releases—or even less regularly updated RSS feeds and candidate websites—without sacrificing validity (Barberá et al. 2018).

Second, validation of our measure using issue-specific vote scores suggests that our relatively simple, transparent, and interpretable approach to tweet classification can achieve satisfactory levels of accuracy across diverse issues. This stands in contrast to most methods of tracking sentiment on social media, which tend to rely on out-of-the-box sentiment analysis tools, unsupervised approaches, and black-box or proprietary methods that can limit interpretability.

Third, Twitter data has a high level of temporal granularity. Given its low cost of access and use, the majority of legislators tweet on at least a daily basis, and many tweet multiple times per day, employing staff to manage and regularly update their social media accounts (Cluverius 2012). This means that we can see how legislators' strategic communication evolves—particularly in response to events of interest—down to the second.

Fourth, we can automate data collection, enabling us to continually update our analysis. Legislators' tweets can be collected as soon as they are posted using Twitter's API at no cost, and researchers can also access historical data, meaning that any study of legislators' strategic communication can be extended to include more legislators, longer time horizons, or different types of content, in any location of interest. Moreover, our data collections can be expanded in the aftermath of elections to include new legislators and our machine learning classifiers can be trained with new data at low cost, enabling us to continue to measure legislators' policy positions with a high level of temporal granularity into the future.

Finally, while most analysis of legislators' strategic communication measures the salience of particular policy issues (Grimmer 2010; Barberá et al. 2018; Munger et al. 2018; McLean and Bartula-Henkle 2019), we go a step further to measure not only shifting issue salience but also how the positions that legislators express on particular issues change over time. While past studies have leveraged social media data to measure longitudinal shifts in publicly expressed mass policy attitudes (Bartlett and Norrie 2015; Williams and Burnap 2015; Guidry et al. 2018; Flores 2017; Grover et al. 2019), mass sentiment or happiness (Hollander and Renski 2015; Tumasjan et al. 2010; Jashinsky et al. 2014; Mitchell et al. 2013; Yu and Wang 2015), or mass attitudes towards products or businesses (Hornikx and Hendriks 2015; Arora, Li and Neville 2015) no study to our knowledge has examined changes in the positions elected officials express over time. This enables us to track broad policy trends as well as how individual legislators' rhetoric may differ strategically from their voting behavior.

2   Building the Congressional Tweet Tracker

Figure 1 displays a schematic detailing the key features of our real-time tracker of legislators' policy positions. There are several steps involved in building the tracker. On the left, in the "Data Progression" column of the chart, we display how our data evolves from collecting legislators' historical tweets to developing our real-time tracker. On the right, in the "Processes and Methods" column we summarize the empirical approaches used at each stage of data collection and analysis.

Figure 1: **Real-time tracker of legislators' expressed policy agendas**. Figure shows a schematic of the main stages to build a real-time tracker of legislators' expressed policy agendas.

Data Progression

Legislators' Past Tweets
Legislator's tweets from 2013-2019

Relevant Tweets
All tweets related to issue area (e.g. immigration)

Training Data
Human-coded data to train classifiers

Classified Relevant Tweets
Relevant tweets classified on issue dimension of interest

New Daily Data
Automatically labeled new data from real-time collection (1/2019-present).

Real-Time Expressed Agenda Tracker
Daily updated plots of issue salience and position frequency from 2013-present.

Processes and Methods

IDENTIFY ACCOUNTS
Identify elected officials' Twitter handles for each country and time period.

Get Pre-Existing Lists
Use everypoliticianR package to find elected officials' Twitter handles.

Crowdsource Missing Data
Use Figure8 to crowdsource the identification of missing handles.

GET PAST TWEETS
Use the Historical Powertrack API to get officials' past tweets.

GET RELEVANT TWEETS
Use a word2vec model trained on full dataset of tweets to build dictionaries to identify relevant tweets.

Dictionary Validation
Validate dictionaries using random samples of tweets from the full dataset and expert coding.

TRAIN CLASSIFIERS
Use human coded training data to train classifiers that tag tweets according to whether they express preference for the issue position of interest or not (e.g. inclusive (or exclusive) towards immigrants/refugees or not)

Code Training Data
Use Figure8 to crowdsource human coding of relevant tweets

Classifier Choice and Validation
Use cross validation and compare precision, recall, accuracy, balanced accuracy, and F1 scores to choose optimal classifier.

GET REAL-TIME DATA
Use Twitter's streaming API to start real-time collections of tweets. Automate tagging of real-time data using word2vec dictionaries and classifiers trained on recent data

Measure Salience
Relevant Tweets / Total Tweets

Add New Legislators
Add new legislator handles to real-time collections after elections.

BUILD / UPDATE PLOTS
Plot salience & issue positions over time by party, region, or other covariates of interest.

Measure Issue Positions
Issue Position A / Relevant Tweets
Issue Position B / Relevant Tweets

Update Dictionaries/Classifiers
Retrain dictionaries and classifiers annually.

6

First we use the `everypoliticianR` package[3] in `R` (R Core Team 2021) to find lists of elected officials and their Twitter handles. If the social media handles of particular legislators are missing, then we use Figure8 (which has since been acquired by Appen), a crowd-sourced data coding platform, to have crowd workers identify legislators' Twitter handles through Google searches and advanced searches on Twitter. Three crowd workers identified each missing Twitter handle to ensure accuracy. After identifying legislators' Twitter handles, we use Twitter's Historical Powertrack API to download legislators' past tweets. This gives us our first dataset of all legislators' tweets from any period of interest up to the time at which we begin our analysis.

We then use a word2vec model (Mikolov et al. 2013) to build enhanced dictionaries to find tweets referencing a topic of interest (for example, immigration). In particular, we begin with a set of stemmed seed words that we identify as being relevant to the concept of interest (e.g. "migrant", "immig", "refugee"). We then use word embeddings (trained on the entire corpus of legislators' tweets) to identify other words that are semantically related to our seed words in the data. Semantic similarity here is based on these words appearing in similar contexts, and can be computed using cosine similarity on the word embedding space (Gurciullo and Mikhaylov 2017). These dictionaries are then limited to the 100 most similar words and we remove overly general or irrelevant terms.[4] We validate these dictionaries by taking random samples of tweets from the full dataset to ensure that our dictionaries do not cause

---

[3]https://github.com/ajparsons/everypoliticianR
[4]This choice of a 100 word threshold is arbitrary, but when we choose a larger threshold the additional terms tend to be irrelevant or rare in our data.

us to systematically miss relevant tweets, as well as that the tweets identified using our dictionary method are in fact relevant to the policy issue of interest.

Next we take a random sample of relevant tweets and use crowd-sourced human coders on Figure8 to classify relevant tweets according to the policy position they express (e.g. inclusive, exclusive, or neither, position toward immigrants/refugees). We then use this human coded data to train classifiers to classify tweets according to the policy position they promote. We use five-fold cross validation and compare precision, recall, accuracy, balanced accuracy, and F1 scores to choose the best performing classifier.[5]

We then start real-time collections of legislators' tweets (from January 2019 forward), which are automatically updated.[6] We classify these tweets as relevant or not using the word2vec enhanced dictionary methods and then use our classifier to label all legislators' past and current relevant tweets (e.g. tweets about immigration) according to whether they express a particular issue position (e.g. inclusive, exclusive, or neither position toward immigrants/refugees). This gives us all of the data we need to build our tracker or real-time plots of issue salience (e.g. daily proportion of legislators' tweets that are relevant to immigration) as well as the prominence of particular issue positions (inclusive vs. exclusive toward immigrants/migrants) over time. These plots can also be disaggregated by party, region, or other covariates

---

[5]See Supplementary Materials for details.

[6]This was originally done using Twitter's Streaming API. All data in this paper from January 2019 forward was collected using the Streaming API. After NYU's Center for Social Media and Politics gained access to the decahose (Twitter's 10% streaming sample) in March of 2020, we began to use the decahose to collect realtime data for our automatically updated website. Because legislators' tweets almost always receive at least 10 retweets, the decahose enables us to collect every original tweet produced by legislators from March 2020 forward.

or interest. They can further be automatically updated by filtering new real-time tweets through our word2vec dictionaries and using our classifier to categorize them. The list of legislators' tweets being collected can be updated in the aftermath of elections and the classifiers can be retrained periodically using new human coded data. Further details about the data, measurement, and classifiers can be found in the Supplementary Materials.

To illustrate the analytical leverage we gain by tracking legislators' expressed policy positions over time with Twitter data, here we use our method to track rhetoric by members of Congress on immigration and climate change from January 2013 to May 2020. We also display the immigration analysis on an automatically updated publicly available interactive website, enabling researchers, journalists, and policy makers alike to explore legislators' shifting rhetoric on immigration in real-time.[7]

3   Validating the Congressional Tweet Tracker

Although tweets are short 140 or 280 character messages, recent research suggests that they can be a reliable source of information about the importance that legislators place on diverse political issues, and that their validity for inferring legislators' ideology is equivalent to that of legislative speeches and roll-call votes (Barberá et al. 2018). To assess the degree to which our method is capturing substantively meaningful measures of legislators' behavior, we compared legislators' positions on

---

[7]The website is updated every Friday with a one week lag. We exclude retweets from the data displayed on the website to make the automatic updating more computationally tractable. The website therefore only displays patterns from legislators' original tweets, and not other content that they may share on Twitter.

immigration and climate change expressed in their tweets to their roll call voting records.

First, we scored each legislator according to how inclusive of an agenda their tweets expressed on immigration and how pro-action (i.e. for environmental regulation) of an agenda their tweets expressed on climate change. For immigration, this score was calculated as inclusive tweets/(inclusive + exclusive immigration tweets), and for climate it was calculated as pro-climate action/(pro-climate action + anti-climate action tweets). This produced scores ranging from 0 to 1 for each legislator, where lower scores represent less inclusive (relative to exclusive) tweets on immigration and less support for taking action on climate change (relative to tweets calling for no action or less regulation). We then compared these scores to legislators' issue-specific voting records.

In the case of immigration, we compared legislators' immigration inclusivity tweet scores to their scores on the National Hispanic Leadership Agenda's (NHLA) Hispanic Congressional Scorecard (HCS). The NHLA builds campaigns around central priorities such as Latinos United for Immigration Reform. They score members of Congress according to how frequently their votes align with positions supported by the NHLA. Their most recent scorecard assessed legislators' voting records across the 113th Congress (From January 2013-January 2015). Panel a of Figure 2 shows the correlation between legislators' immigration tweet scores and their HCS scores. Higher scores suggest more inclusive tweets and vote scores in line with NHLA positions.[8]

---

[8]Limiting the analysis to the 18 legislators with HCS scores between .25 and .75, we still observe a positive correlation with inclusively tweet scores of .43, p<.06.

10

Similarly, for climate change, we compared legislators' tweet scores—the degree to which they express support for taking action on climate change or environmental regulation—to their scores on the National Environmental Scorecard produced annually by the League of Conservation Voters (LCV). Panel b of Figure 2 shows the correlation between legislators' climate tweet scores and their LCV scores. Higher scores suggest legislators' tweets express more pro-action or pro-regulation positions on climate change and vote scores in line with LCV positions.[9]

Figure 2: **Tweet Scores vs. Voting Scorecards** Figure shows correlation between legislators' inclusivity tweet scores on immigration and their vote scorecards by the National Hispanic Leadership Council (Panel a), as well as the correlation between their action tweet scores on climate change and their vote scorecards by the League of Conservation Voters (Panel b). Correlations shown for legislators with at least 10 relevant tweets.

(a) Inclusivity Tweet Score vs. HCS Score   (b) Climate Tweet Score vs. LCV Score

$r = .89, p < .001$ $r = .77, p < .001$

---
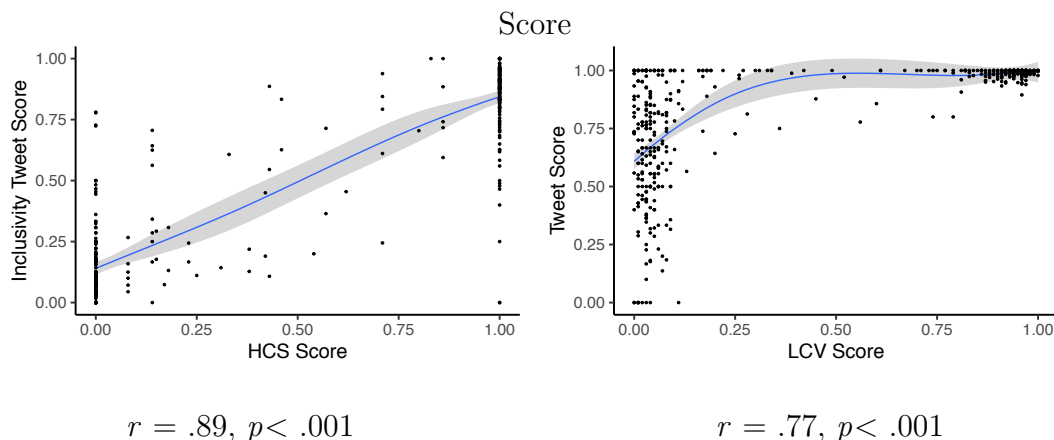
[9]Limiting the analysis to the 22 legislators with LCV scores between .25 and .75, we no longer observe a clear correlation with tweet scores, given that there is so little variation in tweet scores in that range (-.004, p<.983).

As Figure 2 demonstrates, legislators' tweets cores correlate with their vote scores on both immigration and climate change. It is possible, however, that tweet scores are less accurate for legislators who only rarely tweet about either immigration or climate change. We therefore replicate our analysis at different thresholds (legislators who have at least 5 tweets relevant to immigration or climate, 10 relevant tweets, 20 relevant tweets, or 50 relevant tweets. For immigration the correlations are .89 for legislators with at least 5 tweets, .90 for those with at least 10, .91 for those with at least 20, and .96 for those with at least 50. For climate, the correlation coefficients are .75 for those with at least 5 climate tweets, .77 for those with at least 10, .78 for those with at least 20, and .79 for those with at least 50. In both cases, correlations become stronger when the analysis is limited to legislators who tweet about these topics more often, particularly for the immigration tweets. Conducting the same analysis using DW-nominate first dimension scores—rather than issue-specific vote scores—yields similar results, with Pearson's correlation coefficients of -.85 in the case of immigration and -.75 for climate change, for legislators with at least 10 tweets. Correlations between tweet scores and DW-nominate are negative because negative nominate scores represent more liberal ideological positions.

These correlations between legislators' voting behavior and the policy positions expressed in their tweets increase our confidence that our method is identifying substantively meaningful behavior. This is further demonstrated by our analysis of the tweet scores of individual legislators. For example, the top three legislators with the most inclusive immigration tweets were Pramila Jayapal, Adriano Esaillat, and Richard Durban. Pramila Jayapal served as a pro-immigrant civil rights activist

12

before her election to Congress and has frequently called for immigration reform and opposed Trump's zero tolerance policies.[10] Rep. Esaillat was the first formerly-undocumented immigrant to be elected to Congress and has put immigration at the forefront of his agenda.[11] Richard Durban has frequently supported pro-immigrant legislation and worked to protect immigrant children in particualar.[12]

By contrast, the top legislators with the most exclusive tweets on immigration were Paul Gosar, Andy Biggs, Chip Roy, and Ted Cruz. Rep Gosar is a Republican from Arizona known for his hardline opposition to immigration reform.[13] Rep. Biggs is another Republican from Arizona known for his anti-immigrant views.[14] These patterns persist going further down the list of the legislators expressing the most inclusive and exclusive positions on immigration in their tweets, which are displayed in Tables A17 and A18 in the supplementary materials.

Turning to climate, the legislators with the highest climate tweet scores, or those calling for the most action on climate change or environmental regulation were Shel-

---

[10]https://www.nybooks.com/daily/2018/12/03/a-new-moral-imagination-on-immigration/

[11]https://www.vice.com/en/article/9kdzmp/how-two-formerly-undocumented-immigrants-got-elected-to-congress

[12]https://www.durbin.senate.gov/newsroom/press-releases/durbin-attempts-to-pass-legislation-to-protect-immigrant-workers-and-children-in-the-green-card-backlog

[13]https://www.azcentral.com/story/news/politics/immigration/2018/06/07/rep-paul-gosar-calls-immigration-reform-crap-sandwich/679047002/

[14]https://biggs.house.gov/media/press-releases/congressman-biggs-votes-against-attempt-grant-amnesty-illegal-aliens Chip Roy is a Congressman from Texas who has taken hardline positions on immigration, including arguing that Democrats' claims that kids are kept in cages are an attempt to "score political points."[15] Additionally, Senator Cruz has adopted a strong stance on immigration, calling for the end of DACA as a prerequisite for any immigration reform legislation.[16] Rep. Byrne has sponsored bills to ban the admission Syrian refugees, supported President Donald Trump's 2017 executive order travel ban, and has been very supportive of Trump's border wall.[17]

don Whitehouse, Edward Markey, and Donald Breyer. Senator Whitehouse is known for giving hundreds of speeches calling for action on climate and lambasting the fossil fuel industry.[18] Rep. Breyer is a co-chair of the Climate Change Task Force and a vocal advocate of environmental action.[19] Ed Markey has recently cosponsored the Green New Deal and has a long record of fighting for action on climate change.[20]

Legislators with the lowest tweet scores on climate action include Mitch Mc-Connell, Harold Rodgers, and Keith Rothfus, a group that has vocally opposed environmental regulation. For example, Senator McConnell has vocally opposed environmental regulation and argued that the Green New Deal would "[kill] off entire domestic industries [and wind] down millions of jobs—basically [outlaw] the only sources of energy that working-class and middle-class families can actually afford."[21] As with immigration tweet scores, these patterns persist going further down the list of the legislators expressing the most pro and anti-action or regulation climate tweets, which are displayed in Tables A21 and A22 in the supplementary materials.

Beyond validating our method, examining the top legislators whose tweet scores are most different from their vote scores highlights how members of Congress can use Twitter strategically to advocate positions that are not reflected in their voting behavior. Looking at the top legislators whose tweets on immigration were furthest from their vote scores, at the top of the list we see Jon Tester, a Democrat from Mon-

---

[18]https://www.whitehouse.senate.gov/news/release/whitehouse-delivers-200th-time-to-wake-up-climate-speech

[19]https://beyer.house.gov/news/documentsingle.aspx?DocumentID=1210

[20]https://www.newyorker.com/news/daily-comment/why-ed-markey-the-co-sponsor-of-the-green-new-deal-may-be-hopeful-for-its-chances

[21]https://twitter.com/senatemajldr/status/1110557585843908608

tana who has often been criticized for his anti-immigrant stances by progressives.[22] We also see Republicans who have low vote scores from the HCS but discuss the benefit of "high skill" migrants on Twitter such as Jerry Moran. The list of those with a large gap between their HCS vote and inclusivity tweet scores also includes Justin Amash, a Republican whose father came to the US from Palestine as a refugee[23] who regularly tweets about his origin story and the need to protect refugees more broadly.

Turning to legislators' climate tweets, the top legislators whose tweet scores are most different from their vote scorecard from the League of Conservation Voters were all Republicans who tweet about either economic benefits of climate change solutions or their bipartisan work to combat climate change. For example, Republicans Dan Newhouse and Cathy McMorris Rodgers, two of the legislators with the biggest score gap, tweet frequently about the benefits of hydro-power and clean energy in Washington State. Don Bacon, also at the top of the list, tweets about Nebraska's commitment to the environment and clean energy, despite the fact that he opposed the Paris Climate accords and votes against regulation.[24] Similarly, Rep. John Curtis has repeatedly called on the GOP to stop referring to climate change as a hoax, but nonetheless frequently votes against environmental regulation.[25] Indeed many Republicans who have low scores from the League of Conservation voters have

---

[22]https://bordercrossinglaw.com/news/progressives-must-hold-jon-tester-accountable-for-his-dream-act-vote

[23]https://www.washingtonpost.com/powerpost/arab-american-republican-lawmakers-divided-on-trumps-travel-ban/2017/02/01/d75e7d10-e8cf-11e6-80c2-30e57e57e05d_story.html

[24]https://bacon.house.gov/issues/issue/?IssueID=14894

[25]https://www.sltrib.com/news/politics/2020/08/24/rep-john-curtis-says-gop/

increasingly discussed the need for climate change solutions, with a growing number acknowledging that human activity is the primary cause of climate change[26] and calling for bipartisan solutions.[27] This trend may help explain the somewhat lower correlation of vote scores to tweets on climate change, relative to immigration. More generally, the incommensurate vote and tweet scores that our method quantifies opens up new avenues for the study of strategic communication and its relationship to behavior.[28]

4   Mapping Issue Salience and Policy Positions Over Time

Like past work using floor speeches, press releases, and legislators' tweets, our tracker enables us to evaluate how the salience of these issues changes over time. Figure 3 shows the daily proportion of tweets about immigration (Panel a) and climate change/environmental regulation (Panel b) in Members of Congress' tweets from January 2013 to May 2020. Days with the highest volume of tweets are annotated. The bottom of each panel shows the figures broken down by party with Democrats in blue and Republicans in red. The plot suggests that legislators tweet about immigration and climate change quite frequently, and on days when key legislation is announced or external events make these issues salient, up to almost 40% of legislators'

---

[26]https://www.washingtonpost.com/politics/2019/04/03/has-green-new-deal-pushed-republicans-acknowledge-that-man-made-climate-change-is-real/

[27]https://www.theatlantic.com/politics/archive/2017/03/house-republicans-climate-change-global-warming-trump/518430/

[28]A list of the top 10 legislators with the largest gap between tweet and vote scores on both immigration and climate is displayed in Table A22 in the supplementary materials.

daily tweets reference immigration and almost 20% reference climate change. While immigration is approximately equally salient for both Republicans and Democrats, climate change is much more salient for Democrats—particularly following President Trump's political rise and subsequent election in 2016.

Figure 3: **Salience of Immigration and Climate in Members of Congress' Tweets**. Figure shows how the salience of immigration (Panel a) and climate (Panel b) shifts over time in legislators' tweets. The data are all tweets produced by Members of Congress collected with Twitter's Historical Streaming API from January 2013 to May 2020. Tweets have been classified as being relevant to immigration or climate change using word2vec dictionaries trained on the full dataset of legislators' tweets. Days with the highest proportion of relevant tweets are annotated. Top panel shows results for all Members of Congress, bottom panel shows results disaggregated by party.

(a) Immigration
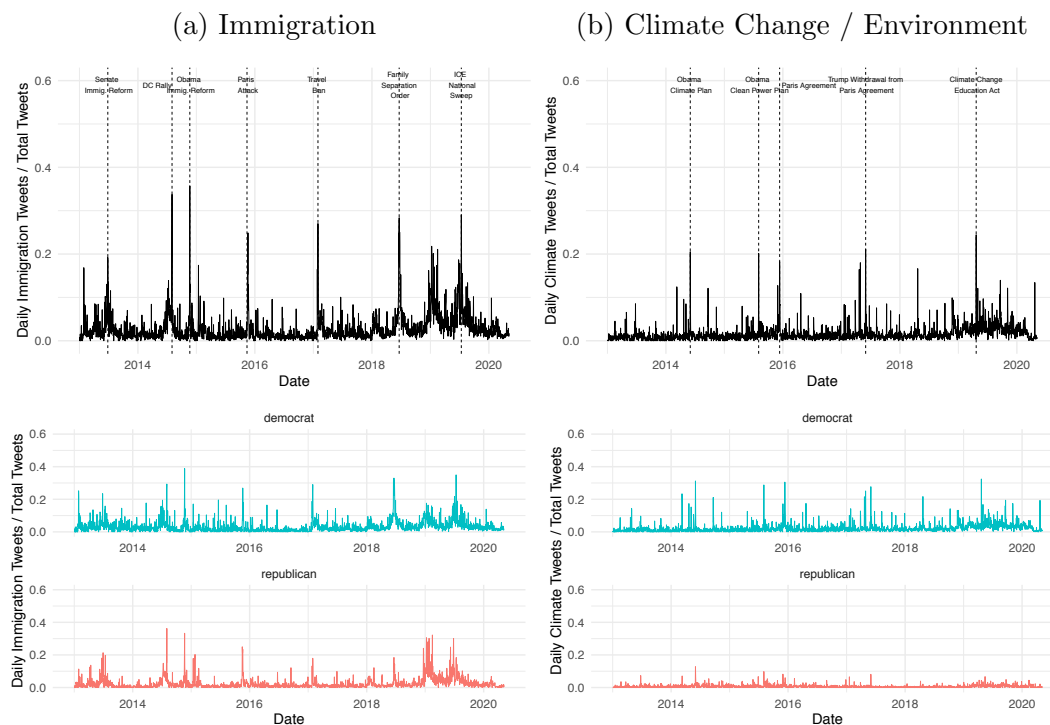
(b) Climate Change / Environment

Figure 4 shows changing issue positions in Members of Congress' tweets about immigration (Panel a) and climate change / environmental regulation (Panel b) over time for the same period in aggregate and again disaggregated by party. With regard to immigration (Panel a), while the proportion of tweets expressing inclusive policy positions towards refugees and migrants increases over the entire time period, the proportion of tweets expressing exclusive policy positions increases until early 2016 and then levels off. In aggregate we see legislators' tweets becoming more inclusive than exclusive after 2016.

Notably, until 2016 fewer than 20% of legislators' tweets actually take an inclusive or exclusive position on immigration, whereas after 2016 about half of legislators' tweets express either an inclusive or exclusive position. Indeed many of legislators' tweets simply reference immigration as a topic discussed at a town hall event or media appearance or share news stories about immigration without expressing clear policy positions. This strategy aligns with past research suggesting that members of Congress tend to make ambiguous rather than clear statements on divisive issues for fear that speaking too boldly on an issue could have negative repercussions if an issue stance proves to be unpopular (Alesina and Cukierman 1990; Milita et al. 2017).

With regard to those tweets taking a stand, most inclusive tweets are produced by Democrats while most exclusive tweets are produced by Republicans. Broken down by party, among legislators who tweet about immigration, we observe Democrats producing an increasingly large proportion of tweets expressing inclusive positions towards refugees and migrants over the entire period. The proportion of exclusive

18

tweets produced by Republicans increases over this time period also increases over this time period, particularly after 2018.

Figure 4: **Issue Positions on Immigration and Climate Change/ Environmental Regulation in Members of Congress' Tweets**. Figure plots daily proportion of legislators' relevant tweets expressing issue positions on immigration (Panel a) and climate change/environmental regulation (Panel b) as local regression lines with loess smoothing and 95% confidence intervals. The data are all tweets produced by Members of Congress collected with Twitter's Historical Streaming API from January 2013 to May 2020 that are relevant to immigration (Panel a) or climate change / environmental regulation (Panel b). Tweets have been classified as being relevant to immigration or climate change using word2vec dictionaries trained on the full dataset of legislators' tweets. Relevant tweets are then classified as inclusive of immigrants/refugees, exclusive of immigrants/refugees, or neither, in Panel a and supporting taking action to address climate change / supporting environmental regulation, opposing taking action to address climate change / opposing environmental regulation, or neither, in Panel b. The data is disaggregated by party in the lower half of the Figure.

(a) Immigration

(b) Climate Change / Environment
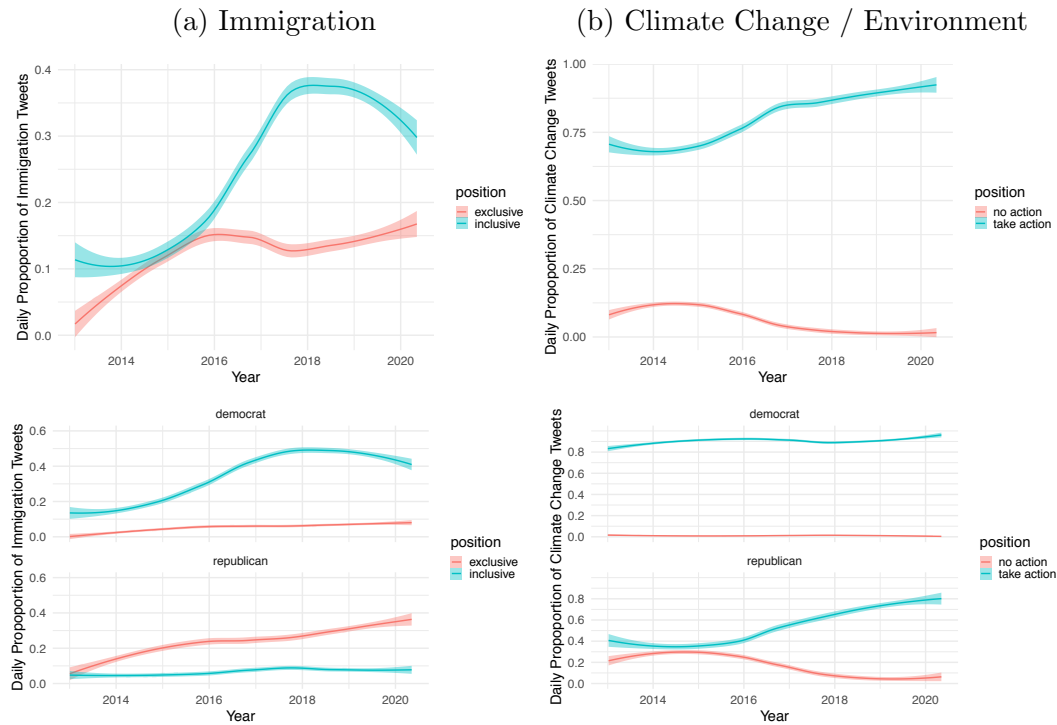
Structural topic models (Roberts et al. 2013) of legislators' tweets (presented in Panel a of Figure 5) highlight the issues encompassed in the inclusive vs. exclusive immigration debate.[29] Inclusive topics, largely emphasized in Democrats' tweets, include opposing Trump's travel ban, the need to reunite families, characterizing refugees as fleeing ISIS or other harms, and celebrating immigrants as beneficial to the economy. By contrast, exclusive topics, primarily referenced in Republicans' tweets, reference crime, the southern border, and Customs and Border Patrol.

Turning to our climate analysis, as Panel b of Figure 4 demonstrates, in aggregate legislators are increasingly calling for action on climate change or environmental regulation, while the proportion of tweets calling for no action on climate change or opposing environmental regulation is decreasing in this period. Unsurprisingly, the vast majority of tweets produced by Democrats are calling to take more action on climate change or advocating for more environmental regulation. Perhaps more surprisingly, Republicans increasingly call for more action on climate change over this period while at the same time reducing the proportion of their tweets that oppose action on climate change or more environmental regulation. This is in line with data from the League of Conservation Voters suggesting that while Republicans' environmental voting records consistently get lower scores than those of Democrats, a plurality of Republicans on the Climate Solutions Caucus have increased their scores over the past several years.[30]

---

[29]The number of topics (k) was set to 25 to run these structural topic models. After assigning a tentative label to each topic, we finalized the labels by reviewing the 10 tweets with the highest probability for each topic and evaluating whether the content of these messages conforms to the labeled theme inferred from the cluster of top keywords.

[30]https://citizensclimatelobby.org/exploring-nuances-lcv-scorecard-climate-solutions-caucus/

Structural topic models of legislators' tweets (presented in Panel b of Figure 5) suggest that most of Republicans' tweets opposing action were criticizing the EPA for overreach under the Obama administration, criticizing Obama's so-called "War on Coal," and criticizing the Paris Climate agreement. Their tweets calling for action or regulation generally are highlighting how green energy will likely create new jobs in their districts, and these tweets have increased over time. By contrast, Democrats' tweets calling for action on climate change or regulation tend to call for people to recognize climate change as a global problem, advocate US participation in international treaties, advocate taking scientific research on climate change seriously, and criticize climate denial.

Figure 5: **Topics by Issue Positions on Immigration and Climate Change/ Environmental Regulation in Members of Congress' Tweets**. Figure shows the results of a structural topic model comparing the topics prevalent in "inclusive" and "exclusive" tweets on immigration (Panel a) and tweets that call for action on climate change or environmental regulation compared to those that call for no action or less regulation (Panel b).

(a) Immigration
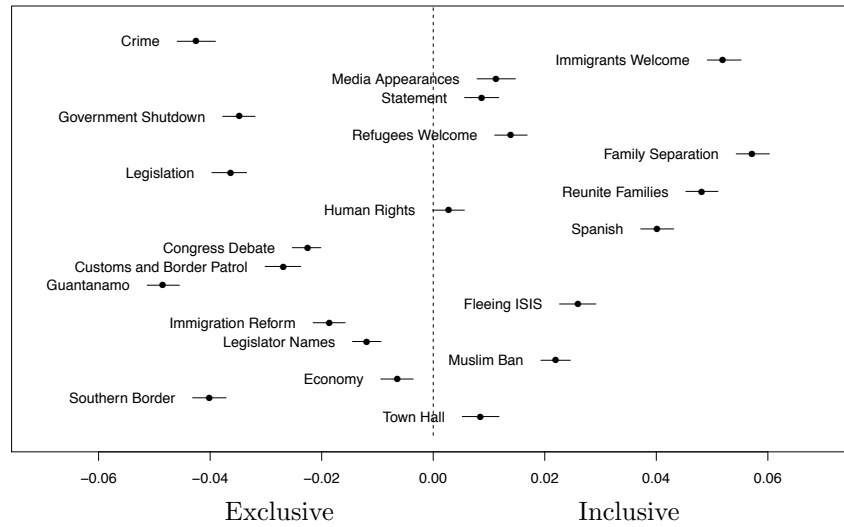
(b) Climate Change / Environmental Regulation
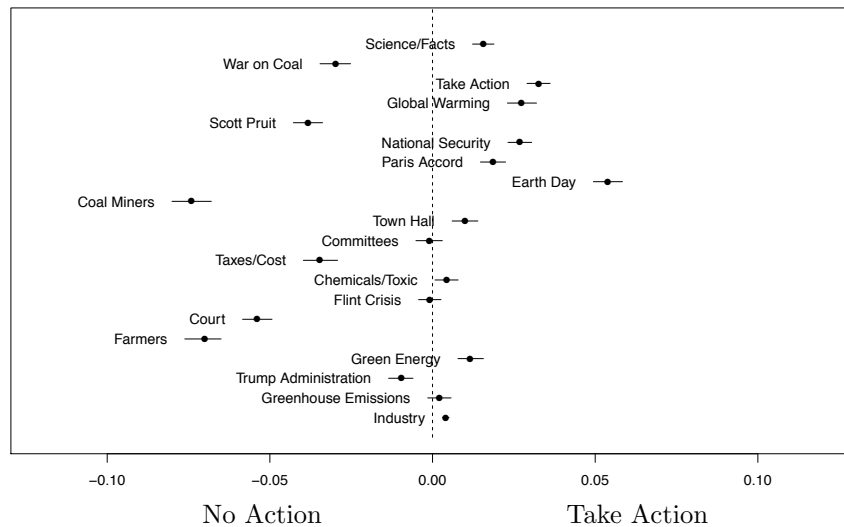
Subsetting our data to members of Congress who tweet across the entire time period, we see that our results are not driven by a change in adoption of Twitter by members of Congress. All senators and the vast majority of members of the House were on Twitter by 2013 and we therefore see that the majority of legislators in our dataset tweet across the entire time period, excluding those who left office or were elected for the first time. As Figure A.1 in the supplementary materials suggests, our results look quite similar excluding those who do not tweet across the entire time period.

5   Discussion and Conclusion

As these examples demonstrate, tracking legislators' expressed policy positions through Twitter data offers detailed, real-time, scalable insights into how elite rhetoric shifts over time. Our approach can be scaled to other legislatures and used to answer longstanding social science questions about elite political behavior and public policy, facilitating tests of hypotheses that may otherwise be constrained by data availability or timeliness. For example, how do issue salience and issue positions shift in the lead up or aftermath of elections? How does issue ownership by parties or factions of parties change over time? How do exogenous events (e.g. terror attacks or environmental disasters) affect the manner in which elites reveal their positions on policy issues? When is elite rhetoric cheap talk and when does it predict voting behavior? When do legislators diverge from their party on salient policy issues? Answering questions that rely on data from diverse groups of legislators with high levels of tem-

poral granularity is much more difficult—and in some cases impossible—using more traditional data sources.

Like any measurement approach, tracking legislators' expressed policy positions with Twitter data has limitations. Because tweets are very short texts, they do not offer the same level of nuanced rhetoric as other forms of strategic communication. However, given that legislators put a great deal of time and resources into managing their social media accounts, their tweets tend to be longer, contain fewer spelling mistakes, and more formal grammar and sentence structure than posts produced by everyday citizens on social media, making content analysis more effective. As a result, recent research has found analysis of legislators' tweets produces valid measures of ideology and issue positions on par with analysis using floor speeches (Barberá et al. 2018).

Second, although collection and analysis of data for the real-time tracker can be automated and scaled limitlessly, the word2vec dictionaries and classifiers do need to be updated periodically in order to ensure that we are not missing new rhetoric that may have entered a debate surrounding a particular issue. If dictionaries and classifiers have not been updated recently, we may miss key events or short-term trends. Similarly, the human labeling in the training data needs to be sufficiently accurate to allow for reliable classification.

Finally, not all legislators use Twitter at the same rate. While this is a concern in measuring any form of strategic communication—as legislators do not all put out press releases or give floor speeches (for example) at the same rate on the same topics—the differences between legislators may be more pronounced on social

media. That being said, any particular analysis can account for which legislators are producing which content as Twitter data is pre-annotated with rich metadata that can be combined with other legislator-specific variables.

Leveraging Twitter data, our real-time policy agenda tracker enables us to measure when and how political actors communicate about salient policy issues with a high level of temporal granularity. As our proof of concept examples mapping members of Congress' daily communication on immigration and climate change from 2013 to 2020 suggest, even simple descriptive plots of this data offer insights into how elite strategic communication evolves over time. The method is transparant and interpretable, achieves satisfactory levels of accuracy, and is scalable over different time periods and issue areas. Moreover, by displaying this data on an automatically updated publicly available interactive website, our tracker enables not only researchers, but also journalists, policy makers, and everyday citizens to gain detailed insight into how elites talk about salient policy issues. We hope that future research will take advantage of this approach to improve our understanding of the microdynamics of political communication on a wide range of topics in diverse comparative contexts.

References

Achen, Christopher H and Larry M Bartels. 2006. It feels like we're thinking: The rationalizing voter and electoral democracy. In Annual Meeting of the American Political Science Association, Philadelphia. Vol. 30.

Adler, E Scott, Chariti E Gent and Cary B Overmeyer. 1998. "The home style homepage: Legislator use of the World Wide Web for constituency contact." Legislative Studies Quarterly pp. 585–595.

Alesina, Alberto and Alex Cukierman. 1990. "The politics of ambiguity." The Quarterly Journal of Economics 105(4):829–850.

Anstead, Nick and Andrew Chadwick. 2008. Parties, election campaigning, and the Internet: Toward a comparative institutional approach. In Routledge handbook of internet politics. Routledge pp. 72–87.

Arora, Deepali, Kin Fun Li and Stephen W Neville. 2015. Consumers' sentiment analysis of popular phone brands and operating system preference using Twitter data: A feasibility study. In 2015 IEEE 29th International Conference on Advanced Information Networking and Applications. IEEE pp. 680–686.

Barberá, Pablo, Andreu Casas, Jonathan Nagler, Patrick Egan, Richard Bonneau, John T Jost and Joshua A Tucker. 2018. "Who Leads? Who Follows? Measuring Issue Attention and Agenda Setting by Legislators and the Mass Public Using Social Media Data.".

Bartels, Larry M. 2005. "Homer gets a tax cut: Inequality and public policy in the American mind." Perspectives on Politics 3(1):15–31.

Bartlett, Jamie and Richard Norrie. 2015. "Immigration on Twitter: understanding public attitudes online." Demos .

Broockman, David E and Daniel M Butler. 2017. "The Causal Effects of Elite Position-Taking on Voter Attitudes: Field Experiments with Elite Communication." American Journal of Political Science 61(1):208–221.

Catalinac, Amy. 2018. From pork to policy: The rise of programmatic campaigning in Japanese elections. In Critical Readings on the Liberal Democratic Party in Japan. BRILL pp. 882–917.

Chong, Dennis and James N Druckman. 2007a. "Framing public opinion in competitive democracies." American Political Science Review 101(4):637–655.

Chong, Dennis and James N Druckman. 2007b. "A theory of framing and opinion formation in competitive elite environments." Journal of Communication 57(1):99–118.

Cluverius, John G. 2012. Legislator technology adoption and presentation of self through Twitter PhD thesis The University of North Carolina at Chapel Hill.

Cormack, Lindsey. 2013. Sins of Omission: Legislator (Mis) Representation in Constituent Communications. In APSA Annual Meeting.

Downs, Anthony et al. 1957. "An economic theory of democracy.".

Druckman, James N, Martin J Kifer and Michael Parkin. 2009. "Campaign communications in US congressional elections." American Political Science Review 103(3):343–366.

Evans, Heather K, Victoria Cordova and Savannah Sipole. 2014. "Twitter style: An analysis of how house candidates used Twitter in their 2012 campaigns." PS: Political Science & Politics 47(2):454–462.

Farrell, Joseph. 1995. "Talk is cheap." The American Economic Review 85(2):186–190.

Fenno, Richard F. 1978. Home style: House members in their districts. Harper-Collins,.

Flores, René D. 2017. "Do anti-immigrant laws shape public sentiment? A study of Arizona's SB 1070 using Twitter data." American journal of sociology 123(2):333–384.

Franklin, Charles H. 1991. "Eschewing obfuscation? Campaigns and the perception of US Senate incumbents." American Political Science Review 85(4):1193–1214.

Golbeck, Jennifer, Brooke Auxier, Abigail Bickford, Lautaro Cabrera, Meaghan Conte McHugh, Stephani Moore, Jacquelyn Hart, Justin Resti, Anthony Rogers and Jenna Zimmerman. 2018. "Congressional twitter use revisited on the platform's 10-year anniversary." Journal of the Association for Information Science and Technology 69(8):1067–1070.

Golbeck, Jennifer, Justin M Grimes and Anthony Rogers. 2010. "Twitter use by the US Congress." Journal of the American Society for Information Science and Technology 61(8):1612–1621.

Graham, Todd, Marcel Broersma, Karin Hazelhoff and Guido Van'T Haar. 2013. "Between broadcasting political messages and interacting with voters: The use of Twitter during the 2010 UK general election campaign." Information, communication & society 16(5):692–716.

Grimmer, Justin. 2010. "A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases." Political Analysis 18(1):1–35.

Grimmer, Justin. 2013a. "Appropriators not position takers: The distorting effects of electoral incentives on congressional representation." American Journal of Political Science 57(3):624–642.

Grimmer, Justin. 2013b. Representational style in Congress: What legislators say and why it matters. Cambridge University Press.

Grimmer, Justin. 2016. "Measuring Representational Style in the House: The Tea Party, Obama, and Legislators' Changing Expressed Priorities.".

Grimmer, Justin, Sean J Westwood and Solomon Messing. 2014. The impression of influence: legislator communication, representation, and democratic accountability. Princeton University Press.

Grose, Christian R, Neil Malhotra and Robert Parks Van Houweling. 2015. "Explaining explanations: How legislators explain their policy positions and how citizens react." American Journal of Political Science 59(3):724–743.

Grover, Ted, Elvan Bayraktaroglu, Gloria Mark and Eugenia Ha Rim Rho. 2019. "Moral and affective differences in us immigration policy debate on twitter." Computer Supported Cooperative Work (CSCW) 28(3-4):317–355.

Guidry, Jeanine PD, Lucinda L Austin, Kellie E Carlyle, Karen Freberg, Michael Cacciatore, Yan Jin and Marcus Messner. 2018. "Welcome or Not: Comparing# Refugee Posts on Instagram and Pinterest." American Behavioral Scientist 62(4):512–531.

Gulati, Jeff and Christine B Williams. 2010. "Communicating with constituents in 140 characters or less: Twitter and the diffusion of technology innovation in the United States Congress.".

Gurciullo, Stefano and Slava J Mikhaylov. 2017. Detecting policy preferences and dynamics in the un general debate with neural word embeddings. In 2017 International Conference on the Frontiers and Advances in Data Science (FADS). IEEE pp. 74–79.

Hemphill, Libby, Jahna Otterbacher and Matthew Shapiro. 2013. What's congress doing on twitter? In Proceedings of the 2013 conference on Computer supported cooperative work. ACM pp. 877–886.

Hollander, Justin B and Henry Renski. 2015. "Measuring urban attitudes using Twitter: An exploratory study." Cambridge: Lincoln Institute of Land Policy .

Hornikx, Jos and Berna Hendriks. 2015. "Consumer tweets about brands: A content analysis of sentiment tweets about goods and services." Journal of Creative Communications 10(2):176–185.

Jacoby, William G. 2000. "Issue framing and public opinion on government spending." American Journal of Political Science pp. 750–767.

Jashinsky, Jared, Scott H Burton, Carl L Hanson, Josh West, Christophe Giraud-Carrier, Michael D Barnes and Trenton Argyle. 2014. "Tracking suicide risk factors through Twitter in the US." Crisis .

Jerit, Jennifer. 2006. "Reform, rescue, or run out of money? Problem definition in the Social Security reform debate." Harvard International Journal of Press/Politics 11(1):9–28.

Kat Devlin, Regina Widjaya and Jeremiah Cha. 2019. "For Global Legislators on Twitter, an Engaged Minority Creates Outsize Share of Content." Pew Research .

Kinder, Donald R and Lynn M Sanders. 1990. "Mimicking political debate with survey questions: The case of white opinion on affirmative action for blacks." Social cognition 8(1):73–103.

Klüver, Heike and Iñaki Sagarzazu. 2016. "Setting the agenda or responding to voters? Political parties, voters and issue attention." West European Politics 39(2):380–398.

Kousser, Thad. 2019. "Tweet style: campaigning, governing, and social media in Australia." Australian Journal of Political Science pp. 1–19.

Lau, Richard R, Lee Sigelman and Ivy Brown Rovner. 2007. "The effects of negative political campaigns: a meta-analytic reassessment." Journal of Politics 69(4):1176–1209.

Leeper, Thomas J. 2013. Interest Groups and Political Attitudes. Technical report Working Paper.

Lenz, Gabriel S. 2009. "Learning and opinion change, not priming: Reconsidering the priming hypothesis." American Journal of Political Science 53(4):821–837.

Lilleker, Darren G and Karolina Koc-Michalska. 2013. "Online political communication strategies: MEPs, e-representation, and self-representation." Journal of Information Technology & Politics 10(2):190–207.

Loewen, Peter John and Daniel Rubenson. 2011. "For want of a nail: Negative persuasion in a party leadership race." Party Politics 17(1):45–65.

Mackie, Diane and Joel Cooper. 1984. "Attitude polarization: Effects of group membership." Journal of Personality and Social Psychology 46(3):575.

Manheim, Jarol B. 1994. Strategic Public Diplomacy and American Foreign Policy: The Evolution of Influence lit. Oxford University Press.

Martin, Lanny W and Georg Vanberg. 2008. "Coalition government and political communication." Political Research Quarterly 61(3):502–516.

Martin, Shane. 2011. "Parliamentary questions, the behaviour of legislators, and the function of legislatures: An introduction." The Journal of Legislative Studies 17(3):259–270.

Mayhew, David R. 1974. Congress: The electoral connection. Vol. 26 Yale University Press.

McLean, Lucy and Madison Bartula-Henkle. 2019. "The Weinstein Effect: The Role of MeToo Movement in Altering the Communication Strategies of Members of Congress.".

Mikolov, Tomas, Kai Chen, Greg Corrado, Jeffrey Dean, L Sutskever and G Zweig. 2013. "word2vec." URL https://code. google. com/p/word2vec .

Milita, Kerri, Elizabeth N Simas, John Barry Ryan and Yanna Krupnikov. 2017. "The effects of ambiguous rhetoric in congressional elections." Electoral Studies 46:48–63.

Mitchell, Lewis, Morgan R Frank, Kameron Decker Harris, Peter Sheridan Dodds and Christopher M Danforth. 2013. "The geography of happiness: Connecting twitter sentiment and expression, demographics, and objective characteristics of place." PloS one 8(5):e64417.

Munger, Kevin, Richard Bonneau, Jonathan Nagler and Joshua A Tucker. 2018. "Elites tweet to get feet off the streets: Measuring regime social media strategies during protest." Political Science Research and Methods pp. 1–20.

Paulussen, Steve and Raymond A Harder. 2014. "Social media references in newspapers: Facebook, Twitter and YouTube as sources in newspaper journalism." Journalism practice 8(5):542–551.

Quinn, Kevin M, Burt L Monroe, Michael Colaresi, Michael H Crespin and Dragomir R Radev. 2010. "How to analyze political attention with minimal assumptions and costs." American Journal of Political Science 54(1):209–228.

R Core Team. 2021. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
URL: *https://www.R-project.org/*

Radford, Jason and Betsy Sinclair. 2016. "Electronic Homestyle: Tweeting Ideology.".

Riker, William H. 1986. The art of political manipulation. Vol. 587 Yale University Press.

Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Edoardo M Airoldi et al. 2013. The structural topic model and applied social science. In Advances in neural information processing systems workshop on topic models: computation, application, and evaluation. Vol. 4 Harrahs and Harveys, Lake Tahoe.

Rojecki, Andrew. 2008. "Rhetorical alchemy: American exceptionalism and the war on terror." Political Communication 25(1):67–88.

Scherpereel, John A, Jerry Wohlgemuth and Audrey Lievens. 2018. "Does Institutional Setting Affect Legislators' Use of Twitter?" Policy & Internet 10(1):43–60.

Sellers, Patrick. 2009. Cycles of spin: Strategic communication in the US Congress. Cambridge University Press.

Shapiro, Matthew A, Libby Hemphill, Jahna Otterbacher and Han Woo Park. 2014. "Twitter and Political Communication in Korea: Are Members of the Assembly Doing What They Say?" Journal of Asia Pacific Studies 3(3).

Tesler, Michael. 2015. "Priming predispositions and changing policy positions: An account of when mass opinion is primed or changed." American Journal of Political Science 59(4):806–824.

Tumasjan, Andranik, Timm O Sprenger, Philipp G Sandner and Isabell M Welpe. 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. In Fourth international AAAI conference on weblogs and social media.

Vavreck, Lynn. 2009. The message matters: The economy and presidential campaigns. Princeton University Press.

Williams, Matthew L and Pete Burnap. 2015. "Cyberhate on social media in the aftermath of Woolwich: A case study in computational criminology and big data." British Journal of Criminology 56(2):211–238.

Wilson, Walter. 2009. "Latino representation on congressional websites." Legislative Studies Quarterly 34(3):427–448.

Yoon, Ho Young and Han Woo Park. 2014. "Strategies affecting Twitter-based networking pattern of South Korean politicians: Social network analysis and exponential random graph model." Quality & Quantity 48(1):409–423.

Yu, Yang and Xiao Wang. 2015. "World Cup 2014 in the Twitter World: A big data analysis of sentiments in US sports fans' tweets." Computers in Human Behavior 48:392–400.

Zaller, John R et al. 1992. The nature and origins of mass opinion. Cambridge university press.

A   Supplementary Materials

A.1   Coding Scheme

The following coding schemes were presented to Figure8 workers who did the human coding of tweets we used to train our classifiers for immigration and climate change / environmental regulation. The immigration coding instructions initially contained two categories (characterizing refugees/immigrants as a contribution, burden, or neither and promoting policies that are inclusive or exclusive toward refugees/migrants). Because contribution/burden tweets were quite rare, we collapsed these categories into: contribution or inclusive and burden or exclusive, which form the broader inclusive and exclusive policy dimensions we reference in the main body of the text.

Immigration

Overview: In this job, you will be presented with tweets and other online posts from politicians that are about refugees and immigration. Review the tweets to determine the relevance to the subject and then answer a series of questions about each tweet.

Coding Instructions

- Read the tweet.

- Determine if the tweet is relevant to the topic.

- If tweet is relevant, answer a series of questions about the content of the tweet.

A1

Question 1: If the tweet is relevant, you will first assess whether the tweet characterizes refugees as contributing to or placing a burden on society by answering the following question:

Does the following tweet characterize refugees or immigrants as contributing to society or placing a burden on society?

- Contribution

- Burden

- Neither / Irrelevant

Contribution: Messages stating that refugees or immigrants contribute to society include those that characterize them as offering cultural or economic benefits such as enriching local culture or bringing new businesses or skills to the economy. They also include tweets pointing out that refugees are heavily vetted and do not represent a danger to society, are not more likely to commit crimes, terror attacks, or otherwise cause harm.

Burden: Messages stating that refugees or immigrants place a burden on society include those that characterize them as damaging the host society's culture or values, hurting the economy, taxpayers, or welfare system, or increasing crime. They also include messages blaming refugees for terror attacks or extremism, describing them as representing a danger or as "troublemakers," "problematic," or otherwise causing harm.

Neither: Messages that neither characterize refugees as contributing to or placing a burden on a host society include news stories or neutral announcements about immigrants or refugees, messages that do not state a position on the impact of refugees on a host society.

Contribution Examples:

- Immigrants make the US stronger. #insideout11M http://t.co/Ne90SnuTYT

- Rescinding DACA would remove 750,000 productive #Dreamers from the economy. Protect the program. #HeretoStay

- We shouldn't be educating drs in the US forcing them 2 leave when we need drs in rural areas. My bipartisan bill extends visas 4 these drs

- Donald Trump's claims about immigrants committing more crimes? You guessed it, they're false

- Immigrant families should not be forced to wait in the shadows. #Fight4DAPA Good for families, good for the economy, good for our country!

Burden Examples:

- Pretty obvious, it's cheaper to keep migrants out than letting them in-ask anyone hurt by illegals committing crimes https://t.co/KrkO7lr6hE

- Introduced bill w senators including @SenFeinstein amp; @JeffFlake to strengthen Visa Waiver Program amp; help stop terrorists from entering U.S.

- #Amnesty for 11 million illegal immigrants costs hard-working American #taxpayers millions of dollars American workers thousands of jobs.

- I'm cosponsoring legislation to stop the influx of refugees since Obama cannot be trusted to verify those entering the US are not terrorists

- Amnesty hurts taxpaying Americans who want good jobs

Neither Examples:

- Thanks @Neighbster 4 having me at #wellstone center in StPaul. My amndts 2 immigration bill would protect more women from domestic violence

- On this date in 1868, 14th Amendment was ratified, granting citizenship to all people born or naturalized in U.S., including former slaves.

- ICYMI, I was on @upwithsteve yesterday talking about Syrian refugee crisis: http://t.co/2pooCeCUCc

Question 2: If the tweet is relevant, you will then assess whether the tweet contains an inclusive or exclusionary message about by answering the following question:

Does the following tweet contain an inclusive or exclusionary message about refugees or immigrants?

- Inclusive

- Exclusive

- Neither

Inclusive: Inclusive messages express hospitality or advocate welcoming more refugees or immigrants, extending rights to refugees or immigrants or protecting their existing rights. These include tweets calling for more inclusive policies towards refugees or immigrants such as more lenient asylum or amnesty policies or those opposing exclusive policies such as travel bans or border closures. They also include tweets expressing sympathy for the plight of refugees.

Exclusive: Exclusive messages express a desire to prevent migrants or refugees from entering or remaining in a host country or advocate restricting the rights of refugees or migrants within a host country. These include tweets calling for more exclusive policies towards refugees or immigrants such as border closures, improving border security, travel bans. They also include tweets stating that refugees should stay in their home countries etc.

Neither: Messages that are neither inclusive nor exclusionary include news stories or neutral announcements about immigrants or refugees, and messages that do not express a clear attitude or stance on accepting immigrants or refugees (or related policies).

Inclusive Examples:

- I will be joining @NancyPelosi and fellow Democrats today on steps of SCOTUS

at 6 p.m. to oppose the refugee ban. #NoBan

- Wanted to share this touching story on the life of one Syrian refugee, a 12-year-old girl named Hana:https://t.co/EvLEfgosu7

- Will #immigration reform include DREAMer's? Al thinks it will. http://t.co/4qedAuMnT1

- Thank you Sen @amyklobuchar: extends U visa protections to elderly suffering abuse. Key in country w/ increasing # of elderly.

Exclusive Examples:

- The first priority of our immigration policy should be securing the border to halt the flow of illegal immigrants. #LASenateDebate #LASEN

- 5,000 plus on tonight's Teletown. Will bring your support for #fairtax, low spending, opposition to #obamacare, amnesty to the debate.

- As Governor I will oppose Syrian refugees being relocated to Arkansas.

- Just voted to protect #NationalSecurity by strengthening #VisaWaiverProgram. https://t.co/Vm8gIOfHdM

- Voted to protect our borders. #Time4Solutions More: http://t.co/PB1yKpJGgA

Neither Examples:

- Thanks Senator @alfranken for your Leadership on #immigration and be at the town hall #timeisnow @SEIU http://t.co/tMBOkBGZa9

A6

- Climate crisis is here - 1st US climate refugees expose the global risk hundreds of millions face in coming decades https://t.co/juOu37g6aY

- Spoke to @ShopFloorNAM Board today on Competitive Agenda: Training workers, more exports, infrastructure, immigration/regulatory/tax reform.

- See my full statement on Syrian refugees below. https://t.co/Y2m7ThqCCK

Climate Change / Environmental Regulation

Overview In this job, you will be presented with tweets and other online posts from politicians that are about climate change, energy, or the environment. Review the tweets to determine the relevance to the subject and then answer a series of questions about each tweet.

Coding Instructions

- Read the tweet.

- Determine if the tweet is relevant to the topic.

- If tweet is relevant, answer a series of questions about the content of the tweet.

If the tweet is relevant, you will first assess whether the tweet is calling for more or less environmental regulation by answering the following question:

Does the following tweet call for more or less environmental regulation?

- More Environmental Regulation (Climate Change is a Problem)

- Less Environmental Regulation (Climate Change is Not a Problem)

- Neither / Irrelevant

More Environmental Regulation (Climate Change is a Problem): Messages calling for more environmental regulation characterize climate change as a problem or a threat, oppose policies that may have harmful effect on the environment, support "clean energy," are supportive of the Environmental Protection Agency (EPA) or other groups calling for more regulation, and/or directly call for more environmental regulation of coal, fossil fuels, drilling for oil, etc.. They may also call for the US to be a part of international treaties such as the Paris Climate Agreement.

Less Environmental Regulation (Climate Change is Not a Problem): Messages calling for less environmental regulation do not characterize climate change as a problem, call for less regulation of industries like coal and oil, are opposed to the Environmental Protection Agency (EPA) or other groups calling for more regulation, and/or directly call for less environmental regulation and encourage the US to leave international treaties such as the Paris Climate Agreement.

Neither: Messages that neither call for more or less environmental regulation include news stories or neutral announcements about climate, energy, or the environment, as well as messages that do not state a position on climate change, energy, or the environment.

More Environmental Regulation (Climate Change is a Problem) Examples:

- @SEEC will continue to fight Trump's assault on environmental protections

- Clean energy makes business and economic sense.

- I urge Congress to reinstate the House Select Committee on Climate Change to take immediate steps to work on this issue.

- When Trump pulled the US out of the climate agreement he announced Make America Retreat Again!

- Please don't drill for oil on the beautiful coastline of Coupeville, WA.

Less Environmental Regulation (Climate Change is Not a Problem) Examples:

- Fighting Back Against President Obama's Anti-Coal Plan

- SCOTUS agrees, EPA does not have unlimited authority to implement whatever ill-conceived regulations it chooses.

- Curb harmful #EPA #regulations

- A #WarOnCoal is a #WarOnJobs

- Obama's #WarOnCoal must no longer be one-sided. Time 2 fight back! #Coal Country Protection Act stops EPA extremism

A.2 Preprocessing Tweets

Before training our classifier, tweets were pre-processed to remove English stopwords, punctuation, and URLs. Hashtags and Twitter handles were not removed. Words occurring in fewer than 5% of tweets were trimmed from the document term matrix.

A.3 Dictionary Validation

To ensure that our word2vec dictionaries allow us to accurately tag tweets that are relevant to immigration or climate change / environmental regulation, we took two stratified random samples of tweets produced by all legislators and coded them according to whether or not they were relevant to immigration or climate change / environmental regulation. In the first sample there were 18 tweets about immigration. Our dictionary method identified 15 of the 18 tweets (83% Recall) and identified one tweet that was not actually relevant to immigration (94% Precision). In the second sample we identified 29 tweets that were relevant to climate change. Our dictionary failed to identify two of these tweets (93% Recall) and identified one tweet that was not actually relevant to climate change / environmental regulation (96% Precision).[31]

---

[31]Throughout the manuscript, precision refers to true positives / (true positives + false positives); recall refers to true positives / (true positives + false negatives); and the F1 score is a harmonic mean or weighted average of precision and recall.

A.4 Classifier Performance

Immigration

Table 1: Naive Bayes Classifier Performance (Exclusive or Not)

Accuracy	0.813
Precision	0.921
Recall	0.853
F1	0.885
Balanced Accuracy	0.738

Table 2: Naive Bayes Classifier Performance (Inclusive or Not)

Accuracy	0.830
Precision	0.920
Recall	0.865
F1	0.892
Balanced Accuracy	0.781

Table 3: XGBoost Classifier Performance (Exclusive or Not)

Accuracy	0.795
Precision	0.836
Recall	0.909
F1	0.871
Balanced Accuracy	0.668

Table 4: XGBoost Classifier Performance (Inclusive or Not)

Accuracy	0.828
Precision	0.861
Recall	0.916
F1	0.888
Balanced Accuracy	0.746

Table 5: Elastic Net Classifier Performance (Exclusive or Not)

Accuracy	0.801
Precision	0.809
Recall	0.967
F1	0.881
Balanced Accuracy	0.615

Table 6: Elastic Net Classifier Performance (Inclusive or Not)

Accuracy	0.821
Precision	0.817
Recall	0.978
F1	0.890
Balanced Accuracy	0.674

Table 7: Lasso Classifier Performance (Exclusive or Not)

Accuracy	0.784
Precision	0.797
Recall	0.962
F1	0.872
Balanced Accuracy	0.586

Table 8: Lasso Classifier Performance (Inclusive or Not)

Accuracy	0.828
Precision	0.826
Recall	0.974
F1	0.894
Balanced Accuracy	0.691

A.4.1 Climate Change / Environmental Regulation

Table 9: Naive Bayes Classifier Performance (No Action or Not)

Accuracy	0.827	0.015
Precision	0.904	0.010
Recall	0.874	0.019
F1	0.889	0.011
Balanced Accuracy	0.742	0.019

Table 10: Naive Bayes Classifier Performance (Take Action or Not)

Accuracy	0.758	0.021
Precision	0.640	0.038
Recall	0.742	0.034
F1	0.687	0.036
Balanced Accuracy	0.746	0.024

Table 11: XGBoost Classifier Performance (No Action or Not)

Accuracy	0.818
Precision	0.880
Recall	0.896
F1	0.888
Balanced Accuracy	0.698

Table 12: XGBoost Classifier Performance (Take Action or Not)

Accuracy	0.736
Precision	0.662
Recall	0.694
F1	0.678
Balanced Accuracy	0.729

Table 13: Elastic Net Classifier Performance (No Action or Not)

Accuracy	0.811
Precision	0.830
Recall	0.963
F1	0.891
Balanced Accuracy	0.575

Table 14: Elastic Net Classifier Performance (Take Action or Not)

Accuracy	0.745
Precision	0.655
Recall	0.764
F1	0.706
Balanced Accuracy	0.748

Table 15: Lasso Classifier Performance (No Action or Not)

Accuracy	0.813
Precision	0.830
Recall	0.965
F1	0.892
Balanced Accuracy	0.576

Table 16: Lasso Classifier Performance (Take Action or Not)

Accuracy	0.738
Precision	0.646
Recall	0.764
F1	0.700
Balanced Accuracy	0.742

A.5 Top Tweeters Validation

Table 17: Most Inclusive Members of Congress on Twitter

	name	inclusive tweets − exclusive tweets
1	Pramila Jayapal	929
2	Adriano Espaillat	652
3	Richard J. Durbin	580
4	Joaquin Castro	482
5	Kamala D. Harris	467
6	Lucille Roybal-Allard	427
7	Carolyn B. Maloney	407
8	Nydia M. Velázquez	380
9	Robert Menendez	362
10	Donald S. Beyer, Jr.	361
11	Judy Chu	350
12	Catherine Cortez Masto	331
13	Yvette D. Clarke	297
14	Raúl M. Grijalva	296
15	Wm. Lacy Clay	290
16	Nanette Diaz Barragán	283
17	Ilhan Omar	265
18	Jeff Merkley	259
19	Tony Cárdenas	259
20	Barbara Lee	242

Table 18: Most Exclusive Members of Congress on Twitter

	name	exclusive tweets − inclusive tweets
1	Paul A. Gosar	338
2	Chip Roy	324
3	Andy Biggs	316
4	Ted Cruz	280
5	John Cornyn	235
6	Bradley Byrne	228
7	Steve Scalise	190
8	Brian Babin	168
9	Jeff Duncan	159
10	Martha McSally	153
11	Scott DesJarlais	123
12	Mo Brooks	115
13	Michael T. McCaul	113
14	Ken Calvert	109
15	Marsha Blackburn	99
16	Steven M. Palazzo	97
17	Jim Jordan	89
18	Jim Hagedorn	88
19	Kevin McCarthy	87
20	Warren Davidson	84

Table 19: Legislators Calling for Most Action on Climate on Twitter

	name	take action − no action tweets
1	Sheldon Whitehouse	2706
2	Donald S. Beyer, Jr.	1481
3	Edward J. Markey	1470
4	Jeff Merkley	1158
5	Paul Tonko	1031
6	Bernard Sanders	1009
7	Kathy Castor	910
8	Mike Quigley	895
9	Brian Schatz	829
10	Scott H. Peters	797
11	Thomas R. Carper	717
12	Suzanne Bonamici	681
13	Frank Pallone, Jr.	674
14	A. Donald McEachin	620
15	Alan S. Lowenthal	603
16	Jared Huffman	477
17	Martin Heinrich	453
18	Pramila Jayapal	448
19	Tom Udall	423
20	Benjamin L. Cardin	415

Table 20: Legislators Calling for Least Action on Climate on Twitter

	name	no action − take action tweets
1	Mitch McConnell	242
2	Harold Rogers	45
3	Keith J. Rothfus	29
4	Rand Paul	28
5	Bill Johnson	27
6	Bob Gibbs	25
7	Alexander X. Mooney	24
8	Sam Graves	23
9	H. Morgan Griffith	17
10	Patrick T. McHenry	17
11	Jason Smith	16
12	Jim Bridenstine	16
13	Steve Daines	14
14	Andy Barr	12
15	Pat Roberts	11
16	Cynthia M. Lummis	10
17	Martha Roby	10
18	Blaine Luetkemeyer	9
19	Chris Collins	9
20	David P. Roe	9

Table 21: Legislators with Largest Difference btw Inclusivity Tweet Score and HCS Score

	name	difference between tweet score and HCS score
1	Jon Tester	1.00
2	John Kennedy	1.00
3	David W. Jolly	0.78
4	Justin Amash	0.78
5	Danny K. Davis	0.75
6	Rodney Davis	0.73
7	Jerry Moran	0.72
8	Richard M. Nolan	0.60
9	Fred Upton	0.57
10	Bennie G. Thompson	0.53

Table 22: Legislators with Largest Difference btw Climate Action Tweet Score and LCV Score

	name	difference between tweet score and LCV score
1	Warren Davidson	0.99
2	John R. Curtis	0.99
3	Joe Wilson	0.94
4	Devin Nunes	0.93
5	Don Bacon	0.92
6	Dan Newhouse	0.92
7	Cathy McMorris Rodgers	0.92
8	Mike Gallagher	0.91
9	Francis Rooney	0.90
10	Bruce Westerman	0.89

Figure A.1: **Issue Positions on Immigration and Climate Change/ Environmental Regulation in Members of Congress' Tweets (Those who tweet the entire time only)**. Figure plots daily proportion of legislators' relevant tweets expressing issue positions on immigration (Panel a) and climate change/environmental regulation (Panel b) as local regression lines with loess smoothing and 95% confidence intervals. The data are all tweets produced by Members of Congress collected with Twitter's Historical Streaming API from January 2013 to May 2020 that are relevant to immigration (Panel a) or climate change / environmental regulation (Panel b). Tweets have been classified as being relevant to immigration or climate change using word2vec dictionaries trained on the full dataset of legislators' tweets. Relevant tweets are then classified as inclusive of immigrants/refugees, exclusive of immigrants/refugees, or neither in Panel a and supporting taking action to address climate change / supporting environmental regulation, opposing taking action to address climate change / opposing environmental regulation, or neither in Panel b.

(a) Immigration

(b) Climate Change / Environment