Using Twitter to Observe Election Incidents in the United States*

Walter R. Mebane, Jr.† Alejandro Pineda‡ Logan Woods§

Joseph Klaver¶ Patrick Wu‖ Blake Miller**

April 2, 2017

*Prepared for presentation at the 2017 Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 6–9, 2017. A previous version was presented at the 2016 Annual Meeting of the American Political Science Association. Thanks to Josh Pasek for letting us use his script used to extract data from Sysomos, to Catherine Morse for advice regarding use of Sysomos and to Joseph Hansel and Barry Snyder for assistance. Work supported by NSF award SES 1523355.

†Professor, Department of Political Science and Department of Statistics, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: wmebane@umich.edu).

‡Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: ajpineda@umich.edu).

§Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: ltwoods@umich.edu).

¶Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: jrklav@umich.edu).

‖Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: pywu@umich.edu).

**Department of Political Science, University of Michigan, Haven Hall, Ann Arbor, MI 48109-1045 (E-mail: blakeapm@umich.edu).

Abstract

Citizen complaints about election administration can be valuable to improve election performance, to help assess how well election forensics methods work, to address interesting behavioral questions and possibly to help establish the legitimacy of an election. In the United States such complaints as submitted through official channels are infrequent and often limited in scope. We use Twitter to extract observations of election incidents by individuals all across the United States throughout the 2016 election, including primaries, caucuses and the general election. To classify Tweets for relevance and by type of election incident, we use automated machine classification methods (e.g., support vector machines in an active learning framework). We demonstrate that for primary election day in one state (California), the distribution of types of incidents revealed by data developed from Twitter roughly matches the distribution of complaints called in to a hotline run on that day by the state. Questions we address include the following: Does the mix of problems vary between the primary and general election periods? Does the mix vary by features of each State, e.g., is it different in battleground states or in states that otherwise exhibit higher quality election administration? How does the content of Tweets change not only over the days before and after election day but also during election day?

1 Introduction

Election forensics is the field devoted to using statistical methods to determine whether the results of an election accurately reflect the intentions of the electors. Most such methods analyze information about voter participation or voters' choices, looking statistically for patterns that suggest frauds occurred (e.g. Myagkov, Ordeshook and Shaikin 2009; Levin, Cohn, Ordeshook and Alvarez 2009; Mebane 2010; Pericchi and Torres 2011; Cantu and Saiegh 2011; Beber and Scacco 2012; Mebane 2014; Montgomery, Olivella, Potter and Crisp 2015; Mebane 2016). It would be useful to draw other information into statistical analysis, both generally to enhance diagnosis of what happened in an election and more specifically to help address the primary challenge for election forensics: trying to tell whether patterns in election results that may appear anomalous in statistical estimates and tests are the results of election frauds or of strategic behavior (Mebane 2013, 2016).

One potential source of supplementary data to fraud statistics is reports from election observers. Indeed election observation, particularly that performed by international monitors, has become a global norm and some evidence has shown that it can improve the quality of elections (Hyde 2011). Election observation can be conducted either by international or domestic groups (Bjornlund 2004). Such monitoring is far from perfect. There is little in the way of international standards for election observation missions and the nature of this fragmentation can lead to biases in monitoring practices (Kelley 2012). These missions are also frequently limited in scope and can simply displace fraudulent activities (Ichino and Schuendeln 2012).

While most monitoring is performed by international organizations, numerous countries possess domestic institutions that enable citizens or domestic political parties to file formal election disputes, essentially deputizing these groups into the role of informal election observers. Such information is potentially useful due to the limited scope of more formal election observation. Mebane, Klaver and Miller (2016) and Mebane and Wall (2015) use such data, respectively from German citizens and from Mexican parties. In Germany the auxiliary data comes from citizen complaints about the federal election filed with a committee of the *Bundestag*, and in Mexico the additional information comes from petitions parties filed to try to nullify the votes counted in particular ballot boxes. In both cases the auxiliary data facilitate seeing that election forensics statistics are responding to strategic behavior or to parties' tactical actions, as well as perhaps to frauds.

Performing such analysis for elections in the United States faces difficulties. Leaving aside the difficulty of obtaining suitable low-level (say precinct-level) information about voters' participation and choices in the United States, information about citizens' observations of election incidents is difficult to obtain for several reasons. Election complaint processes in most states are convoluted and characterized by multiple possible channels for disputes, depending on which particular election law was supposedly violated. These channels could include submitting a complaint or dispute via an online portal, reporting an incident via phone, printing out a particular form and submitting a hard copy, or even simply emailing the relevant election authority. In many cases the process for filing a dispute, even once the appropriate channel has been determined, is itself burdensome, leading to few complaints being submitted. For instance, all complaints submitted in compliance with the Help America Vote Act of 2002 must be notarized. Consequently very few complaints are submitted via this process. Few states make what complaint data exist from other official channels publicly available.

To get information about American election administrative performance from individual observers throughout the country, we turn to data from Twitter. In this paper we describes steps we have taken to construct infrastructure to develop information from Tweets that contains election observations by individuals all across the United States. Here we focus on the presidential primary elections and caucuses held across the country in 2016 as well as on the Fall general election. The system involves extracting Tweets using keyword filters, collecting information about election officials' and other leading actors' Twitter accounts, classifying Tweets for relevance and coding Tweets for type of incident. For the classification tasks we apply automated machine classification methods to Tweet texts, although images associated with Tweets are important for classification decisions.

We demonstrate that for primary election day in one state (California), the distribution of types of incidents revealed by data developed from Twitter roughly matches the distribution of complaints called in to a hotline run on that day by the state. In terms of clarity of type definition and in terms of number and geographic dispersion of incidents, the data derived from Tweets may be superior to the officially collected hotline data. For the general election period we currently have coded by hand a random sample of Tweets that will be an initial training set for an active learning process. The sample suggests that long lines are the most frequent kind of incident observed in Twitter data, and that incidents may be associated with election administration features such as how early voting and absentee ballots are handled and with demographic features such as the racial composition and educational attainment of state populations.

2 Election Complaints in the United States

Nothing like the centralized system of election complaints that exists in Germany exists in the United States. Each state implements its own complaint process, or none. The Help America Vote Act (HAVA) mandates that each state create a complaint process (42 U.S.C. §402 (a) (1), (2002)), and state codes implementing HAVA frequently state that resolved complaints will be made available to the public online. But the HAVA infrastructure is moribund and useless. Using the methods described in Appendix section 5.1, we reviewed information about complaint procedures in each state. Table 1 summarizes information about the online presence of complaint procedures: whether links could be found to state HAVA forms or to other complaint-related forms; whether it is possible to submit complaints online; and whether some information about submitted complaints appears to be available online. As can be seen, few states allow complaints to be submitted online. Moreover the HAVA complaint process is burdensome, requiring the submission of notarized documents. The submission process in many states is equally burdensome. Submitted complaints are not actually available online for all states that have online links supposedly to the complaints, although some states do post some version of the official rulings on selected complaints.

*** Table 1 about here ***

The inutility of the HAVA complaint procedures is exemplified by the number of HAVA complaints that exist. Table 2 reports the number of HAVA complaints we were able to obtain after contacting all the states. We obtained HAVA complaints for five states. Most states did not respond substantively to written inquiries.[1] The numbers of HAVA complaints we received from states is miniscule. For instance, Ohio sent the single HAVA complaint submitted regarding the 2012 election. Perhaps the reason for the general lack of responsiveness is that frequently a state had no HAVA complaints. The general lack of existence of HAVA complaints was confirmed by a commissioner of the Election Assistance Commission.[2] HAVA complaints cannot be used as a measure of the quality election administration or the existence of possible election frauds.

*** Table 2 about here ***

Many states have alternative avenues for election complaints outside of the HAVA-mandated procedures. These processes produce more extensive and more widely available documents, but even so the state-level complaints are not sufficient to provide citizen observation data across the

[1] Although the plan to implement HAVA in California states, "The determination must be posted on the Secretary of State's website, unless such posting might compromise a criminal investigation or other enforcement action" (California Help America Vote Act State Plan 2010 Update, §254(a)(9)(4), (2010)), California refused to provide any HAVA complaints they may have. They wrote,

> "It is the longstanding policy and practice of the Secretary of State, like all agencies with investigative functions, to treat complaints and information associated with those functions as exempt from the requirements of disclosure pursuant to the California Public Records Act. Specifically, Government Code section 6254(f) allows the non-disclosure of investigative information because the disclosure of the information would impair an agency's ability to conduct investigations in the public interest.
>
> "Consistent with guidelines issued by the California Attorney General, material covered by the non-disclosure exemption of Government Code section 6254(f) includes, 'Records of complaints, preliminary inquiries to determine if a crime has been committed, and full-scale investigations, as well as closure memoranda...'
>
> "For these reasons, the Secretary of State is unable to provide the information you have requested. However, we are open to speaking to you to discuss ways in which we might be able to provide helpful information to you and your research." (Secretary of State, Constituent Affairs 2015)

We had requested copies of any HAVA complaints going back to 2006.

[2] Personal communication from Matt Masterson on March 10, 2016, at the Election Verification Network (EVN) Conference, Washington, DC.

whole of the United States. Table 3 lists states for which we have been able to obtain documents regarding state complaints. Despite what the "Complaints Online" column in Table 1 may suggest, to date we have obtained documents from only 11 states. In some cases documents are online, but in other cases we received documents (some electronic, some on paper) in response to written requests. Beyond the limited availability of any information regarding the complaints, the "complaints" we have obtained are often of limited utility. In a few cases we have the original text of complaints, but in most cases the documents reflect the results of administrative adjudication of each complaint. In Germany Mebane, Klaver and Miller (2016) also have only the reports of official adjudications by the *Bundestag* committee, so such indirection is not necessarily a problem. More problematic is that in most states all or most complaints made available concern allegations of campaign finance violations or of procedural irregularities committed by parties or by candidates. Few complaints concern citizen observation of other particulars of the electoral process, including especially actions by elected or appointed election officials. The state complaints also cannot generally be used to assess election administration or to detect election frauds.

*** Table 3 about here ***

Some states operate election-day hotlines, usually phone hotlines. The states we identified as having such a hotline operation in the most recent election are listed in Table 4. Just fewer than half the states have a hotline.[3] Some states that operate hotlines say they do not record information about the calls they deal with.[4] Hotline data also cannot be used to generate data for the whole country. Below we will use data provided by California from the hotline that state operated for the 2016 primary election.

*** Table 4 about here ***

[3]In some states some counties operate hotlines, and in others voters are simply instructed to contact the elections division main phone number.

[4]We distributed a survey to state election officials asking them about their voter hotline policies but received scant responses.

3 Using Twitter to Capture Election Observations

We begin to construct infrastructure to allow information in Tweets to be used to build data regarding election observations by individuals in the United States. We focus on the presidential primary and caucus elections in all states in 2016 and on the 2016 general election. For the primary/caucus period we collect Tweets from within date windows beginning ten days before and ending ten days after each election day. For states that allow absentee (mail-in) voting in primaries, we begin collection on the first day that absentee ballots can be submitted as votes. For the general election period we collect Tweets continuously starting on October 1 and ending on November 8.

3.1 Collecting Twitter Data

We used two modalities for collecting Tweets: the Sysomos MAP (Sysomos 2016) search tool and archive and the Twitter API (Twitter, Inc. 2016*b*).

We used Sysomos for the primary/caucus Tweets because Sysomos supports searching for Tweets using keywords for a period going back 12 months.[5] With Sysomos MAP (Sysomos 2016) we used state names in the "location" field along with search terms to obtain Tweets. Initially we used more extensive keyword sets when downloading Tweets manually, while the more limited sets are used when downloading using a script in an automated process. The initial states are Arizona, California, Colorado, Connecticut, Illinois and Washington. The search terms used in these cases are listed in Table 5. To define the search terms in the less extensive sets of terms, we first obtained a list of election official, party and other Twitter accounts ("handles") (see Appendix section 5.2 for details regarding compiling the list).[6] We combined the most productive kinds of keywords found in performing the manual searches (e.g., "azprimary") with search terms that would capture Tweets sent to officials (e.g., "to:CASOSvote"). The resulting sets of terms are listed in Table 6. Finally for all states we ran Sysomos searches using the keywords listed in the

[5] We began downloading Tweets on June 20, 2016.

[6] The proportion of county election offices that have an affiliated Twitter account varies greatly across states.

note.[7]

*** Tables 5 and 6 about here ***

Using the Twitter REST API (Twitter, Inc. 2016*b*) we downloaded the timelines of 493 Twitter accounts. Use of the API gave more control over the data than the use of Sysomos, as Sysomos only returns certain fields, and the data returned is from a random sample (which we cannot be certain is truly random). The API returns more comprehensive and more complete data. To access the API, we registered an application with Twitter.com, giving us the security tokens necessary to query data from Twitter's database. Our goal was to pull entire timelines from 493 accounts (for perspective, one California account had over eleven thousand Tweets in their timeline). Further details about building the list of accounts and about the process of extracting Tweets using the API are in Appendix section 5.2.

Table 7 shows the number of unique Tweet texts downloaded from each state for the primary/caucus period. Retweets are excluded. We use the `location` specified to Sysomos to determine the state for each Tweet. California has the most unique Tweet texts (60,350), followed by Hawaii (25,256) and Iowa (21,520). For each other state there are less than 20,000 unique Tweet texts. Montana has the smallest number of unique Tweet texts (300).

*** Table 7 about here ***

For the general election period we used data from officials' timelines along with data from the Twitter Streaming API (Twitter, Inc. 2016*a*). Keywords we used to select Tweets are shown in the note.[8] In all during October 1 through November 8 we downloaded 44,329 Tweets from timelines

[7]Sysomos Keywords: Line to Vote, Long Line to Vote, Problems Voting, Voting Rights, Right to Vote, Election Fraud, Corruption, Voter Fraud, Stole Election, Election Stealing, Voter ID, Voter Identification, Election Complaint, Broken Voting Machine, Election Officials, Disenfranchised, Campaign Finance, Primary Election, General Election, Voter Complaint, Polling Place, (State)Vote, Vote(State), (State)Election, (State)Primary.Caucus, Una Fila Para Votar, Larga Fila Para Votar, Problemas de Votacion, Derecho Al Voto , Derecho Al Votar, Fraude Electoral, Corrupcion, Colegio Electoral, Elecciones Robo, Robo Electoral, Identificacion Del Elector, La Identificacion Del Votante, Queja Electoral, Maquina De Votacion Roto, Funcionarios Electorales, Privados De Sus Derechos, Financion De Las Campanas, Eleccion Primaria, Eleccion General, Quejas De Elector.

[8]Twitter API Keywords: line to vote, long line to vote, wait to vote, absentee voting, early voting, problems voting, voting rights, right to vote, election fraud, corruption, voter fraud, stole election, stolen election, rigged,

and 16,221,304 Tweets via the Streaming API. Removing retweets leaves 6,163,892 unique Tweets which contain 4,566,333 unique Tweet texts. Only 598,783 Tweets have `place` and `fullname` information (see Appendix 5.2), which is needed to be able to locate any incident observation reliably in geography, which means to place it in a state, city or neighorhood. We drew a sample of 20,942 Tweet texts from the Tweets that have `place` and `fullname` information and labeled them by hand as containing an incident observation ($n = 1,591$) or not. Table 8 reports the distribution of the sampled incident observations over states. If the "hit rate" (.076) in our sample matches the rate throughout the whole body of Tweets, we might expect to observe 345,321 incidents once we complete the process of mining all the Tweets. Applying the hit rate to the subset with `fullname` information we may expect to find 45,507 more general election incident observations that can be reliably located geographically.

*** Table 8 about here ***

3.2 Categorizing Twitter Data

To determine whether a downloaded Tweet includes any relevant observations of the electoral process and then to say what types of incidents are being reported, we augment, clean and classify the Tweets.

3.2.1 Text Augmentation

We augment the text "content" of each Tweet in two ways. In general we get the resource, if any, located at each URL the content contains. If that resource contains any text, we capture that text

---

election stealing, tamper, manipulate, voter id, voter identification, election complaint, election problem, broken voting machine, election officials, electronic voting, election audit, election observer, poll watch, vote protection, election protection, disenfranchised, campaign finance, election system, primary election, general election, voter complaint, polling place, registration database, statevote, votestate, stateelection, vote count, vote tabulation, voter database, voter registration, voter suppression, voting machine, voting machine hacked, vote not counted, vote, US election, American election, not enough ballots, absentee ballot, voter intimidation, voter harassment, mail in ballot, vote by mail, voter hotline, waiting to vote, precinct, precinct officials, precinct captain, replacement ballot, ballot selfie, my ballot, my vote, eleccion, fila para votar, derecho al voto, derecho al votar, fraude electoral, maquina de votacion, funcionarios electorales, colegio electoral, neo-nazi, white supremacist, white nationalist, alt-right.

and append it to the original content.[9] If that resource contains an image, we capture the image's URL.[10] We plan to augment the original Tweet content also with a text description of the image, but this process has not yet been fully operationalized. See Appendix section 5.4 for a description of the process we have used to date with a sample of Tweets. As it is, images often affect human coders' judgments regarding any information Tweets may contain, but the machine classification algorithm currently does not have access to images or descriptions of images.

3.2.2 Tweet Text Cleaning

Cleaning the augmented Tweet content involves removing nonprintable characters, stray HTML codes, internal quotation marks and the '*' character. For the version of the contents used in machine classification and active learning processes, we also removed URLs and made some frequently occurring text strings generic instead of specific to each state. The latter changes replaced some state-specific strings with strings like "#XXvotes," "#XXprimary," "#XXcaucus" and "#XXvoterfraud," where "XX" originally was the postal code abbreviation for a state. We did this to enhance the comparability of Tweets across states for the machine classification algorithm.

3.2.3 Classifying Tweets

To determine whether each downloaded Tweet includes relevant observations, we began by using humans[11] to examine the raw Tweets directly. A Tweet that contains relevant observations about electoral processes is coded to be a "hit." Each hit was also classified into one or more categories. For the primary/caucus period the classification rules are based on the incident type categories in EIRS and in Mebane, Klaver and Miller (2016). Through several rounds of coding, discussion and recoding of random samples of Tweets from Arizona, California, Colorado, Connecticut and Washington[12] we developed consensus criteria for deciding that a Tweet is a hit and for what types to use to classify incidents. For the general election period the classification rules are

[9]Specifically, we capture any text in the `og:description` field in the resource's HTML code.
[10]Specifically, we capture any URL in the `og:image` field in the resource's HTML code.
[11]The human coders were subsets of this paper's authors.
[12]In Washington Tweets come from both the Democratic caucuses and the Republican primary elections.

modified to refer to all observed incidents without emphasizing "complaint" observations. The procedure we developed for humans to use when making hit determinations for the primary/caucus data is shown in Figures 4 and 5. The procedure for humans to use when making hit determinations for the general election data is shown in Figures 6, 7 and 8. The background for these flowcharts is discussed in Appendix section 5.5.

*** Figures 4 and 5 about here ***

The coding rules for categorizing the incidents to which hits refer are described in Appendix section 5.6 (for primary/caucus Tweets) and in Appendix section 5.7 (for general election Tweets). For coding primary/caucus incidents by type there are 15 categories: Absentee, Mail-In, or Provisional Ballot Issue; Registration Issues; Disability/Accessibility Problem; Improper Outside Influence; Other Ballot Problems; Election Official Complaints/Incidents; Electoral System; Voter Fraud; Voter Identification Issues; Long Lines/Crowded Polling Place; Polling Place Problems; Voting Machine complaints; Unspecified Other; Positive; and Ambiguous. These categories collapse several EIRS categories into each other, and definitions of categories are modified accordingly. Categories are collapsed into each other when they are thematically related. For example, the categories regarding mail-in, provisional, and absentee ballots are combined. An additional category, Not Hit, is used when a human coding a Tweet the machine classification algorithm classified as a hit decides the Tweet is not a hit.

For coding general election Tweets there are twelve main categories: Outside Influence; Disability/Accessibility Issue; Line Length, Waiting Time, Polling Place Crowding; Polling Place Event; Electoral System; Absentee, Mail-In, or Provisional Ballot Issue; Election Official; Voter Identification; Registration; Voter Fraud; Ballot and Voting Technology; Unspecified Other. For most of these categories we also record which "adjective" modifies the incident. For example, for the Line Length, Waiting Time, Polling Place Crowding category adjectives distinguish no lines from short lines from long lines. See Appendix section 5.7 for details regarding the definition of these adjectives. Many adjectives reflect judgments about things working well or poorly, but our coding scheme does not depend on and is not intended to measure any kind of emotional sentiment. For example, many express warm feelings when encountering a very long line to vote: we record the long line and ignore how the person Tweeting said they felt about it.[13]

To produce a training set to use to start the machine classification algorithm with the primary/caucus Tweets, we used a stratified random sample[14] of Tweets from the manual Sysomos downloads from Arizona, California, Colorado, Connecticut and Washington. The Tweets in that sample were coded as "hit" or "not a hit" based on whether at least three of five human coders agreed (upon coding the Tweets again) that the Tweet is a hit or, for Tweets that did not attract such agreement, by using the flowchart. This produced an initial training set containing 192 hits and 806 not-hits.

We use a support vector machine (SVM) in an active learning framework to classify Tweets. Active learning allows us to build a training set with fewer labeled observations, and a better balance between classes, which is useful because of the scarcity of the positive "hit" class (Miller and Klaver 2016). Prior to classification, we first preprocess the text of each Tweet's augmented content. This involves removal of all duplicate texts, stemming, stop word removal and the eventual transformation of text data into a numeric matrix. After stemming, we remove stop words. Stop words are words we believe have low relative information value, or words we believe are uninformative to our classification task. Such words include "and," "the," etc. Once we have preprocessed the words in each document, we use a word n-gram model for Tweet text and a character n-gram model for hashtags to convert the Tweet corpus into a document term matrix: each row represents a Tweet and each column represents a unique character or word unigram, bigram or trigam. Cell values represent the count of each word in the document. Finally we do a TF-IDF transformation of the raw count matrix (Leopold and Kindermann 2002; Lan, Tan, Low and Sung 2005). Because the feature space is high dimensional, and we want to avoid overfitting, we select features using the coefficients of a linear SVM with ℓ_1 norm penalty. Features with SVM coefficients lower than the mean of all coefficients are discarded (Rakotomamonjy 2003).

[13] We plan to recode the primary/caucus Tweets using the general election scheme that's concerned with "observations" and not more narrowly "complaints."

[14] For a description of the sample see Appendix section 5.3.

Our initial classificatory scheme to discriminate hits from not-hits is based on an initial human-labeled training set. Because the initial training set is small, we use active learning, an iterative supervised machine learning technique (Settles 2010). This framework uses uncertainty sampling to identify observations that we should label by hand to provide the most useful new input to the next iteration of the classifier. At each iteration, we train a SVM on labeled Tweet texts. We use the distance from the SVM's separating hyperplane to measure model uncertainty. We then iteratively label the texts closest to the hyperplane and refit a model until acceptable average precision, recall and F-measure are achieved.

To date for the primary/caucus date humans have manually labeled 9,417 Tweet texts, which includes texts from the 998 Tweets in the initial training set. Among the human-labeled texts, 1,204 are hits and 8,213 are not-hits. Over all unlabeled Tweet texts the SVM classifies 44,272 texts as hits and 270,488 as not-hits. SVM performance measures, based on a weighted cross-validation method,[15] are shown in Table 9. Overall we achieve average precision, recall and F-measure of .78, .79 and .77, respectively.

*** Table 9 about here ***

3.3 Characteristics of Primary/Caucus Tweet Contents and Incidents

According to the SVM, both hits and not-hits occur in every state in the primary/caucus period. As Table 10 shows, in every state the number of not-hits exceeds the number of hits. For California we also have a breakdown of hits versus not-hits by county, as is shown in Table 11. The number of Tweet texts and of hits is largest in Los Angeles, although overall both numbers seem roughly proportional to the population of each county. In every county except Alpine, Inyo and Mono counties the number of not-hits is greater than the number of hits, although the number of Tweet texts in those three counties is extremely small.

*** Table 10 and 11 about here ***

[15]Because the number of hits is so much smaller than the number of not-hits, sample sizes for cross-validation are constrained so that the expected number of not-hits sampled is the same as the number of hits.

While in the future we plan to use machine classification to classify incidents by type, for the moment we have humans performing such classifications manually, according to the scheme described in Appendix section 5.6. From the Tweet texts with locations in California that the SVM classifies as hits, we selected a simple random sample of $n = 600$ to classify by type manually. Table 12 shows these type frequencies. Among both the unique Tweet texts and the unique Tweets that have those texts (for which $n = 700$), Polling Place Problems are the most frequent type of incident, followed by Improper Outside Influence, Absentee Mail-in or Provisional Ballot Issues, Long Lines/Crowded Polling Place, and Electoral System concerns.

*** Table 12 about here ***

Notable is that human coders decided that 249 of the 600 sampled Tweet texts that the SVM had classified as hits were actually not-hits. A proportion of $.585 = 1 - 249/600$ is a bit smaller than the .66 precision value for hits reported in Table 9. It may be that such a discrepancy reflects variation in classifier performance across states, but in any case it suggests that the number of human-labeled texts should be increased.

Polling Place Problems remain the most frequent type of incident in California when we consider only the subsample of texts from Tweets on election day (June 7, 2016). Table 13 shows the election-day type frequencies. Omitting texts that express positive evaluations of the remarked situation, on election day Absentee Mail-in or Provisional Ballot Issues are second-most frequent in the subsample, while Long Lines/Crowded Polling Place and Improper Outside Influence are tied for third. If the sample size for the comparison between proportions is taken to be $n = 103$, then the proportion of Polling Place Problems among texts that are not Positive ($n = 34$) is significantly greater than the proportion of Absentee Mail-in or Provisional Ballot Issues ($n = 19$), but the proportion of Absentee Mail-in or Provisional Ballot Issues is not significantly greater than the proportion of Long Lines/Crowded Polling Place or Improper Outside Influence incidents ($n = 14$).

*** Table 13 about here ***

3.3.1 Comparisons to the California Hotline

On primary election day in 2016 California operated a statewide voter hotline (Plummer 2016). The distribution of complaints recorded by hotline operators appears in Table 14. Because no codebook for the California categories is available to explain their meaning,[16] it is difficult to say how the distribution of hotline complaints compares to the distribution of election-day Tweet texts presented in Table 13. Nonetheless Poll Worker Problem alone is the most frequent hotline complaint, Polling Location is the second most frequent and Closed Polling Place is fifth. Perhaps those frequencies are a match for Polling Place Problems being the most frequent type of incident in the election-day Tweet texts. Voter Registration concerns are 11.4 percent of hotline complaints but Registration Issues describe less than five percent of election-day Tweet texts. Provisional Voting and Vote by Mail Ballot together are less than five percent of hotline complaints (Voting Process Issue complaints are another 3.9 percent), while Absentee Mail-in or Provisional Ballot Issues are 18.4 percent of election-day Tweet texts that are not Positive. On the whole there are many differences between the hotline complaints distribution and the distribution of election-day incidents that Tweet texts point to, but the distributions are not utterly unlike one another.

*** Table 14 about here ***

An important difference between the hotline complaints and the election-day Tweet text data is the latter have more extensive geographic coverage across the state. Table 15 shows that hotline complaints come from 31 counties, with most complaints coming from Los Angeles and other large population counties. A pattern in which large population counties have the most observations also occurred for the Tweet texts that are hits, as shown in Table 11 for a time period that includes but is not restricted to election day. Table 16 shows that on election day Tweet texts that are classified as hits occur in 41 counties as well as in the "Bay Area" (which includes "East Bay") and in "Silicon Valley" (without reference to a particular county). The tendency for more hits to occur in more populous counties continues to occur.

[16]Codings were left to the discretion of the individual hotline operators (Pancharian 2016).

*** Tables 15 and 16 about here ***

Not all the instances classified as hits will prove to be hits on closer inspection—recall that only 58.5 percent of classified hits proved to be hits upon examination by a human (59.3 percent in Table 13, for election day). But the machine classification performance will very likely improve once a greater number of Tweets are labeled by a human in the active learning process. Even with likely reductions in the number of hits, more incidents and more widely dispersed incidents are likely to be identified by the Twitter data than there are complaints in the hotline data.

3.4 Characteristics of General Election Tweet Contents and Incidents

For the general election we have classification results for the Tweets in the sample used to develop a training dataset. Table 8 reports the distribution of counts of the sampled incident observations over states, and Table 17 shows the distribution in terms of observations per million persons. While the sample is too small for the distribution estimates to be reliable, on a per capita basis Texas and California do not appear to be the states with the most incident observations. Instead the District of Columbia stands out followed by Nevada and North Carolina. Bivariate regression analyses show the per capita observation values conditionally depend on several variables: on a state requiring some form of photo or non-photo identification; on a state allowing no excuse absentee voting; on a state not having a complaint process outside of HAVA; on there being at least one way (HAVA, non-HAVA, online portal) for voters in a state to submit complaints online; on the general election voting-eligible-population (VEP) turnout in the state; and perhaps on a state allowing early voting or in-person absentee voting. It will be interesting to see whether these or other associations appear when the sample of classified Tweets is larger.

*** Table 17 about here ***

Plotting incident observations in the training sample by day shows that the most observations occur on election day (see Figure 1).[17] Figures 2 and 3 show the distributions over time of

[17]The last bar in the histogram in Figure 1 corresponds to November 9, which is the date associated with some Tweets due to our expressing all times in Eastern Standard Time units.

incident observations by type. Long lines or waiting times to vote are the most frequent observed incident, although many also observe that lines or waiting times are not very long on election day (Figure 2(a)). The next most frequent kind of observation is a report of success voting on election day, during early voting or by absentee ballot, although many also report problems affecting voting or polling places (Figure 2(b)). Reports of problems with voter registration are nearly as frequent as reports of success with voter registration (Figure 2(c)), and reported problems with voter identification are only slightly more frequent than reported successes with voter identification (Figure 2(d)). Problems with ballots or voting technology are reported both during early voting and on election day (Figure 3(a)). Complaints about election officials occur frequently on particular days (Figure 3(b)): the two peaks in the figure refer to actions by officials in Florida (on October 7) and Wisconsin (on October 25), and given the similarity in language among the referent Tweets these may stem from Tweets by bots. Both complaints and praise of aspects of the election system are apparent (Figure 3(c)). A few observations refer to problems with voter fraud (Figure 3(d)).

*** Figures 1, 2 and 3 about here ***

On a per capita basis some types of incident conditionally depend on state-level characteristics. For example three-category observed line length depends on a state allowing no excuse absentee voting (longer lines), on a state requiring identification (shorter lines) and on the general election VEP turnout (longer lines). Success voting depends on a state allowing no excuse absentee voting (more success).

Some types of incident on a per capita basis also conditionally depend on state-level demographics characteristics. For example line length depends on the proportion of the population that identifies as white-alone, as black-or-african-american-alone or as hispanic-or-latino; the interaction between hispanic-or-latino and black-or-african-american-alone proportions is statistically significant as is the interaction between hispanic-or-latino and white-alone proportions, so the marginal associations are not straightforward to characterize. Success voting depends on the hispanic-or-latino proportion and on the black-or-african-american-alone proportion separately, with the interaction between the two proportions perhaps being statistically significant. All these associations also interact significantly with the bachelor's-degree-or-higher proportion.

4 Discussion

Every indication is that Twitter can be used to develop data containing individuals' observations of how American elections are conducted, data that cover the entire country with extensive and intensive local detail. Observations for each day can be gathered, and observations can be even more finely resolved in time (using the timestamps on Tweets). The frequency and likely the diversity of observations may vary depending on how many people care about an election and want to participate in it, observe it and comment on it. Some Tweets seem like shouts into the void (although maybe such a view underestimates the importance of "Twitter followers"), but others are messages directed specifically at election officials. One question we will eventually investigate is whether those two types of Tweets typically convey information about different kinds of election incidents.

An important immediate step for development is to try better to exploit the geolocation information that exists for a small proportion of Tweets. `place` information is available for some Tweets obtained via the Twitter API. Here we have illustrated how for such Tweets geography can be reliably resolved to states, but in fact in many cases resolution is possible to the city, neighborhood or even building. We envision using such geographic identifications to place Tweets in particular election districts. Ideally we would like to associate Tweets with particular polling places, but for most Tweets that will not be possible.

`place` information is not available for most Tweets from the Twitter API, and for Tweets obtained via Sysomos "location" information appears to come from user profiles. Such "location" data usually reflects the location associated generally with (and chosen by) the sender of the Tweet, not the place whence the Tweet originated. Perhaps in cases where voting happens in person, we can rely on selected locations to correspond both to where the sender lives and to the place where the sender is trying to vote—but clearly such is not a generally reliable assumption. Perhaps geolocation data can be used to develop models to estimate the likelihood that Tweets that do not have reliable geolocation information actually come from the place the "location" indicates. "Location" information is also often vague, which makes it challenging to associate incidents with particular election districts. That presents a challenge for the goal to combine such information with information about votes.

We don't know what observational biases affect the set of incidents observed using Twitter data. An obvious bias is that Tweets come only from individuals with a smartphone who use Twitter, and such individuals may not be be as frequently present at every place from which we would like to observe election incidents. Privacy settings in Twitter also limit the number of tweets we see, and incidence of (for us) adverse settings may vary across time and space. When we rely on Tweets at election officials we may be biasing our data to include more observations from states with high degrees of professionalization in their county governments.

Also it is entirely voluntary to send a Tweet, so the availability of Tweets depends in unknown ways on individual characteristics. In the future we hope to get some purchase on the characteristics of people who Tweet incident observations, by examining their timelines and their networks of fellow Twitter users.

In most cases we cannot know whether purported incidents actually occurred, although in a few cases incidents alleged in Tweets can be verified by information obtained from other channels such as news reports or official reports. Many other questions will arise regarding observations derived from Twitter, but at this point it seems better to get the data then critically appraise them rather than not obtain the data at all.

5 Appendices

5.1 State Election Resources

We are attempting to find and acquire copies of resolved official complaints filed with each state. Although state Help America Vote Act code frequently states that resolved complaints will be made available online to the public, this is rarely the case. In order to facilitate our searches for resources, we went to the official elections (or substantively similar) website for each state and collected what data we could regarding the available official complaint forms for each state, the contact email for each state election director, the mailing address for each state elections division, the phone number for each state elections division, the "voter hotline" phone number for each state as available, and the link to each state's online complaint portal as available.[18] We also obtained the election laws for each state. These data were collected over the span of the past year, and the accuracy of these data relies upon each state's official elections website. As is suggested by Table 1, we have collected many state-level resources for each state and the District of Columbia.[19]

We contacted states by reaching out to the state election division directors (or equivalent role) directly. We avoided general public e-mail accounts or hotlines. The e-mails of the state election directors were found in a publicly available National Association of State Election Directors (NASED) roster, last revised January 16, 2016. Below is an example of a generic e-mail sent out. E-mails were modified with links to state complaint forms (where available), state HAVA complaint forms (where available), and always referenced the publicly available state complaints of the State of Colorado. E-mails were almost always sent during the mornings or early afternoons on weekdays in an attempt to increase the probability that they will be seen by the state elections director. Any requests for formal FOIA requests were completed as well.

[18] The most recent update to the state election directors contact information was found through the National Association of State Election Directors at `https://www.nased.org/`.

[19] We plan to build an elections resources website to publicize the information we have collected, to be made available for the 2016 general elections.

Subject: [State Name] Voter Complaints Data

Dear [Elections Director Name Here],

I am a graduate student at the University of Michigan's Department of Political Science gathering data related to election administration in the United States. Specifically, I am collecting data on election and voter complaints. I was wondering if the State of [State Name Here] has any collection of voter complaints available on its website, similar to what the State of Colorado currently has (`http://www.sos.state.co.us/pubs/elections/complaints/index.html`).

If not, I am interested in receiving copies of complaints that were either submitted through state-complaint processes ([Link to state-complaint form]) or complaints that were filed as HAVA complaints ([Link to HAVA complaint form]). I am interested in all years where election complaints are available.

Thank you, in advance, for your assistance with my research project. I look forward to hearing from you very soon.

Best,

5.2 Twitter API Data

To access the Twitter API (Twitter, Inc. 2016*b*), we registered an application with Twitter.com, giving us the security tokens necessary to query data from Twitter's database.[20] In order to collect Tweets to and from election officials on and around the respective Election Days, we first had to find the Twitter accounts for those election officials.

These Twitter accounts were found in two ways: first, the Election Assistance Commission has collected information regarding the social media accounts of election officials at both the state and county levels across the United States, with varying degrees of completeness of data across

[20]We used a combination of Python modules, mainly Twython and Tweepy. Code was adapted from (Bonzanini 2015; Moujahid 2014; Saxton 2014; Dolinar 2015)

states.[21] The second way these Twitter accounts were obtained was by manually searching Twitter for terms associated with the office of election officials, such as "election official," "county clerk," "department of elections" and "county auditor." Along with manually searching for election officials, user-created lists of election officials were searched for previously not-found election officials.[22] We used similar methods to find the Twitter accounts of state-level Republican and Democratic Parties, state-level Leagues of Women Voters, and state-level ACLUs. In order to facilitate these searches, we created a Twitter account affiliated with this research project.[23]

Our goal was to pull entire timelines from 493 accounts (for perspective, one California account had over eleven thousand Tweets in their timeline). A few challenges arose in querying that much data. First, user timelines are not static: a user can post Tweets while our application queries the data, which would effect the results; we had to recursively pull Tweets twenty at a time, starting with the user's most recent Tweet and ending with the first Tweet posted (in some cases dating back to 2007). Second, the sheer size of the query would occasionally break the script, so we had to pull timelines in batches; that is, we could not pull all 493 accounts at once, but rather, pull them fifty at a time. For perspective, a single batch would return hundreds of thousands of Tweets. Finally, Twitter places rate limits[24] on applications that query data from the API, so we had to design the script to pause in between requests. This way, we would not exceed rate limits, and the script could complete each query.

Part of the data collection was to identify tweets by their unique identification number, allowing us to quickly identify and omit duplicate tweets from our final dataset. The data returned are formatted in JSON[25], so we had to identify the specfic fields of interest (in this case, the

[21]The list of resources can be found at `http://www.eac.gov/voter_resources/state_and_local_election_office_social_media_list.aspx`.

[22]An example of one of these user-created lists can be found at `https://twitter.com/EACgov/lists/us-election-officials/members`.

[23]The Twitter user name for this account can be found at `https://twitter.com/election_ballot`.

[24]Enforced on a "per access token" basis, Twitter limits users to fifteen requests per fifteen-minute window, although this number varies with the object being called; for more information on Twitter rate limits, see `https://dev.twitter.com/rest/public/rate-limiting`.

[25]JavaScript Object Notation, a data format represented by simple text, used to transfer data objects that consist of atrribute-value pairs; for more information on the format of Twitter data, see: `https://dev.twitter.com/rest/reference/get/search/tweets`.

unique identification number of each Tweet, its content, its timestamp, the name and location of each user, and the place whence the Tweet was sent, which was missing in most Tweets) and write them to a .csv file. Additionally, we were interested in obtaining geo-location data from each tweet (returned in the form of coordinates) but Twitter's privacy settings are such that, this kind of data is not readily available for most users.

The bulk of the content was from outside of our time range, so it was not used. For the primary/caucus period we made sure that the data used from the Twitter API were from the same time frame as the data obtained via Sysomos. Part of the data collection was to identify Tweets by their unique identification number, allowing us to quickly identify and omit duplicate Tweets from our final dataset.

The second phase of data collection started in October 2016, corresponding with the beginning of early voting in the general election. Because we were now streaming data, we could use keywords as filters to capture tweets of interest. These keywords signaled issues with voting—voter complaints, registration issues, long lines, broken machines, etc. To supplement the absence of geo-location data, we pulled data from the `place` object. This object is part of the JSON metadata, but is associated with individual tweets rather than with a users' profile.[26].

The `fullname` field is used to do a reverse lookup of the state. Our code uses the GeoPy module in Python to access the Nominatim search tool used by OpenStreetMap. The tool itself allows for non-standard search of places and returns a standard dictionary of addresses and latitude/longitude coordinates.[27] The GeoPy module also offers the use of Google Maps, Bing Maps, or Yahoo BOSS, but the Nominatim geolocation service has the advantage of breaking down addresses into key-attribute pairs (Python dictionaries), whereas the other services rely solely on comma separated values. As addresses are not standardized, this can be problematic because the `state` field will not be in the same location for every query. Search results were

[26]`place` is specified at the time a user posts a tweet: users are asked if Twitter can access location information, and if they respond yes, the object is attached to the tweet: "*Places* are specific, named locations with corresponding geo coordinates. They can be attached to Tweets by specifying a *place id* when tweeting. Tweets associated with places are not necessarily issued from that location but could also potentially be about that location" (Twitter 2017)

[27]OpenStreetMap is an open source, collaborative project that seeks to produce geographical data provided by users. Companies that use OpenStreetMap data include: Apple Inc., Flickr, MapQuest, and Craigslist (OpenStreetMap 2017).

checked by the authors to ensure the states returned matched the addresses provided in the Twitter metadata.

5.3 Stratified Sampling for Tweets in Training Set

The stratfied random sample used for the initial primary/caucus training set contained 1,001 Tweets of which n = 998 are unique Tweets (unique based on the 18-digit Twitter ID number). The population used for sampling was the union of the distinct samples drawn previously for use in developing the coding schemes. Strata were defined by state, by type of search terms used to find Tweets and by whether any human identified the Tweet as a hit in the initial round of coding (that is, before the flowchart of Figures 4 and 5 existed). The stratum labels derived from state and search terms are AZ, CA, CT, CO, WAd, WAr, CAeo, COeo and WAeo, where the first two letters are a state's postal code, "WAd" refers to the Democratic caucus, "WAr" refers to the Republican primary and the "eo" suffix means search terms focused on election officials. Table 18 shows the number of Tweets in each of the state-term strata in the full set of Tweets manually downloaded from Sysomos, as well as the breakdown by hit-or-not-hit preliminary classification. Because the hit strata are much smaller than the not-a-hit strata, sampling was weighted to include approximately 30 percent hits and 70 percent not-hits, with a minimum of two observations in the sample from each of the 18 strata. Stratum sample sizes appear in Table 18.

*** Table 18 about here ***

5.4 Image Classification

Images were often included as a part of the Tweet. Many of these Tweets with images often imply a complaint with the image. For instance, there are a large number of Tweets with images of crowded caucus locations or long polling lines. In order to classify these Tweets as hits, we uploaded the direct links to these images on Google's Reverse Image Search and classified the Tweet as a hit or not based on Google's guess of the subject of the image. Google's Reverse

Image Search uses a proprietary image identification system that utilizes a mix of pre-existing images online and the contents of webpages containing similar images. At the moment, for the primary/caucus data we have completed image classification of all images in our training set as well as a random sample of 500 Tweets in the non-training set data.

5.5 Flowchart Development

The primary/caucus hits flowchart (Figures 4 and 5) was developed over the course of several individual handcoding sessions. Tweets with three or more agreements as "hits" (among five coders) were designated core Tweets; a random sample of Tweets with two or fewer agreements as "hits" were reviewed and collectively discussed. After the discussion, we used both the core Tweets and the discussion of the marginal Tweets to create what we call the "hits flowchart." The flowchart was developed to standardize hits classification among the authors and avoided a simple definitional basis for classifying hits. The first half of the hits flowchart lays out what a hit is *not* (for instance, a hit is *not* an endorsement of a candidate); the second half of the hits flowchart engages with the substantive content of the Tweet and classifies the Tweet as a hit or not. This flowchart was used to create the training set, and coders currently use the flowchart to engage with the Tweets given by the active learning framework.

The general election hits flowchart (Figures 6, 7 and 8) reflects modifications to refer to all observed incidents without emphasizing "complaint" observations.

5.6 Coding Scheme for Primary/Caucus Tweets

Updated 8/21/16 (Version 5)

5.6.1 Instructions

After deciding whether the Tweet in question is a "hit" or not according to the flowchart, use the categories listed below to classify that hit. These categories and definitions also may help decide if a Tweet is a hit or not, if you are having trouble. A Tweet can be appropriately classified into multiple categories. For example, a Tweet that reads "For some reason there was a problem with my voter registration, but the workers at my polling place were very helpful!" would fall within the `registration problems` category and the `positive` category.

> 0 or blank: Tweet does not fit within this category
> 1: Tweet fits within this category

5.6.2 Categories for Coding

1. Absentee, Mail-In, or Provisional Ballot Issue: This category applies to hits relating to problems with absentee or mail-in ballots, including ballots not being received by the voter or ballots not being counted. This category also applies to incidents relating to provisional ballots, such as a voter having to vote provisionally (or not being allowed to). This category corresponds to the "provisional ballot abuse" and "Non-receipt of requested absentee ballots" EIRS category.

2. Other Ballot Problems: This category includes complaints or incidents regarding the design of the ballot, including layout and foldability. This category also applies to individuals being given the incorrect ballot, as well as a voter's preferred candidate or party not appearing on the ballot. This category corresponds to the other ballot problems EIRS category.

3. Disability/Accessibility Problem: Tweets that fall under this category would include complaints or observations about some aspect of the election that is not accessible for those with disabilities–for example, a polling place not offering special ballots or assistance to voters who are blind, or a polling place not being wheelchair accessible. This corresponds to the lack of disability access EIRS category.

4. Election Official Complaints/Incidents: Complaints that accuse governmental, election workers (including poll workers), or election officials of corruption, malfeasance, ignorance, being unhelpful or non-responsive, being rude, or some other complaint. This includes allegations of mis-managing the election. A Tweet that falls in this category and the positive category might not that a pollworker or election official was helpful, or the staff managed the polling place well. This category is analogous to the EIRS categories for 'pollworker malfeasance/ineptitude' as well as "other election worker problem."

5. Electoral System: This includes complaints relating to the specific aspects of the American electoral system, such as the first-past-the-post system, top-two electoral systems, caucuses, or open/closed primary elections. This category also includes complaints or hits that do not criticize a specific aspect of the American electoral system such as non-proportional representation. This also includes complaints about improper district boundaries and gerrymandering.

6. Improper Outside Influence: This category includes cases where the complainant encountered improper campaign advertising, such as advertising too close to a polling place. This category also includes complaints or observations alleging candidates', parties', or outside entities such as PACs' campaign practices violate the spirit or letter of the law. Also included in this category are allegations of police misconduct relating to the administration or outcome of the election, as well as complaints or incidents regarding the media. For example, an individual might complaint that the media called the election while people were still

in line to vote, or reporters may be improperly interviewing voters. This category is in part analogous to the "Improper Outside Influence" EIRS category.

7. Long Lines/Crowded Polling Place: This category refers to a complaint, incident, or report that states a long line or crowded polling place, including statements about the polling place being too small. Other examples of this category include a person referencing how long they have had to wait to vote, or reporting that their caucus has been moved outside due to crowding. This corresponds to the "polling place chaos and crowding" EIRS category.

8. Polling Place Problems: This category includes problems or incidents related to the polling place, such as the set-up of the voting booths and other election structures. Another example of a problem that would fit in this category is the presence of security cameras observing how individuals vote. Furthermore, this category includes voters being told an incorrect location for their polling place or precinct line. Finally, this category includes complaints or reports that allege intimidation by polling place officials or other persons (non-police) that occurred while the relevant person was casting his or her ballot, approaching the polling place, or in the polling place. This category does not include corruption, malfeasance, impropriety, or other comments regarding poll workers. It partially corresponds to the "Incorrect polling place/precinct information" and "Voter Intimidation" EIRS categories.

9. Registration Issues: Voters or prospective voters encountered difficulty registering to vote or had problems registering with their preferred party. It could also include instances of registration records being incorrect. This corresponds to "Incorrect registration lists/non-receipt of registration cards" EIRS category.

10. Voter Fraud: This category refers to instances or alleged instances of voter fraud, including a voter being told that he or she has already voted. This category is analogous to EIRS category "Voter fraud."

11. Voter Identification Issues: The voter or prospective voter had issues relating to voter identification requirements. This might include an election official improperly asking for identification, problems acquiring identification, or being rejected at the polls due to lack (or accused lack) of necessary identification. This corresponds to the "Improper ID requirements" EIRS category.

12. Voting Machine Complaints: This category includes voting machines being inoperable as well as unclear instructions regarding how to use the voting machines. Examples could include machines misreading scanned ballots, not printing receipts, or machines being difficult to use. This category is similar ot the "Machine malfunction/usage problem" EIRS category.

13. Unspecified Other: Includes complaints of which the nature is unclear as well as non-sequitur complaints. Analogous to the EIRS category "Other."

14. Positive: This category indicates that the complaint or incident was positive in nature: for example, complimenting an election official on being helpful, or there not being a long line to vote. In the latter case, it is appropriate to both mark the "Long Lines" category and the "Positive" category.

15. Ambiguous: This category notes that the wording of a Tweet or complaint is unclear and it is not possible to ascertain if it is complaint or hit. As such, it warrants further examination. For example, a Tweet might be worded such that it could be taken as a joke or as a serious comment on the election system, depending on the reader.

16. Not Hit: For the purposes of coding the machine-coded hits, mark this category if the Tweet in question is not a hit (that is, it was mistakenly defined as a hit by the machine classification algorithm).

5.7 Coding Scheme for General Election Tweets

Instructions
After you have decided whether the Tweet in question is a "hit" according to the flowchart, use the categories and subcategories listed below to classify that hit. These categories and definitions also may help you decide whether or not a Tweet is a hit, if you are having trouble. A Tweet can be appropriately classified into multiple categories. For example, a Tweet that reads "For some reason there was a problem with my voter registration, but the workers at my polling place were very helpful!" would fall within the "Election Official" and "Registration " categories, with Adjectives 2 and 0, respectively.

Once you have placed a Tweet in its appropriate category(s), you will also note which adjective applies to the Tweet. A Tweet stating "The line at my polling place was long" would be coded as a 2. So your task is both to place the Tweet in its appropriate categories as well as choosing the appropriate adjective that more specifically describes the content of the Tweet. These adjectives are either dichotomous (0 or 2) or trichotomous (0, 1, or 2)

Importantly, at this time we are not concerned with any sentiment or emotion contained within the Tweet. We are concerned with some statements that are evaluative or normative. We are concerned with describing the factual (or purported factual) event or item to which the Tweet refers.

Coding Scheme for Categorization

0 or blank: Tweet does not fit within this category
1: Tweet fits within this category

5.7.1 Categories for Coding

1. Outside Influence

This category includes cases where the complainant encountered improper campaign advertising, such as advertising to close to a polling place. This category also includes observations alleging the campaign practices of candidates, parties, or outside entities such as PACs violate the spirit or letter of the law. Also included in this category are allegations of police misconduct relating to the administration or outcome of the election, as well as complaints or incidents regarding the media. For example, an individual might complain that the media called the election while people were still in line to vote, or reporters may be improperly interviewing voters. This category is in part analogous to the "Improper Outside Influence" EIRS category.

Adjective: N/A.

2. Disability/Accessibility Issue

Tweets that fall under this category would include observations about some aspect of the election that is accessible or not accessible for those with disabilities—for example, a polling place not offering special ballots or assistance to voters who are blind, or a polling place not being wheelchair accessible. This relates to the lack of disability access EIRS category.

Adjective:

0: The aspect of the election was inaccessible

2: The aspect of the election was accessible.

3. Line Length, Waiting Time, Polling Place Crowding

This category refers to the length of a line or time to wait to vote or register, or to a crowded or empty polling place, including statements about the polling place being too small. Other examples of this category include a person referencing how long they have had to wait to vote, or reporting that their caucus has been moved outside due to crowding. This relates to the "polling place chaos and crowding" EIRS category.

Adjective:

0: There is no crowd or line at the polling place;

1: There was a small crowd or short line or wait;

2: The polling place was crowded or there was a long line or wait (20 minutes or longer).

4. Polling Place Event

This category includes incidents related to the polling place, such as the set-up of the voting booths and other election structures. Another example of a problem that would fit in this category is the presence of security cameras observing how individuals vote. Furthermore, this category includes a voter being told a correct or incorrect location for their polling place or precinct's line, or not knowing where to go to vote. Statements about the convenience of a polling place are included in this category. "I voted" statements referring to actions on election day are included in this category. Finally, this category includes complaints or reports that allege intimidation by persons other than polling place officials that occurred while the relevant person was casting his or her ballot, approaching the polling place, or in the polling place. This category does not include corruption, malfeasance, impropriety or other comments regarding poll workers. It partially relates to the "Incorrect polling place/precinct information" and "Voter Intimidation" EIRS categories.

Adjective:

0: The polling place did not function as expected or information is incorrect

1: The Tweet describes the polling place without noting whether it or an aspect functioned correctly or incorrectly

2: The polling place did function correctly or information is correct

5. Electoral System

This includes observations relating to specific aspects of the American electoral system, such as voluntary participation, the necessity to register to vote (e.g., registration deadlines), the first-past-the-post system, top-two electoral systems, caucuses, open/closed primary elections or non-proportional representation. This also includes comments about improper district boundaries and gerrymandering. Finally, comments about the integrity of the voting process due to hacking or hacking concerns are included here.

Adjective:

0: the electoral system did not function appropriately

1: the Tweet makes a neutral statement about the electoral system without an indication of if it functioned appropriately

2: the electoral system functioned appropriately

6. Absentee, Mail-In, or Provisional Ballot Issue

This category relates to features of absentee or mail-in ballots, including ballots being received or not being received by the voter, ballots being mailed or ballots not being counted. Early voting incidents are also included: "I voted" statements referring to actions during early voting are included in this category.. This category also applies to incidents relating to provisional ballots, such as a voter having to vote provisionally (or not being allowed to). This category relates to the "provisional ballot abuse" and "Non-receipt of requested absentee ballots" EIRS category.

Adjective:

0: the absentee, mail-in, or provisional ballot system did not function appropriately

1: the Tweet makes a neutral observation or statement about the absentee, mail-in, or provisional ballot system without noting it having functioned correctly or incorrectly

2: the absentee, mail-in, or provisional ballot system functioned correctly

7. Election Official

Comments that accuse governmental, election workers (including poll workers), or election officials of corruption, malfeasance, ignorance, being unhelpful or non-responsive, being rude, or some other complaint. This includes allegations of mismanaging the election. This category includes reports that allege intimidation by polling place officials that occurred while the relevant person was attempting to register, casting his or her ballot, approaching the polling place, or in the polling place. A Tweet that falls in this category might instead note that a pollworker or election official was helpful, or the staff managed the polling place well. This category is analogous to the EIRS categories for "pollworker malfeasance/ineptitude" as well as "other election worker problem."

Adjective:

0: The Tweet notes that the election officials did not perform their duties

1: the Tweet makes a neutral observation about election officials without noting them having performed or not performed their duties

2: the Tweet notes that the election officials performed their duties

8. Voter Identification

The voter or prospective voter had issues relating to voter identification requirements. This might include an election official improperly asking for identification, problems or no problems acquiring or using identification, or being rejected at the polls due to lack (or accused lack) of necessary identification. This relates to the "Improper ID requirements" EIRS category.

Adjective:

0: the Tweet notes that there were problems with the voter identification process or application

1: the Tweet makes a neutral observation about voter identification policies

2: the Tweet indicates that the voter identification process or application functioned appropriately

9. Registration

Voters or prospective voters encountered difficulty registering to vote, had problems registering with their preferred party or registered without difficulty. It could also include instances of registration records being incorrect, or positive or neutral statements about the registration process. This also includes an individual noting that he or she has been able to register. Also included is information about registration deadlines or processes. This relates to “Incorrect registration lists/non-receipt of registration cards” EIRS category.

Adjective:

0: The Tweet indicates that an individual was not able to register to vote

1: the Tweet makes a neutral observation about the voter registration process without noting if the individual in question registered or not

2: the Tweet notes that the individual was able to register to vote

10. Voter Fraud

This category refers to instances or alleged instances of voter fraud, including a voter being told that he or she has already voted. This also includes an individual noting that another individual has voted twice or is impersonating another eligible voter. This category is analogous to EIRS category “Voter fraud.” Need to update this language–will look at previous categories (EIRS, Germany)

Adjective:

0: The Tweet indicates that some form of voter fraud did occur

2: the Tweet indicates that some form of voter fraud did not occur

11. Ballot and Voting Technology

This category includes complaints or incidents regarding the design of the ballot, including layout and foldability, or the design or operation of voting technology. The category includes voting technologies working well or being inoperable as well as clear or unclear instructions regarding how to use the voting technology. Also included are observations about the security of the technology. Examples could include machines misreading scanned ballots, not printing receipts, or machines being difficult to use. This category also applies to individuals being given the incorrect ballot, as well as a voter’s preferred candidate or party not appearing on the ballot. This category relates to the “other ballot problems” and “Machine malfunction/usage problem” EIRS categories.

Adjective:

0: the ballot or voting technology was confusing or defective

1: the Tweet makes a neutral observation or statement about the ballot or voting technology without noting it having functioned correctly or incorrectly

2: the ballot or voting technology was well-designed or functioned correctly

12. Unspecified Other

Includes complaints of which the nature is unclear as well as non-sequitur complaints. Analogous to the EIRS category “Other.”

Adjective: N/A

References

Beber, Bernd and Alexandra Scacco. 2012. "What the Numbers Say: A Digit-Based Test for Election Fraud." *Political Analysis* 20(2):211–234.

Bjornlund, Eric C. 2004. *Beyond Free and Fair: Monitoring Elections and Building Democracy*. Washington, DC: Woodrow Wilson Center Press.

Bonzanini, Marco. 2015. "Mining Twitter Data with Python (Part 1: Collecting data)." `https://marcobonzanini.com/2015/03/02/mining-twitter-data-with-python-part-1/`. Accessed: 2016-06-15.

Cantu, Francisco and Sebastián M. Saiegh. 2011. "Fraudulent Democracy? An Analysis of Argentina's Infamous Decade Using Supervised Machine Learning." *Political Analysis* 19(4):409–433.

Dolinar, Sean. 2015. "Collecting Twitter Data: Using A Python Stream Listener." `http://stats.seandolinar.com/collecting-twitter-data-using-a-python-stream-listener/`. Accessed: 2016-06-15.

Hyde, Susan D. 2011. *The Pseudo-Democrat's Dilemma*. New York: Cornell University Press.

Ichino, Nahomi and Matthias Schuendeln. 2012. "Deterring or Displacing Electoral Irregularities? Spillover Effects of Observers in a Randomized Field Experiment in Ghana." *Journal of Politics* 74(1):292–307.

Kelley, Judith. 2012. *Monitoring Democracy: When International Election Observation Works and Why it Often Fails*. Princeton: Princeton University Press.

Lan, Man, Chew-Lim Tan, Hwee-Boon Low and Sam-Yuan Sung. 2005. A Comprehensive Comparative Study on Term Weighting Schemes for Text Categorization with Support Vector Machines. In *Special interest tracks and posters of the 14th international conference on World Wide Web*. ACM pp. 1032–1033.

Leopold, Edda and Jörg Kindermann. 2002. "Text Categorization with Support Vector Machines. How to Represent Texts in Input Space?" *Machine Learning* 46:423–444.

Levin, Inès, Gabe A. Cohn, Peter C. Ordeshook and R. Michael Alvarez. 2009. "Detecting

Voter Fraud in an Electronic Voting Context: An Analysis of the Unlimited Reelection Vote in Venezuela." 2009 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections, Montreal, August 10–11.

Mebane, Jr., Walter R. 2010. "Fraud in the 2009 Presidential Election in Iran?" *Chance* 23:6–15.

Mebane, Jr., Walter R. 2013. "Election Forensics: The Meanings of Precinct Vote Counts' Second Digits." Paper presented at the 2013 Summer Meeting of the Political Methodology Society, University of Virginia, July 18–20, 2013.

Mebane, Jr., Walter R. 2014. Can Votes Counts' Digits and Benford's Law Diagnose Elections? In *The Theory and Applications of Benford's Law*, ed. Steven J. Miller. Princeton: Princeton University Press pp. 206–216.

Mebane, Jr., Walter R. 2016. "Election Forensics: Frauds Tests and Observation-level Frauds Probabilities." Paper presented at the 2016 Annual Meeting of the Midwest Political Science Association, Chicago, April 7–10, 2016.

Mebane, Jr., Walter R. and Jonathan Wall. 2015. "Election Frauds, Postelection Legal Challenges and Geography in Mexico." Paper presented at the 2015 Annual Meeting of the American Political Science Association, San Francisco, CA, September 3–6.

Mebane, Jr., Walter R., Joseph Klaver and Blake Miller. 2016. "Frauds, Strategies and Complaints in Germany." Paper presented at the 2016 Annual Meeting of the Midwest Political Science Association, Chicago, April 7–10, 2016.

Miller, Blake and Joseph Klaver. 2016. "Active Learning for Costly Text Classification Problems and Exploratory Data Analysis." Unpublished paper draft.

Montgomery, Jacob M., Santiago Olivella, Joshua D. Potter and Brian F. Crisp. 2015. "An Informed Forensics Approach to Detecting Vote Irregularities." *Political Analysis* forthcoming.

Moujahid, Adil. 2014. "An Introduction to Text Mining using Streaming API and Python." `http://adilmoujahid.com/posts/2014/07/twitter-analytics/`. Accessed: 2016-06-15.

Myagkov, Mikhail, Peter C. Ordeshook and Dimitry Shaikin. 2009. *The Forensics of Election Fraud: With Applications to Russia and Ukraine*. New York: Cambridge University Press.

OpenStreetMap. 2017. "Nominatim." http://wiki.openstreetmap.org/wiki/Nominatim. Accessed: 2017-01-15.

Pancharian, Nick. 2016. "personal communication." phone conversation on August 22, 2016, with Nick Pancharian, Constituent Affairs, Secretary of State.

Pericchi, Luis Raúl and David Torres. 2011. "Quick Anomaly Detection by the Newcomb-Benford Law, with Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela." *Statistical Science* 26(4):502–516.

Plummer, Mary. 2016. "Numbers are in—poll workers top list of California voter complaints." *89.2 KPCC* . July 12, Internet Edition.

Rakotomamonjy, Alain. 2003. "Variable Selection Using SVM-based Criteria." *Journal of Machine Learning Research* 3(Mar):1357–1370.

Saxton, Gregory. 2014. "Using Python to Grab Twitter User Data." http://social-metrics.org/twitter-user-data/. Accessed: 2016-06-15.

Secretary of State, Constituent Affairs. 2015. "RE: HAVA Complaint Data Availability." Email message received November 19, 2015.

Secretary of State, Constituent Affairs. 2016. "Data Relating to the California Voter Hotline." File CA Election Day Complaints Report _060716.xls, August 9, 2016.

Settles, Burr. 2010. "Active learning literature survey." from Computer Sciences Technical Report 1648 http://burrsettles.com/pub/settles.activelearning.pdf.

Sysomos. 2016. "Social Technology: Turning Insights Into Action . . . Before Anyone Else." URL https://sysomos.com/.

Twitter, Inc. 2016*a*. "Documentation." URL https://dev.twitter.com/overview/documentation.

Twitter, Inc. 2016*b*. "REST APIs." URL https://dev.twitter.com/rest/public.

Twitter, Inc. 2017. "Twitter Development Documentation: Places." https://dev.twitter.com/overview/api/places. Accessed: 2017-01-15.

United States Census Bureau. 2016. "State Population Totals Tables: 2010-2016." URL https://www.census.gov/data/tables/2016/demo/popest/state-total.html, accessed March 12, 2016.

Table 1: Election Complaint Modes and Online Availability in Each State

		Modes						Modes			
State	Initial Contact[a]	HAVA Form	Other Form	Online Portal	Complaints Online	State	Initial Contact	HAVA Form	Other Form	Online Portal	Complaints Online
AL	7/21	Y	Y	Y		MT	5/13	Y[c]	Y[c]		Y
AK	7/21	Y			Y	NE	5/5	Y			
AZ	11/16[b]	Y				NV	4/29		Y		
AR	7/21		Y			NH	5/5	Y[c]	Y[c]		
CA	10/26[b]	Y	Y		Y	NJ	5/13				Y
CO	11/16[b]	Y	Y		Y	NM	5/2	Y[c]	Y[c]		
CT	11/16[b]		Y		Y	NY	5/9				
DE	7/21	Y			Y	NC	5/5	Y	Y		
FL	7/21	Y	Y		Y	ND	5/4	Y			
GA	7/21		Y	Y	Y	OH	5/13	Y	Y		Y
HI	5/5		Y			OK	5/4				
ID	3/24	Y				OR	3/24				
IL	5/10	Y	Y			PA	5/5	Y		Y	
IN	5/5	Y	Y			RI	5/5				
IA	5/4	Y	Y	Y		SC	5/9	Y			
KS	11/16[b]	Y	Y			SD	5/4				
KY	5/4	Y				TN	5/9	Y	Y		
LA	5/4	Y				TX	5/4	Y	Y		
ME	5/5	Y				UT	5/2				
MD	5/9	Y			Y	VT	5/9				
MA	5/5					VA	7/21	Y[c]	Y[c]	Y	
MI	5/13				Y	WA	3/24	Y		Y	Y
MN	5/13	Y			Y	WV	5/9	Y[c]	Y[c]		
MS	5/4	Y				WI	5/9			Y	
MO	5/13	Y	Y		Y	WY	5/4	Y[c]	Y[c]		

Note: An "Online Portal" is a dedicated online way to submit complaints; "Complaints Online" refers to an online repository of complaints available for viewing. [a] dates in 2016; [b] date in 2015; [c] dual HAVA and state forms.

Table 2: Number of HAVA Complaints from Each State

State	# HAVA Complaints
Illinois	4
Louisiana	1
New Hampshire	6
Ohio	1
Wyoming	4

Table 3: State Complaints Received

State	Years
Colorado	2013, 2014
Connecticut	1991, 1994, 2000, 2006, 2007–2016
Delaware	2009–2015
Florida	1993–2016
Indiana	2006–2007
Montana	1990–2016
New Hampshire	2012, 2013, 2014, 2015
New Jersey	1993, 1995–2015
New Mexico	2013–2016
Ohio	2000–2016
Washington	2011, 2012
Wyoming	2004, 2014

Table 4: State Voter Hotlines

Hotline	States
Yes	AZ, CA, CT, FL, GA, IL, IN, IA, KS, KY, LA, ME, MD, MS, MO, MT, NE, NJ, NC, TN, VT, VA
No	AL, AK, AR, CO, DE, HI, ID, MA, MI, MN, NV, NH, NM, NY, ND, OH, OK, OR, PA, RI, SC, SD, TX, UT, WA, WV, WI, WY

Table 5: Search Terms Used for Particular States with Sysomos

Arizona eleccion, spanishterms, eleccionprimaria, campaignfinance, corruption, azprimary, azvote, disenfranchised, election fraud, electionfraud, electionofficials, electionstealing, generalelection, linetovote, pollingplace, primaryelection, problemsvoting, righttovote, spanishtweets, voteaz, voterfraud, voterid, voteridentification, votingrights

California eleccion, spanishterms, eleccionprimaria, linetovote, caprimary, corruption, brokenvotingmachine, caelection, campaignfinance, cavote, disenfranchised, electioncomplaint, #electionfraud, electionfraud, electionstealing, generalelection, pollingplace, primaryelection, problemsvoting, righttovote, stoleelection, voteca, voterfraud, voterid, votingrights, caprimary, longlinetovote, caprimary, caprimaryANDNOTvote, caprimaryANDvote, caprimary, corruption

Colorado campaignfinance, caucus, corruption, disenfranchised, electionfraud, electionofficials, generalelection, linetovote, longlinetovote, pollingplace, primaryelection, probvoting, righttovote, spanishterms1, spanishterms2, statevote, voterfraud, voterid, voteridentification, votingrights

Connecticut campaignfinance, corruption, ctprimary, ctvote, disenfranchised, electionfraud, electionofficials, electionstealing, generalelection, linetovote, pollingplace, primaryelection, problemsvoting, righttovote, spanishtweets, votect, voterfraud, voterid, voteridentification, votingrights

Illinois campaignfinance, corruption, disenfranchised, electioncomplaint, electionfraud, electionofficials, ilprimary, ilvote, linetovote, longlinetovote, outofballots, pollingplace, primaryelection, problemsvoting, righttovote, stoleelection, twill, voteil, voterfraud, voterid, voteridentification, votersuppression, votingrights

Washington campaignfinance, corruption, disenfranchised, electioncomplaint, electionfraud, electionofficials, electionstealing, generalelection, longlinevote, pollingplace, primaryelection, problemvoting, righttovote, spanishterms, stolenelection, voterfraud, voteridentification, votingrights, wacaucus, wavoteetc

Note: examples of search terms used for a few states in searches using Sysomos within windows of ten days around each election/caucus day or election period (for states with absentee voting). “spanishterms” refers to a collection of election-related terms in Spanish.

Table 6: Search Terms Used to Cover All States with Sysomos

(akprimary, akprimary, akcaucus, akcaucus), (alprimary, to:alasecofstate),
(arprimary, to:Mark_Martin), to:ARSecofState, (azprimary, to:SecretaryReagan),
(caprimary+AND+vote, caprimary+AND+NOT+vote, to:CASOSvote),
(coprimary, cocaucus, to:ColoSecofState, to:juddchoate), (ctprimary, to:SOTSMerrill),
(dcprimary, dccaucus, to:DCBOEE, to:SecretaryofDC, dcprimary), (deprimary, to:SecretaryDE),
(flprimary+AND+vote, flprimary+AND+NOT+vote to:KenDetzner),
(gaprimary+AND+vote, gaprimary+AND+NOT+vote to:BrianKempGA),
(hicaucus, hiprimary, hicaucus, hiprimary),
(iacaucus, iaprimary, to:IowaSOS, to:PateforIowa), (idcaucus, idprimary, idprimary),
(ilprimary+AND+vote, ilprimary+AND+NOT+vote, to:ILSecOfState),
(inprimary, to:SecretaryLawson, to:IndianaSOS),
(kscaucus, ksprimary, to:BACaskey, to:KansasSOS),
(kycaucus, kyprimary, to:KySecofState, kyprimary, to:KySecofState),
(laprimary, to:Louisiana_sos), (maprimary, to:MrVoterReg),
(mdprimary, to:SOSMaryland, to:md_sbe), (mecaucus, meprimary, to:MaineSecOfState),
(miprimary+AND+vote, miprimary+AND+NOT+vote to:MichSoS, to:RJ4MI),
(mncaucus, mnprimary, to:MNSteveSimon, to:MNSecofState),
(moprimary, to:JasonKander, to:MissouriSOS),
(msprimary, to:DelbertHosemann, to:MississippiSOS), (mtprimary, to:SOSMcCulloch),
(ncprimary+AND+vote, ncprimary+AND+NOT+vote, to:Elaine4NC, to:NCSBE),
(ndcaucus, ndprimary, to:VoteND), (necaucus, neprimary, to:NESecJGale), nhprimary,
(njprimary, to:KimGuadognoNJ), nmprimary,
(nvcaucus, nvprimary, to:NVElect, to:NVSOS),
(nyprimary+AND+vote, nyprimary+AND+NOT+vote, to:NYSDOS, to:NYSBOE),
(ohprimary+AND+vote, ohprimary+AND+NOT+vote, to:JonHusted, to:OhioSOSHusted),
(okprimary, to:OKelections), (orprimary, to:oregonelections, to:OregonSoS),
(paprimary+AND+vote, paprimary+AND+NOT+vote, to:PAStateDept),
(riprimary, to:RI_BOE, to:RISecState, to:NellieGorbea),
(scprimary, to:scvotes), (sdprimary, to:shantelkrebs), (tnprimary, to:SecTreHargett),
(txprimary+AND+vote, txprimary+AND+NOT+vote, to:VoteTexas, to:TXsecofstate),
(utcaucus, utprimary, to:ElectionsUtah),
(vaprimary, to:vaELECT), (vtprimary, to:VermontSOS),
(waprimary, wacaucus, to:secstatewa), (wiprimary, to:Wisconsin_GAB, to:DougLaFollette),
(wvprimary, to:NatalieTennant),
(wycaucus, wyprimary, wycaucus, wyprimary, to:EdMurrayforWyo)

Note: search terms used in searches using Sysomos within windows of ten days around each election/caucus day or election period (for states with absentee voting). Parentheses group terms bearing on the same state.

Table 7: Number of Primary/Caucus Tweets by State

State	Count	State	Count	State	Count
Arizona	11,212	Kansas	8,011	Nevada	16,330
California	60,350	Kentucky	4,472	New York	17,155
Colorado	10,187	Louisiana	2,464	Ohio	18,866
Connecticut	4,561	Massachusetts	9,583	Oklahoma	820
Washington	15,599	Maryland	5,105	Oregon	3,592
Illinois	19,252	Maine	1,291	Pennsylvania	7,027
Alaska	5,281	Michigan	4,362	Rhode Island	1,267
Alabama	4,208	Minnesota	4,508	South Carolina	9,251
Arkansas	5,610	Missouri	5,459	South Dakota	594
DC	5,371	Mississippi	12,140	Tennessee	1,565
Delaware	1,114	Montana	300	Texas	12,871
Florida	9,782	North Carolina	4,776	Utah	1,108
Georgia	2,590	North Dakota	16,758	Virginia	1,327
Hawaii	25,256	Nebraska	1,237	Vermont	4,021
Iowa	21,520	New Hampshire	12,419	Wisconsin	8,564
Idaho	846	New Jersey	1,716	West Virginia	11,520
Indiana	16,754	New Mexico	2,564	Wyoming	575

Note: Number of unique Tweet texts (excluding retweets) by State obtained via Sysomos for the primary/caucus period.

Table 8: Number of General Election Incident Observations in Training Sample by State

State	Count	State	Count	State	Count
Alabama	14	Kentucky	15	North Dakota	3
Alaska	3	Louisiana	23	Ohio	69
Arizona	37	Maine	4	Oklahoma	24
Arkansas	6	Maryland	39	Oregon	15
California	148	Massachusetts	54	Pennsylvania	46
Colorado	17	Michigan	33	Rhode Island	3
Connecticut	8	Minnesota	30	South Carolina	21
Delaware	6	Mississippi	6	South Dakota	2
District of Columbia	43	Missouri	44	Tennessee	40
Florida	88	Montana	3	Texas	163
Georgia	60	Nebraska	7	Utah	9
Hawaii	1	Nevada	33	Vermont	1
Idaho	3	New Hampshire	4	Virginia	45
Illinois	61	New Jersey	19	Washington	14
Indiana	51	New Mexico	4	West Virginia	6
Iowa	7	New York	113	Wisconsin	28
Kansas	8	North Carolina	107	Wyoming	1
				Virgin Islands	1

Note: Number of Tweets observing incidents in human-labeled training sample (n = 1591) by State obtained via Twitter Streaming API for the general election period.

Table 9: Primary/Caucus Machine Classifier (Support Vector Machine) Performance

Class	**Precision**	**Recall**	**F-Measure**	**Support**
Not a hit	.82	.92	.87	1065
Hit	.66	.43	.52	380
Average/Total	.78	.79	.77	1445

Table 10: "Hit" Classification of Primary/Caucus Tweets by State

Unique Tweet Texts[a]						All Tweets[b]					
	Hit?			Hit?			Hit?			Hit?	
State	no	yes	State	no	yes	State	no	yes	State	no	yes
AK	395	162	MT	851	37	AK	841	175	MT	1,186	44
AL	1,217	9	NC	7,210	204	AL	1,892	29	NC	9,522	301
AR	1,050	55	ND	349	11	AR	1,385	65	ND	449	12
AZ	4,953	3,756	NE	2,201	111	AZ	6,613	4,065	NE	3,151	143
CA	33,190	11,903	NH	4,543	573	CA	44,463	13,064	NH	6,372	713
CO	4,308	3,916	NJ	3,796	41	CO	5,531	4,087	NJ	4,789	65
CT	2,663	677	NM	1,753	52	CT	3,563	796	NM	2,287	78
DC	8,858	229	NV	3,186	832	DC	11,476	309	NV	4,286	858
DE	578	19	NY	29,715	2,017	DE	795	29	NY	39,756	2,285
FL	13,387	440	OH	8,277	397	FL	20,969	594	OH	10,822	450
GA	5,672	143	OK	1,297	53	GA	8,090	225	OK	1,735	72
HI	1,348	256	OR	2,199	72	HI	1,910	319	OR	2,967	100
IA	15,469	2,543	PA	12,156	863	IA	20,253	3,113	PA	16,349	1,179
ID	1,966	149	RI	1,529	117	ID	2,392	191	RI	2,476	149
IL	10,902	3,366	SC	11,298	741	IL	15,571	3,680	SC	16,611	981
IN	8,443	379	SD	199	1	IN	12,517	550	SD	231	1
KS	1,395	520	TN	3,442	107	KS	1,779	526	TN	4,604	132
KY	2,974	742	TX	15,651	525	KY	4,037	922	TX	22,853	682
LA	1,276	30	UT	1,003	109	LA	1,819	35	UT	1,293	130
MA	3,859	208	VA	3,781	219	MA	5,309	265	VA	4,992	269
MD	4,682	321	VT	395	39	MD	5,914	407	VT	548	47
ME	553	100	WA	8,826	3,244	ME	839	132	WA	11,294	3,602
MI	4,801	238	WI	10,281	1,090	MI	6,813	310	WI	14,653	1,363
MN	3,404	1,097	WV	1,860	229	MN	4,807	1,356	WV	2,941	274
MO	3,262	77	WY	484	13	MO	4,491	103	WY	652	15
MS	855	39				MS	1,358	52			

Note: Number of Tweets (excluding retweets) classified as "hits" by State. [a] Counts using the unique texts across all Tweets. [b] Counts using all unique (by 18-digit ID code) Tweets.

Table 11: “Hit” Classification of California Primary Tweets by County

	Hit?			Hit?	
County	no	yes	County	no	yes
Alameda	1,200	323	Placer	117	46
Alpine	2	2	Plumas	6	3
Amador	9	3	Riverside	372	97
Butte	84	20	Sacramento	939	287
Calaveras	2	0	San Benito	22	6
Colusa	2	1	San Bernardino	352	99
Contra Costa	340	106	San Diego	2,781	674
Del Norte	2	1	San Francisco	3,926	873
El Dorado	97	9	San Joaquin	149	43
Fresno	228	52	San Luis Obispo	111	61
Glenn	15	3	San Mateo	223	37
Humboldt	103	23	Santa Barbara	179	38
Imperial	28	6	Santa Clara	660	179
Inyo	0	1	Santa Cruz	110	42
Kern	112	24	Shasta	49	18
Kings	3	0	Siskiyou	10	2
Lake	31	15	Solano	58	12
Lassen	5	2	Sonoma	184	31
Los Angeles	12,035	3,271	Stanislaus	91	19
Madera	13	8	Sutter	7	5
Marin	191	21	Tehama	20	6
Mariposa	2	0	Trinity	2	1
Mendocino	16	2	Tulare	56	15
Merced	27	6	Tuolumne	8	3
Modoc	1	0	Ventura	191	48
Mono	0	1	Yolo	219	65
Monterey	103	41	Yuba	1	0
Napa	22	3	Bay Area	265	131
Nevada	16	2	Silicon Valley	131	28
Orange	981	225			

Note: Number of Tweets (excluding retweets) classified as “hits” by county in California. Counts using the unique texts across all Tweets. “Bay Area” and “Silicon Valley” locations, which span multiple counties, are also shown.

Table 12: Frequency of Incidents by Type in Sample of California Primary "Hits"

Type	Unique Tweet Texts[a] Count	Percent	All Tweets[b] Count	Percent
Absentee Mail-in or Provisional Ballot Issue	44	7.3	49	12.7
Ballot Problems	10	1.7	12	3.1
Disability/Accessibility	0	0.0	0	0.0
Election Official Complaints/Incidents	18	3.0	20	5.2
Electoral System	38	6.3	40	10.4
Improper Outside Influence	51	8.5	54	14.0
Long Lines/Crowded Polling Place	43	7.2	47	12.2
Polling Place Problems	64	10.7	66	17.1
Registration Issues	14	2.3	19	4.9
Voter Fraud	31	5.2	36	9.3
Voter Identification Issues	10	1.7	11	2.8
Voting Machine Complaints	4	.7	4	1.0
Unspecified Other	4	.7	4	1.0
Positive	34	5.7	37	9.6
Ambiguous	29	4.8	33	8.5
Not an Incident	249	—	314	—

Note: "Count" shows the number of sampled Tweets that are of the indicated type, and "Percent" shows the percentage of the 351 Tweet texts (or 386 Tweets) that refer to an incident that are of the indicated type.
The sample (n = 600) is of Tweets drawn from all California Tweets classified as "hits" either by a human or by a machine classification algorithm (n = 15 Tweets in the sample are human-coded as incidents). All Tweets are associated with California based on "California" (or a synonym) being included in search terms or by the `location` in the Tweet mentioning a place in California. Coding by type is performed directly by humans.
[a] Counts using a sample of the unique texts in California Tweets (n = 600). [b] Counts using all replicas of sampled Tweet texts in California (n = 700).

Table 13: California Primary Election-day Incidents by Type, Tweet Sample

Type	Unique Tweet Texts[a] Count	Unique Tweet Texts[a] Percent	Omit Positive[b] Count	Omit Positive[b] Percent
Absentee Mail-in or Provisional Ballot Issue	19	15.7	19	18.4
Ballot Problems	1	0.8	1	1.0
Disability/Accessibility	0	0.0	0	0.0
Election Official Complaints/Incidents	12	9.9	12	11.7
Electoral System	8	6.6	8	7.8
Improper Outside Influence	14	11.6	14	13.6
Long Lines/Crowded Polling Place	18	14.9	14	13.6
Polling Place Problems	36	29.8	34	33.0
Registration Issues	5	4.1	5	4.9
Voter Fraud	2	1.7	2	1.9
Voter Identification Issues	0	0.0	0	0.0
Voting Machine Complaints	1	0.8	1	1.0
Unspecified Other	3	2.5	3	2.9
Positive	18	14.9	—	—
Ambiguous	2	1.7	2	1.9
Not an Incident	83	—	83	—

Note: "Count" shows the number of sampled Tweets that are of the indicated type, and "Percent" shows the percentage of the 351 Tweets that refer to an incident that are of the indicated type. The sample (n = 600) is of Tweets drawn from all California Tweets classified as "hits" either by a human or by a machine classification algorithm. This table shows only the subsample of election-day Tweets (n = 121). All Tweets are associated with California based on "California" (or a synonym) being included in search terms or by the `location` in the Tweet mentioning a place in California. Coding by type is performed directly by humans. [a] Counts using a sample of the unique texts in California Tweets (n = 121). [b] Omitting "positive" Tweets.

Table 14: California Primary Election-day Hotline Complaints

Type	Count	Percent
Closed Polling Place	36	6.3
Electioneering	6	1.1
ID Issue	9	1.6
Other	10	1.8
Poll Worker Problem	234	41.1
Polling Location	71	12.5
Provisional Voting	17	3.0
SOS Election Day Observer Allegation	2	0.4
Vote by Mail Ballot	9	1.6
Voter Registration	65	11.4
Voting Materials	23	4.0
Voting Process Issue	22	3.9
Voting System Equipment	66	11.6

Note: "Count" denotes the number of complaints of a given type submitted to the hotline. The category of complaint was determined on a case-by-case basis by the individual hotline operators. Source: Secretary of State, Constituent Affairs (2016)

Table 15: California Primary Election-day Hotline by County

County	Count	County	Count
Alameda	14	Sacramento	21
Butte	1	San Bernardino	14
Colusa	1	San Diego	25
Contra Costa	17	San Francisco	10
Fresno	7	San Joaquin	1
Humboldt	2	San Mateo	1
Imperial	1	Santa Barbara	1
Kern	3	Santa Clara	9
Kings	1	Santa Cruz	1
Los Angeles	367	Solano	5
Madera	1	Sonoma	3
Marin	1	Tulare	1
Mendocino	1	Tuolumne	1
Napa	1	Ventura	7
Orange	19	Yolo	5
Riverside	28		

Source: Secretary of State, Constituent Affairs (2016)

Table 16: "Hit" Classification of California Primary Election-Day Tweets by County

	Hit?			Hit?	
County	no	yes	County	no	yes
Alameda	211	112	Riverside	55	34
Alpine	0	2	Sacramento	213	82
Amador	3	1	San Benito	2	1
Butte	12	10	San Bernardino	99	38
Contra Costa	36	21	San Diego	700	210
Del Norte	1	0	San Francisco	721	305
El Dorado	8	2	San Joaquin	45	12
Fresno	57	15	San Luis Obispo	18	4
Glenn	6	1	San Mateo	37	14
Humboldt	9	2	Santa Barbara	56	10
Imperial	4	2	Santa Clara	120	65
Kern	18	12	Santa Cruz	34	13
Lake	8	7	Shasta	8	5
Lassen	4	0	Siskiyou	3	0
Los Angeles	2,313	1,355	Solano	16	4
Madera	0	4	Sonoma	59	10
Marin	35	6	Stanislaus	23	11
Mariposa	1	0	Sutter	1	0
Mendocino	4	2	Tehama	10	1
Merced	10	3	Trinity	1	1
Monterey	28	11	Tulare	12	1
Napa	3	0	Ventura	31	20
Nevada	2	0	Yolo	35	11
Orange	231	64	Bay Area	79	46
Placer	25	16	Silicon Valley	29	13
Plumas	2	0			

Note: Number of election-day Tweets (excluding retweets) classified as "hits" by county in California. Counts use the unique texts across all Tweets. "Bay Area" and "Silicon Valley" locations, which span multiple counties, are also shown.

Table 17: Per Capita General Election Incident Observations in Training Sample by State

State	Count	State	Count	State	Count
Alabama	2.88	Kentucky	3.38	North Dakota	3.96
Alaska	4.04	Louisiana	4.91	Ohio	5.94
Arizona	5.34	Maine	3.00	Oklahoma	6.12
Arkansas	2.01	Maryland	6.48	Oregon	3.66
California	3.77	Massachusetts	7.93	Pennsylvania	3.60
Colorado	3.07	Michigan	3.32	Rhode Island	2.84
Connecticut	2.24	Minnesota	5.43	South Carolina	4.23
Delaware	6.30	Mississippi	2.01	South Dakota	2.31
District of Columbia	63.13	Missouri	7.22	Tennessee	6.01
Florida	4.27	Montana	2.88	Texas	5.85
Georgia	5.82	Nebraska	3.67	Utah	2.95
Hawaii	0.70	Nevada	11.22	Vermont	1.60
Idaho	1.78	New Hampshire	3.00	Virginia	5.35
Illinois	4.77	New Jersey	2.12	Washington	1.92
Indiana	7.69	New Mexico	1.92	West Virginia	3.28
Iowa	2.23	New York	5.72	Wisconsin	4.85
Kansas	2.75	North Carolina	10.55	Wyoming	1.71

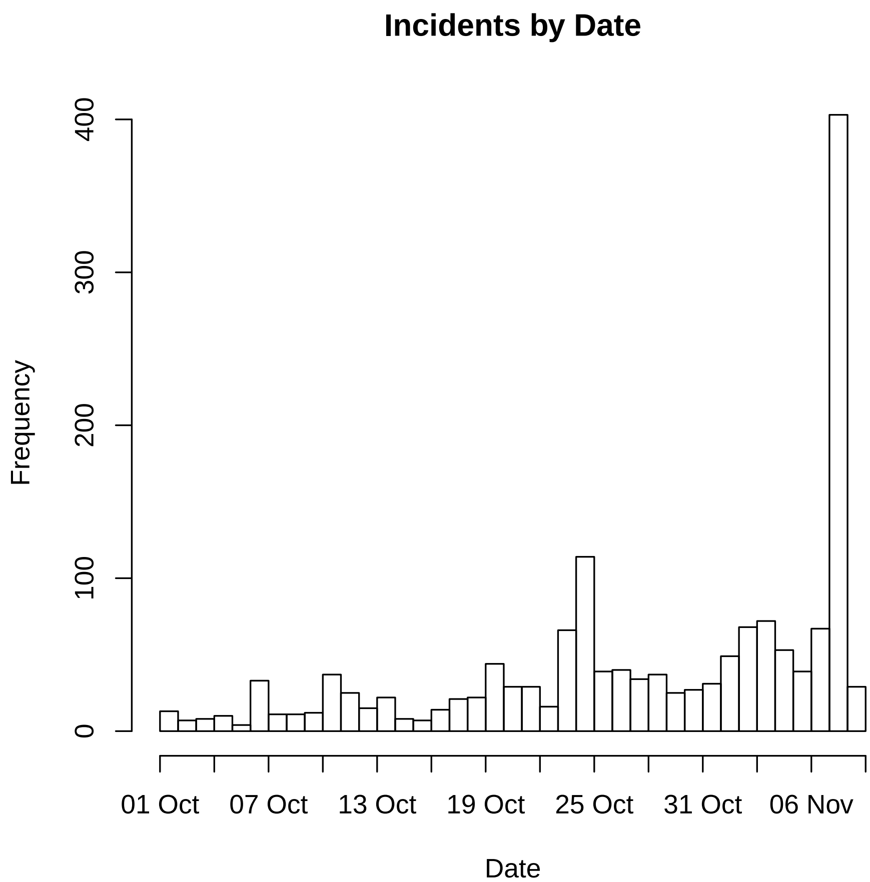
Note: Number per million persons of Tweets observing incidents in human-labeled training sample (n = 1591) by State obtained via Twitter Streaming API for the general election period. 2016 population data source: United States Census Bureau (2016).

Table 18: Strata for Sampling Tweets to use in Initial Primary/Caucus Training Set

		population		sample	
state+	total[a]	not hit	hit	not hit	hit
AZ	9,890	1,607	478	62	109
CA	52,296	10,774	271	414	62
CT	3,537	712	24	27	6
CO	8,388	1,511	261	58	59
WAd	10,062	1,958	169	75	39
WAr	2,910	608	7	23	2
CAeo	3,041	558	72	21	16
COeo	177	68	6	3	2
WAeo	505	105	3	4	2

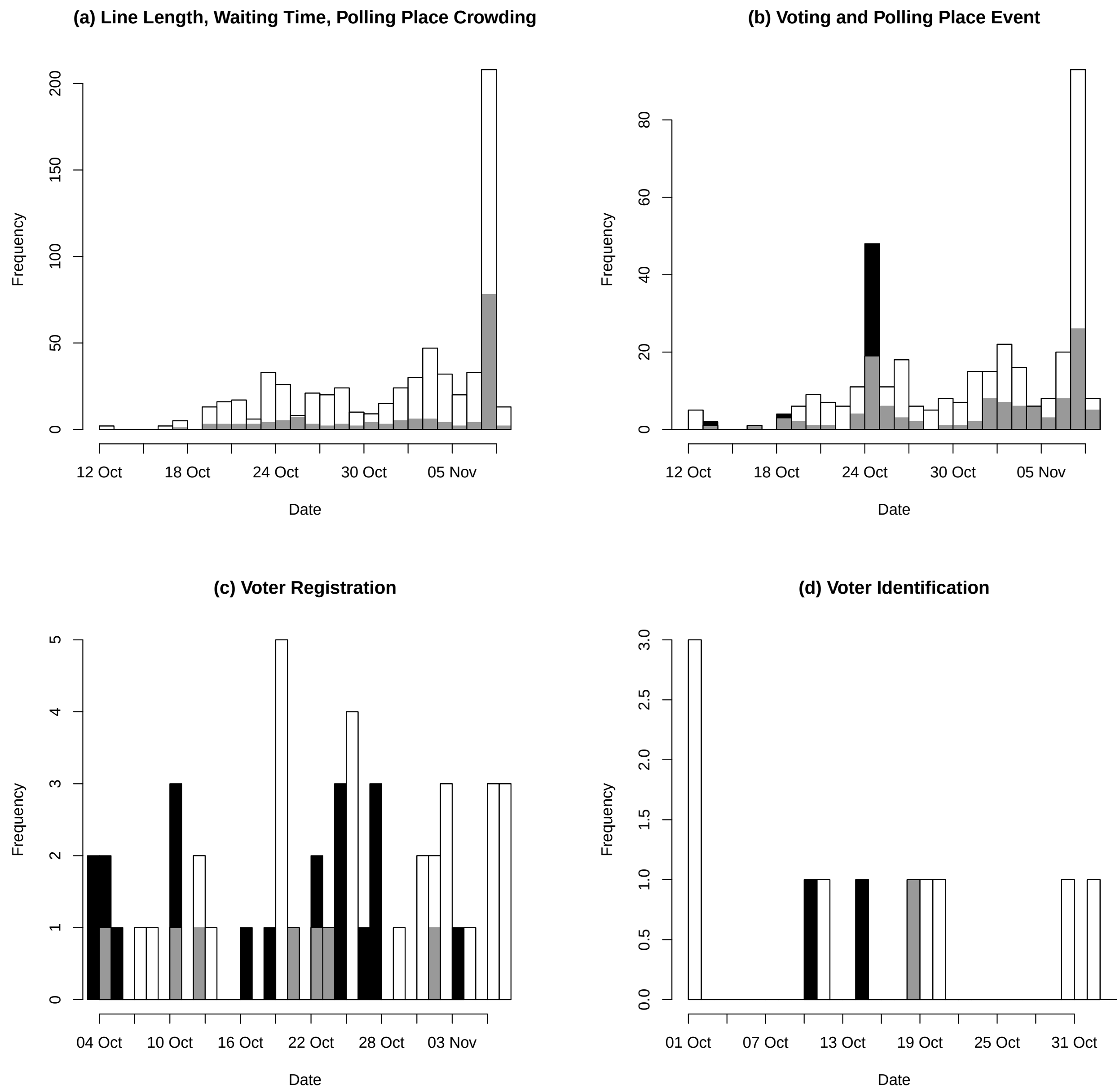
Note: [a] "total" values are the numbers of unique Tweets (no retweets) in each stratum in the set of Tweets manually downloaded using Sysomos.

Figure 1: General Election Incident Observations in Training Sample by Date



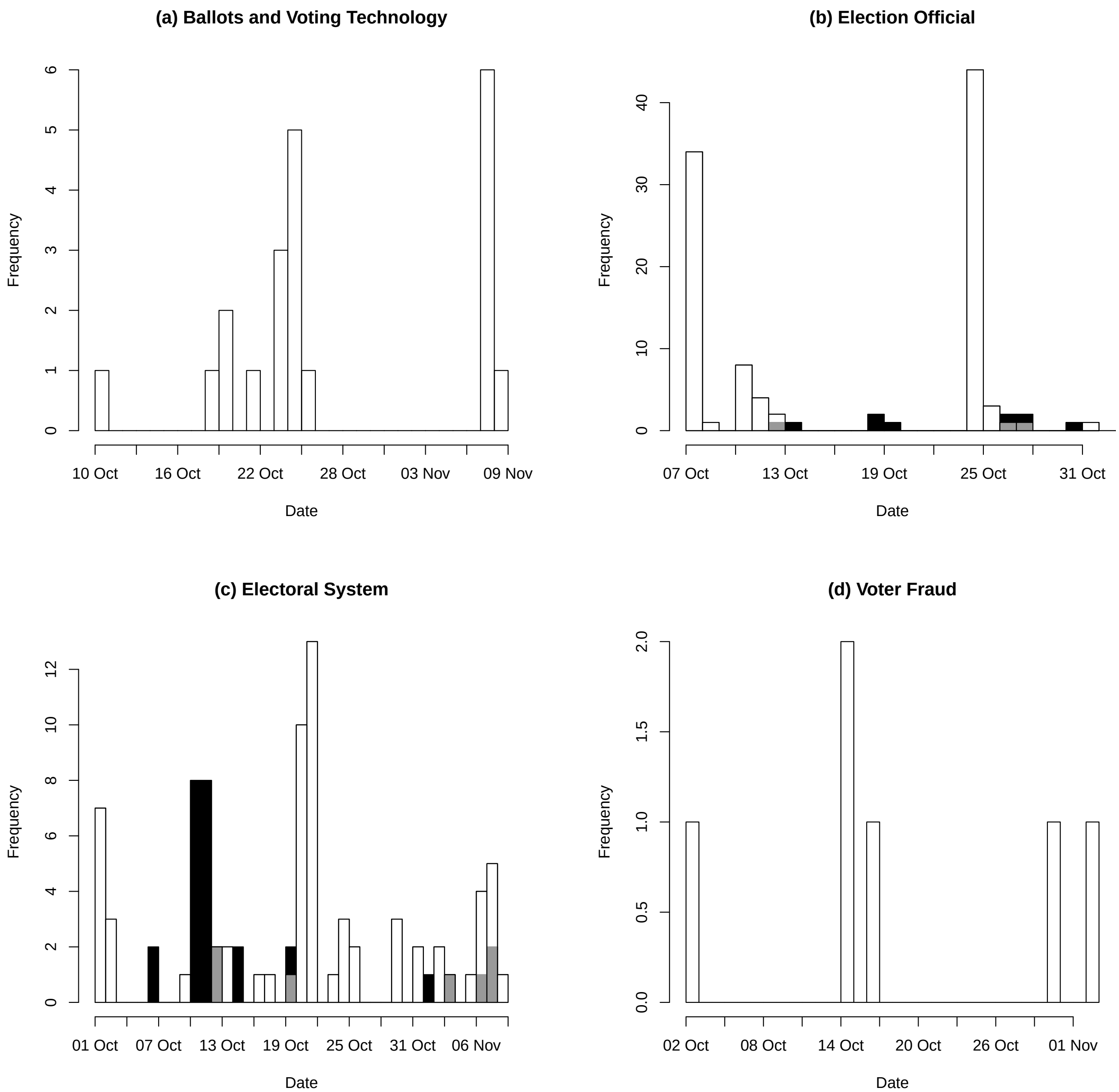
Note:

Figure 2: General Election Incident Observations in Training Sample by Type and Date



Note: (a) white, long line or wait; black, no wait/line or short wait/line. (b) white, in-person, early or absentee voting success; black, voting problem. (c) white, voter registration problem; black, registration success; (d) white, voter identification problem; black, voter identification success.

Figure 3: General Election Incident Observations in Training Sample by Type and Date



Note: (a) ballots and voting technology problem. (b) white, election official problem; black, election official sucess. (c) white, election system problem; black, election system success. (d) voter fraud problem.

Figure 4: Primary/Caucus Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 1

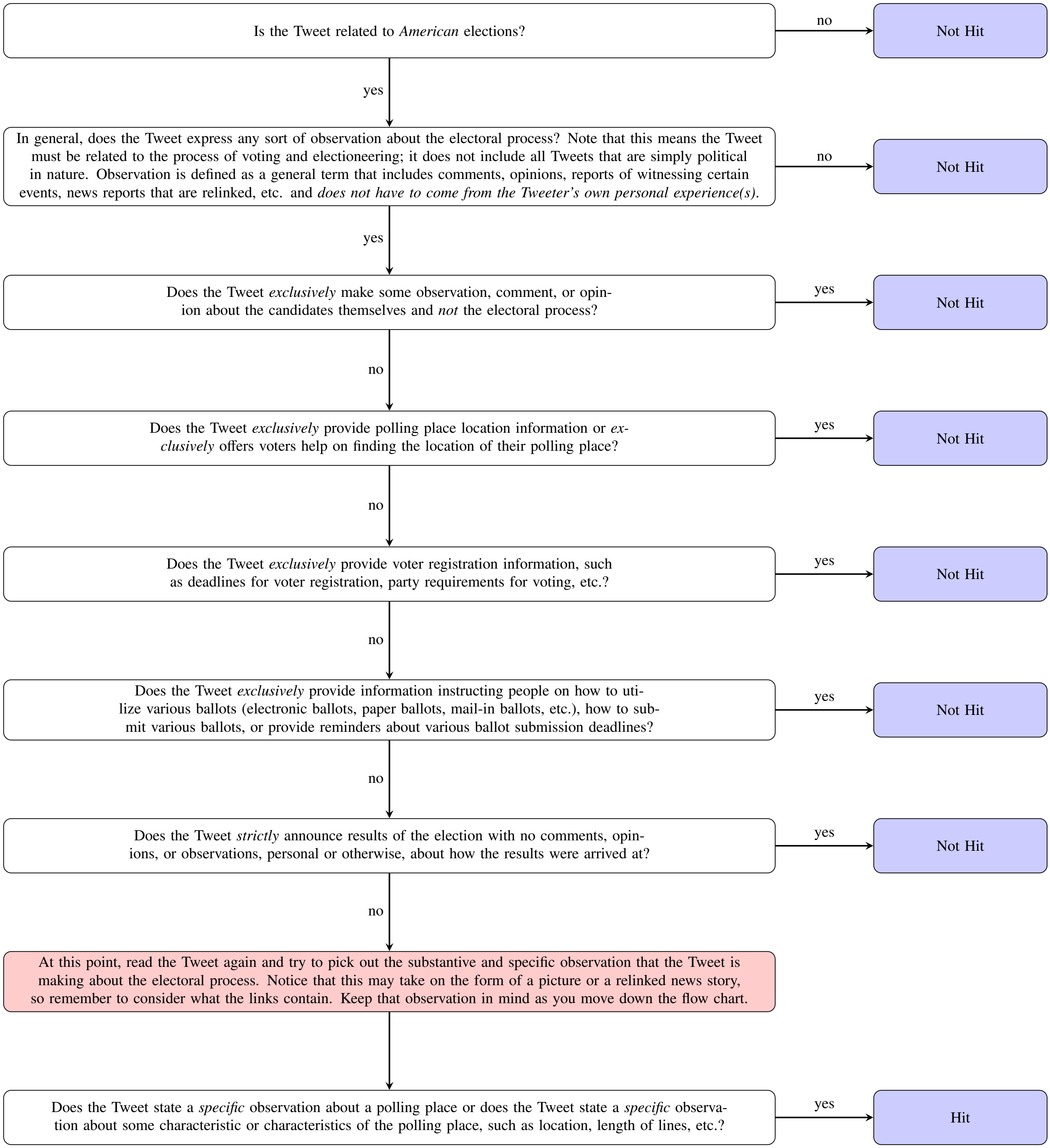


Figure 5: Primary/Caucus Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 2

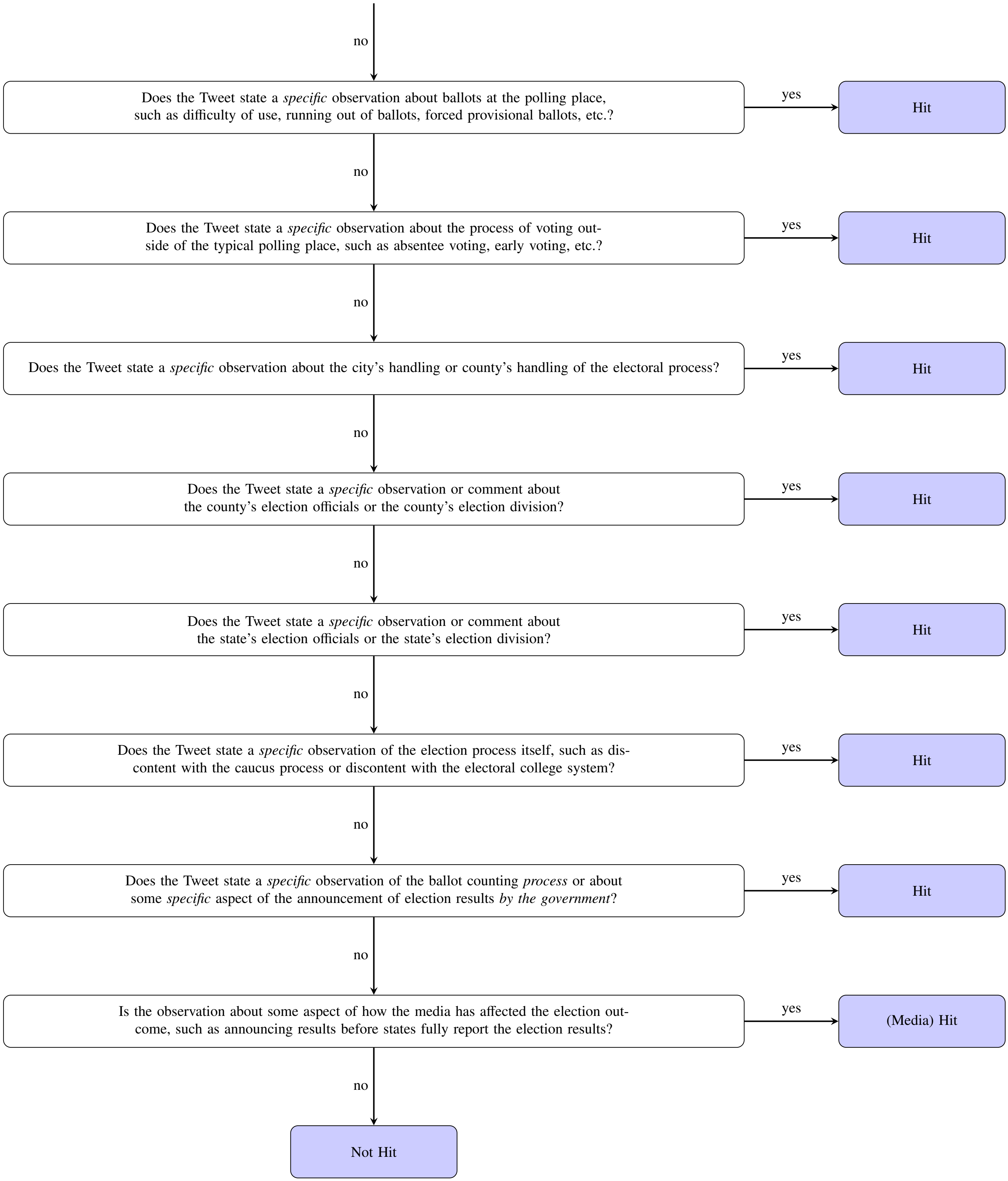


Figure 6: General Election Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 1

An **observation** is a statement that refers to a (probably) real situation with which the Tweeter (probably) had personal familiarity: either the person witnessed the situation or personally knew the person who did; in cases where there is ambiguity about the directness of the personal involvement, the observation report will be treated as if it were personal. So descriptions that are entirely about news reports are generally excluded, but if it's not clear that the item comes from a news report we'll include it.

- Personal involvement does not mean the observation refers to a personal experience: statements about collective situations such as the electoral system, voter registration procedures and electoral administration are also to be included.
- The observation may be embedded in an opinion, comment or advocacy statement, but advocacy statements per se are to be excluded. The observation may be adjacent to a news report that is relinked but news reports per se are to be excluded.
- Notice that an observation may take the form of an image, so remember to consider what any URLs contain. Keep that point in mind as you move down the flow chart.
- If the Tweet contains editorializing comments, be sure to identify the specific observation about the electoral process that the comments may be making.

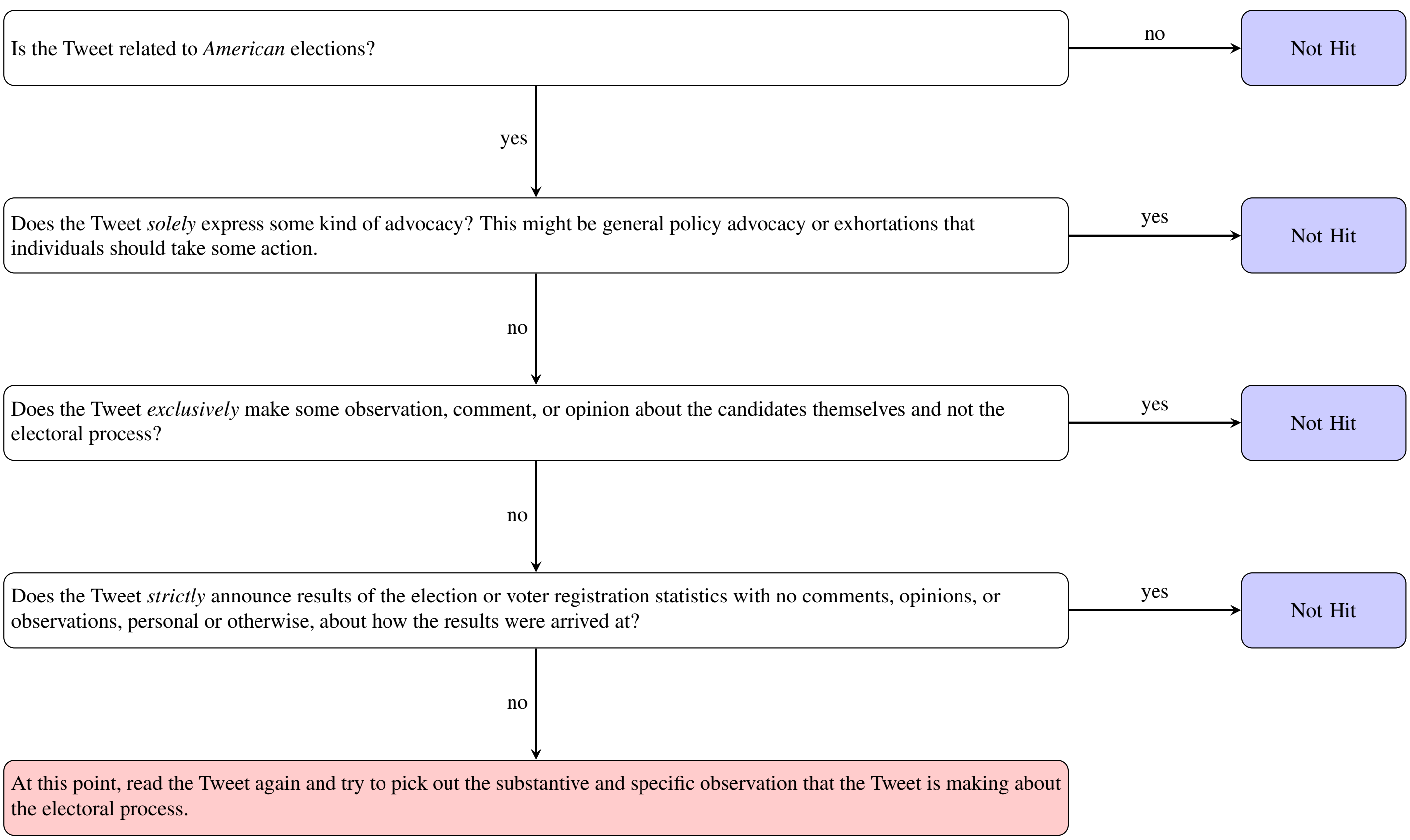


Figure 7: General Election Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 2

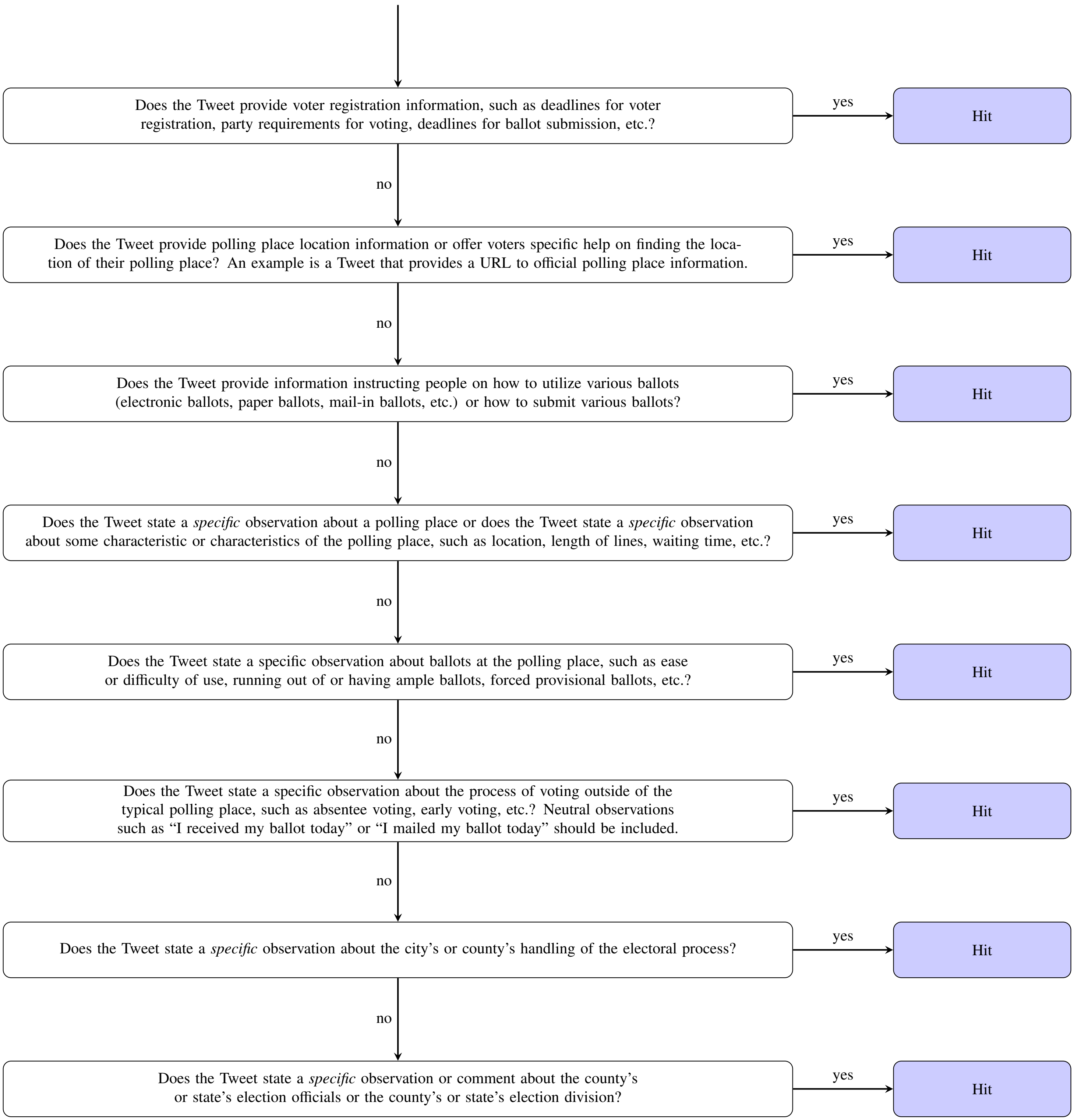


Figure 8: General Election Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 3