Using Twitter to Observe Election Incidents in the United States*

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

Joseph Klaver¶ Patrick Wu|| Blake Miller**

August 20, 2017

*Prepared for presentation at the 2017 Annual Meeting of the American Political Science Association, San Francisco, CA, August 31–September 3, 2017. Previous versions were presented at the 2017 Annual Meeting of the Midwest Political Science Association and 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 Preston Due, 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

Individuals' observations 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 observations cannot be gathered through official channels. 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 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. For the general election we develop hundreds of thousands of incident observations that occur at varying rates in different states, that vary over time and by type and that depend on state election and demographic conditions. Thousands of observations concern long lines, but even more celebrate successful performance of the election process—testimonies that ”I voted!” proliferate.

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; Rozenas 2017). 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).

Problems in elections that are not due to frauds may also stem from legal or administrative decisions. Long waiting times or crowded polling place conditions (Stewart and Ansolabehere 2015; Herron and Smith 2016), for example, are themselves concerns and may also produce distortions in turnout or vote choice data. As another example, the deployment of badly designed ballots (Lausen 2007; Quesenbery and Chen 2008) or defective election equipment (Herrnson, Niemi, Hanmer, Bederson, Conrad and Traugott 2008; Jones and Simons 2012) is inherently interesting and may also cause distortions in other election data. Another example is the number of polling stations opened for an election and where the polling stations are located (Brady and McNulty 2011).

Observing how individual people—voters or would-be voters—interact with such conditions is a challenge. In some countries systems for recording citizen complaints or the findings of observers are robust (e.g. Mebane and Wall 2015), but not for instance in the United States (Mebane, Pineda, Woods, Klaver, Wu and Miller 2017). Survey data cannot produce information with sufficient granularity to locate potential problems throughout an entire electoral system—at every polling station throughout the entirety of a multi-day election, for example. Either for further use in election forensics or because of their inherent interest as causes or consequences of political behavior, it can be useful to obtain observations that originate from ordinary individuals of how elections are administered and of how individuals respond to election circumstances.

We use data from Twitter to get information about American election administrative performance from individual observers throughout the country: the beginnings of a "Twitter Election Observatory." We describe election observations—extracted from Twitter—by individuals during the 2016 election from across the United States. While we do not address how to integrate these data in an election forensics analysis, we do show how various observed phenomena—such as individuals waiting in long lines or having difficulties in casting votes—are associated with state-level election procedures and demographic variables.

We describe both a preliminary complaint-oriented scheme focused on the presidential primary elections and caucuses held across the country in 2016 and a subsequent observation-focused scheme used with Tweets during the Fall general election. The system involves extracting Tweets using keyword filters, collecting information about election officials' and other leading actors' Twitter accounts, and classifying Tweets for relevance and for type of incident. For the classification tasks we apply active learning techniques with automated machine classification methods to Tweet texts, although both images and text 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 show that hundreds of thousands of incident observations can be recovered from Tweets gathered during the election period, observations that get at many different aspects of election performance. Incidents vary over states and over time, and they are 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.

Monitors, Observers and Official Complaints in the United States: Another potential source of data to supplement forensic 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. Mebane, Klaver and Miller (2016) and Mebane and Wall (2015) use such data, respectively from German citizens and from Mexican parties. In Germany data come from citizen complaints about the federal election filed with a committee of the *Bundestag*, and in Mexico 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.

For several reasons it is difficult to obtain information about citizens' observations of election incidents in the United States. Election complaint processes in most states are convoluted and characterized by multiple possible channels for disputes, and they usually depend on particular election laws allegedly being violated. These channels may 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 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 official channels publicly available. In Mebane et al. (2017) we detail the unavailability of official data about citizen complaints in the United States.

The impossibility of obtaining citizen observations of election incidents through such means for the United States prompts us to turn to social media. We find that voluminous observations can be gathered from Twitter. The biggest challenges with such data concern whether observations are reliable, whether the location of reported incidents can be determined and whether the observations we are able to collect accurately represent the full set of all incidents that occur.

2 Using Twitter to Capture Election Observations

We construct infrastructure to allow 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.

2.1 Collecting Twitter Data

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

We used Sysomos for the primary/caucus Tweets because Sysomos supports searching for Tweets using keywords for a period going back 12 months and we were downloading Tweets starting at the beginning of summer, 2016.[1] 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 1. 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 4.1 for details regarding compiling the list).[2] 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 2. Finally for all states we ran Sysomos searches using the keywords listed in the note.[3]

*** Tables 1 and 2 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

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

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

[3] 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.

Tweets using the API are in Appendix section 4.1.

Table 3 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 3 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.[4] In all during October 1 through November 8, 2016, we downloaded 44,329 Tweets from timelines and 16,221,304 Tweets via the Streaming API. Removing retweets leaves 6,163,890 unique Tweets which contain 4,541,097 unique Tweet texts. Only 598,783 Tweets have `place` and `fullname` information (see Appendix 4.1), 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. Among these Tweets there are 505,112 unique Tweet texts. We drew a sample of 19,789 Tweet texts from this collection of 505,112 Tweets and labeled them by hand as containing an incident observation ($n = 2,610$) or not ($n = 17,179$). This is the initial sample of human-labeled Tweets we use to begin the active learning process described in section 2.2.1.

Table 4 reports the distribution of the initially sampled incident observations over states. "States" includes Puerto Rico (PR) and the United States Virgin Islands (VI). All states are

[4]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, 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.

covered, and in general but not always states that are larger in population have more Tweets with incident observations.

*** Table 4 about here ***

2.2 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.

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 and append it to the original content.[5] If that resource contains an image, we capture the image's URL.[6] Human coders examine any images associated with a Tweet when labeling it, but currently the machine learning algorithms we use use only the augmented text. Images often decisively affect human coders' judgments regarding any information Tweets may contain—e.g. an image of a person wearing an "I Voted" sticker or an image of many people in line at a polling place—but the machine classification algorithm currently does not have access to images or descriptions of images.

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.

[5]Specifically, we capture any text in the `og:description` field in the resource's HTML code. For general election type-of-incident classification we also append to the text the date (month, day and year) and `place$fullname` of the Tweet.

[6]Specifically, we capture any URL in the `og:image` field in the resource's HTML code.

To determine whether each downloaded Tweet includes relevant observations, we began by using humans[7] 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 (see Appendix section 4.4) derive from the incident type categories in Mebane, Klaver and Miller (2016) and the Election Incident Reporting System (EIRS) (Verified Voting Foundation 2005; Hall 2005; Johnson 2005). Through several rounds of coding, discussion and recoding of random samples of Tweets from Arizona, California, Colorado, Connecticut and Washington[8] 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 (see Appendix section 4.5) are 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, and the procedure for the general election data is shown in Figures 6, 7 and 8. The background for these flowcharts is discussed in Appendix section 4.3. The coding rules for categorizing the incidents to which "hits" refer are described in Appendix section 4.4 (for primary/caucus Tweets) and in Appendix section 4.5 (for general election Tweets).

*** Figures 4, 5, 6, 7 and 8 about here ***

As detailed in Appendix section 4.4, 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

[7]The human coders were subsets of this paper's authors and two undergraduate assistants.

[8]In Washington Tweets come from both the Democratic caucuses and the Republican primary elections.

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."

As detailed in Appendix section 4.5, 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 4.5 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 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.[9]

2.2.1 Active Learning

To produce a training set to use to start active learning with the primary/caucus Tweets, we used a stratified random sample[10] 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."

To produce a training set to use to start active learning with the general election Tweets, we used a sample of 19,179 Tweets from the streaming API. For a description of how the sample was

[9]We plan to recode the primary/caucus Tweets using the general election scheme.
[10]For a description of the sample see Appendix section 4.2.

drawn see Appendix section 4.2. The ”hits” in this sample were initially produced by several coders but then all were checked by one pair of coders working in tandem.

To grow the initially small training sets we use active learning, an iterative supervised machine learning technique (Settles 2010). Active learning allows us to build training sets with fewer labeled observations and a good balance between classes, which is useful because of the scarcity of the some classes (Miller and Klaver 2016). 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 support vector machine (SVM) on labeled Tweet texts. We use the distance from the SVM's separating hyperplane to measure model uncertainty. We iteratively label the texts closest to the hyperplane and refit a model until acceptable average precision, recall and F-measure are achieved.

2.2.2 Classification

For both the active learning SVM and the algorithms we use for the final classification step[11] we first preprocess each Tweet's augmented text. This involves removal of all duplicate texts. In the primary/caucus data we use stemming and stop word removal but in the general election data we do not. For classification we use a word n-gram model for the preprocessed text and a character n-gram model for hashtags to convert the Tweet corpus into a document term matrix.[12] Each row of the matrix represents a Tweet and each column represents a unique character or word n-gram. Cell values represent the count of each n-gram 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). For the

[11]The classification algorithms we use from `sklearn` (Pedregosa, Varoquaux, Gramfort, Michel, Thirion, Grisel, Blondel, Prettenhofer, Weiss, Dubourg, Vanderplas, Passos, Cournapeau, Brucher, Perrot and Duchesnay 2011) are `linear_model.LogisticRegression`, `naive_bayes.MultinomialNB` and `svm.LinearSVC` as estimators in `ensemble.VotingClassifier`.

[12]We allow up to 5-grams for words and 2-, 3- and 4-grams for characters in hashtags.

final classification step we use a randomized search to select parameters for the various algorithms.[13]

For the primary/caucus data humans 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 we classify 43,169 texts as "hits" and 277,941 as "not-hits." Classification performance measures, based on a weighted cross-validation method,[14] are shown in Table 5. Overall we achieve average precision, recall and F-measure of .78, .79 and .77, respectively.

*** Table 5 about here ***

For general election "hit" labeling humans manually labeled 5,224 Tweet texts with `place` information selected in active learning, for a total of 25,013 human-labeled Tweets. Among the human-labeled texts, 3,689 are "hits" and 21,324 are "not-hits." Over all unique Tweet texts with `place` information we classify 40,687 texts as "hits" and 464,425 as "not-hits." Over all unique Tweet texts with or without `place` information we classify 315,180 texts as "hits" and 4,225,917 as "not-hits."[15] Classification performance measures (Table 6) for the set of Tweets that have `place` information show that overall we achieve average precision, recall and F-measure of .88, .89 and .88, respectively.[16] Notice that classification performance is assessed as similar when done both without stemming and with stemming. Indeed, every Tweet with `place` information is classified identically in both cases, even though algorithm parameters vary when stemming is enabled.[17]

*** Table 6 about here ***

[13]To execute the search we use `RandomizedSearchCV` from `sklearn` (Pedregosa et al. 2011).

[14]We use `model_selection.train_test_split` in `sklearn` (Pedregosa et al. 2011). 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 approximately the same as the number of "hits."

[15]Before classifying all 4,541,097 Tweet texts regardless of whether a Tweet has `place` information, we use active learning to human-label an additional 100 Tweets from the pooled corpus of all 4,541,097 Tweet texts.

[16]Results for the larger set of Tweets, which includes 100 more human-labeled Tweets, are nearly the same.

[17]For instance, without stemming the randomized search finds for words it is best to use up to 3-grams while with stemming it is best to use up to 5-grams.

2.2.3 General-election Type Coding

To determine what type of incident is represented by each of the 40,687 general election Tweets texts with `place` information that are classified as "hits," we begin by manually labeling a random sample of 1,419 of the texts then augment the initial sample using binarized active learning. While each Tweet may mention several types of incidents, the distribution of individual types of incidents in this initial sample is shown in Table 6. A few types are scant, and some possible "adjectives" do not occur in the initial sample. To try to boost a few of the type frequencies before beginning machine-assisted sampling, we hand-labeled a few Tweets located by doing keyword searches in the set of 40,687 Tweet texts.[18]

*** Table 7 about here ***

For binarized active learning we use the SVM approach we used for "hits" for each type and each type adjective separately. For instance one step of the process includes Tweet texts in the sample for human labeling if they are near the separating hyperplane for the "Polling Place Event" incident versus all other types of incidents. Samples are weighted using the inverse relative frequency of occurrence among the human-labeled texts, so that texts that are uncertain members of less frequent classes are sampled more frequently. Types or adjectives that occur too infrequently are not used to determine sampling, although labels for these too-scarce classes may still be assigned by human coders. Table 8(a) shows F-measure classification performance statistics for each class used to determine sampling, as assessed at the end of the active learning process for the Tweets that have `place` information. By the end of active learning there are 4,018 human-labeled Tweet texts.

*** Table 8 about here ***

For both the set of Tweets that have `place` information and the larger set of Tweets, we use a binarized approach with the ensemble classifier for final classifications.[19] We predict classes only

[18]In particular we searched for the strings "disabl," "handicap," "technology" and "electronic." By this method we added 18 type 2 incidents and ten type 11 incidents, along with a scattering of incidents of other types. We did not label as "not hits" Tweets we located through these keyword searches that did not actually report an incident.

[19]For details about the classifier see note 11.

for those classes that have a reasonably large set of human-labeled instances. Table 8(b) shows F-measure classification performance statistics for each such class.

2.3 Characteristics of Primary/Caucus Tweet Contents and Incidents

Based on `location` information, which does not reliably indicate whence the Tweet was sent, "hits" occur in every state in the primary/caucus period. As Table 9 shows, California then Colorado have the largest numbers of Tweets classified as complaint Tweets while the Dakotas and Wyoming have the smallest. For California we also have a breakdown of hits versus not-hits by county, shown in Table 10. 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 9 and 10 about here ***

In the future we may use machine classification to classify incidents by type, but for the moment we have humans performing such classifications manually, according to the scheme described in Appendix section 4.4. From the Tweet texts with locations in California that the are classified as hits, we selected a simple random sample of $n = 600$ to classify by type manually. Table 11 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 11 about here ***

Notable is that human coders decided that 249 of the 600 sampled Tweet texts that were 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 5. 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 12 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 12 about here ***

2.3.1 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 13. Because no codebook for the California categories is available to explain their meaning,[20] it is difficult to say how the distribution of hotline complaints compares to the distribution of election-day Tweet texts presented in Table 12. 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

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

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 13 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 14 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 10 for a time period that includes but is not restricted to election day. Table 15 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.

*** Tables 14 and 15 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 12, for election day). But the machine classification performance will 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.

2.4 Characteristics of General Election Tweet Contents and Incidents

Incidents occur in every state in the general election period. As Table 16(a) shows, among the Tweets that have `place` information, the highest count of Tweet texts that are labeled or classified as incident observations occur in California, Texas, Florida and New York and the smallest in Wyoming, North Dakota, South Dakota and Montana.[21] Table 16(b) shows these same states have the largest and smallest counts of incidents among the larger set of all Tweets:[22] Hawaii has fewer incident-observing Tweet texts than does Montana.

*** Table 16 about here ***

The rate of incidents in the sense of incidents per person is not the same across states. To adjust the counts of "hits" for the populations of the various states, Table 17 shows the distributions in terms of observations per million persons in each state. In both the set of Tweets that have `place` information and in the larger set of Tweets, on a per capita basis the District of Columbia stands out with the highest rate followed by Nevada and North Carolina. Wyoming is lowest.

*** Table 17 about here ***

Plotting incident observations by day shows that the most observations occur on election day. Figure 1(a) uses the 40,678 Tweets that have `place` information and Figure 1(b) uses all 315,180 Tweets either with or without `place` information to display histograms of the number of classified "hit" Tweets on each day during October 1 through November 8, 2016.[23] Both histograms show the same pattern of variation over days. The similarity between the histograms provides some evidence that the set of incidents is similar regardless of whether the `place` identifying option had been enabled by the Twitter user.

*** Figure 1 about here ***

[21] For 255 of the Tweets with `place` information that information neither allowed the state to be identified nor indicated the Tweet did not originate in the United States. For all but 65 of these Tweets we used `location` information to identify the state. The `location` information places six of these 65 Tweets outside the United States, eight in "United States," two in one of three states (e.g., "DC MD VA #DMV"), and the rest have information that is geographically uninterpretable.

[22] For Tweets that lack `place` we attempted to recover state locations from `location` information. The `location` information describes the user and is written by the user, so the entries are idiosyncratic. Even if the `location` describes a real geographic location, that location is not necessarily the place from which the Tweet was sent.

[23] The last bar on the right in the histograms 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.

Figures 2 and 3 show the distributions over time of incident observations by type. A report of success voting on election day, during early voting or by absentee ballot is the most frequent observed incident, with more than ten thousand Tweets, although hundreds also report problems affecting voting or polling places (Figures 2(b) and 3(b)). The bulk of the success Tweets are "I voted!" declarations (often images of "I Voted" stickers). Long lines or waiting times to vote are the next most frequent kind of observation, with thousands of incidents on election day alone, although hundreds also observe that lines or waiting times are not very long on election day (Figures 2(a) and 3(a)). Reports of success with voter registration are slightly more frequent than reports of problems with voter registration in early October, a pattern that is reversed by election day (more Figure 3(d) than Figure 2(d)). For most of the period after October 1 praise of aspects of the election system is more frequent than reports of problems, although by election day the number of problems mentioned is nearly on par with the number of mentions of correct electoral system functioning (Figures 2(c) and 3(c)).

*** Figures 2 and 3 about here ***

Bivariate regression analyses show the type of incident observations depend on several variables. Included are variables that describe aspects of election administration in each state: whether a state requires some form of photo or non-photo identification ("Voter ID"); whether a state allows no excuse absentee voting ("No Excuse Absentee"); whether a state allows early voting or in-person absentee voting ("EV+In-person Abs."); whether a state has a complaint process outside of Help America Vote Act (HAVA); and whether there is at least one way (HAVA, non-HAVA, online portal) for voters in a state to submit complaints online. The type of incident also depends on a state's general-election turnout—measured in terms of the voting-eligible-population (VEP). State demographic variables such as race, ethnicity and educational attainment also relate to the type of incident.

Table 18 reports regressions that illustrate a few of these associations. Outcome variables are formed from the adjectives that describe three types of incidents: Line Length, Waiting Time; Polling Place Event (denoted "Voting"); and Absentee or Early Voting Issue. Levels of each adjective are associated with the numbers 0, 1 and 2: the value 2 represents a very long line (for Line Length), successful polling place operations or voting (Voting), or successful absentee or early voting operations (Absentee). In the regressions each type-of-incident variable is divided by the state population, so relationships concern the rate of incident reporting.[24] The table shows three models that include the Voter ID variable in interaction with three process variables: whether a state allows early voting ("Early Voting"); EV+In-person Abs.; and No Excuse Absentee. In all three cases the coefficients for Voter ID and for the other process variable have significant positive signs while the interaction has a significant negative sign. The fourth model in Table 18 includes the proportion White and the proportion with at least a bachelor's degree plus the interaction between these two variables. The proportion White and the proportion with at least a bachelor's degree each has a positive coefficient and the interaction has a negative coefficient. This means that, for instance, lines/wait-times are said to be shorter in states with high proportions of both whites and college graduates but otherwise longer.

*** Table 18 about here ***

Associations like these are hard to interpret, but at least they suggest that the incident measures we have recovered from Tweets measure potentially interesting phenomena.

3 Discussion

Every indication is that Twitter can be used to develop data about 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: times can be resolved to the millisecond 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

[24] Most covariates also relate to the unadjusted counts.

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—and more generally whether different types of users Tweet different kinds of observations.

An important immediate step for development is to try better to exploit geolocation information. `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. Some Tweets contain exact information about the polling location in the Tweet text (or image), and we plan to investigate how to organize such information.

`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 necessarily 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.

Another important development will be to add capabilities for machine classifiers jointly to use text and image information. Classifier performance for incidents such as line lengths and success at voting is good, but we expect that it would improve significantly if the classifier algorithms were able to interpret both images and text. Many Tweets that humans label for such instances have text that boils down to "look at this!" with an image clearly displaying a polling place, a long line or a smiling person wearing an "I voted" sticker. In fact we're a bit surprised at how well the classifiers perform given that human judgments so frequently depend on images to which the classifiers have no access.

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 privacy 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 general 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 so they can be critically appraised rather than not obtain the data at all.

4 Appendices

4.1 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.[25] 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 states.[26] 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.[27] 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.[28]

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

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

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

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

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

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[29] 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[30], so we had to identify the specfic fields of interest (in this case, the 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

[29] 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`.

[30] 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`.

JSON metadata, but is associated with individual tweets rather than with a users' profile.[31].

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

4.2 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 19 shows the number of Tweets in each of the state-term strata in the full set of Tweets manually

[31]`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)

[32]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).

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 19.

*** Table 19 about here ***

The sample used for the initial general election training set was drawn in stages. The population is the Tweets from the streaming API that have `place` information. Initially we drew a simple random sample of 2,000 Tweets. Then from the remaining Tweets we added another sample of 10,000 Tweets. "Hits" being sparse—there were only 247 "hits"—we used a nearest-neighbor algorithm (trying to match the "hits") to select an additional 2,969 Tweets from 482,485 unique Tweet texts. Then from the remaining Tweets we added first a sample of 5,000 Tweets then a sample of 10,000 Tweets stratified to include 5,000 Tweets during Oct 1-Nov 3 and 5,000 Tweets from Nov 4—Nov 8. Finally based on all the "hits" found in the previous samples, we added another 4,140 nearest neighbors. Dropping duplicate Tweet texts and Tweets whose `uniqueid` values had become corrupted we labeled in all 20,139 Tweet texts as "hits" or "not hits." Eliminating Tweets for which the `place` object exists but the `place$placename` value is missing leaves us with an initial training set of 19,179 Tweets.

4.3 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.

4.4 Coding Scheme for Primary/Caucus Tweets

Updated 8/21/16 (Version 5)

4.4.1 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

4.4.2 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 to 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).

4.5 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 to choose 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

4.5.1 Categories for Coding

1. 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 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

1: Neutral mention of disability/accessibilty concerns.

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 or by in-person early voting 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

1: the Tweet makes an unspecific assertion about voter fraud.

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. Also situations involving electronic pollbooks. 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.

Brady, Henry E. and John E. McNulty. 2011. “Turning Out to Vote: The Costs of Finding and Getting to the Polling Place.” *American Political Science Review* 105(1):115–134.

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.

Hall, Joseph Lorenzo. 2005. “Towards Work-Practice Modeling of Elections and the Election Incident Reporting System (EIRS).” URL `https://josephhall.org/papers/eirs_brahms_mims.pdf` (accessed July 5, 2006).

Herrnson, Paul S., Richard G. Niemi, Michael J. Hanmer, Benjamin B. Bederson, Frederick G. Conrad and Michael W. Traugott. 2008. *Voting Technology: The Not-So-Simple Act of Casting a Ballot*. Washington: Brookings.

Herron, Michael C. and Daniel A. Smith. 2016. “Precinct Resources and Voter Wait Times.” *Electoral Studies* 42:249–263.

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.

Johnson, Jeff. 2005. “The EIRS 2004 and 2005 Voting Problem Incident Data: A Valuable Public Resource.” URL `http://triton.towson.edu/~jlazar/EIRS_factsheet.doc` (accessed July 10, 2014).

Jones, Douglas W. and Barbara Simons. 2012. *Broken Ballots: Will Your Vote Count?* Chicago: University of Chicago Press.

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.

Lausen, Marcia. 2007. *Design for Democracy: Ballot and Election Design*. Chicago: University of Chicago Press.

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., Alejandro Pineda, Logan Woods, Joseph Klaver, Patrick Wu and Blake Miller. 2017. "Using Twitter to Observe Election Incidents in the United States." Paper presented at the 2016 Annual Meeting of the Midwest Political Science Association, Chicago, April 6–9, 2017.

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.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay. 2011. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research* 12:2825–2830.

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.

Quesenbery, Whitney and Margaret Chen. 2008. "Better Ballots." Brennan Center for Justice. URL: `https://www.brennancenter.org/publication/better-ballots`.

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

Rozenas, Arturas. 2017. "Detecting Election Fraud from Irregularities in Vote-Share Distributions." *Political Analysis* 25(1):41–56.

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. 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`.

Stewart, III, Charles and Stephen Ansolabehere. 2015. "Waiting to Vote." *Election Law Journal* 14(1):47–53.

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.

Verified Voting Foundation. 2005. "Election Incident Reporting System." URL https://voteprotect.org/index.php?display=EIRMapNation&tab=ED04 (accessed July 5, 2006).

Table 1: 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 2: 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 3: 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 4: Number of General Election Incident Observations in Training Sample by State

State	Count	State	Count	State	Count
Alabama	23	Kentucky	19	North Dakota	4
Alaska	4	Louisiana	42	Ohio	116
Arizona	62	Maine	9	Oklahoma	29
Arkansas	18	Maryland	71	Oregon	25
California	245	Massachusetts	92	Pennsylvania	72
Colorado	28	Michigan	44	Rhode Island	5
Connecticut	13	Minnesota	42	South Carolina	38
Delaware	6	Mississippi	10	South Dakota	2
District of Columbia	64	Missouri	48	Tennessee	65
Florida	157	Montana	5	Texas	296
Georgia	99	Nebraska	12	Utah	19
Hawaii	7	Nevada	58	Vermont	1
Idaho	4	New Hampshire	7	Virginia	61
Illinois	85	New Jersey	34	Washington	43
Indiana	69	New Mexico	5	West Virginia	12
Iowa	15	New York	188	Wisconsin	25
Kansas	18	North Carolina	175	Wyoming	2
Puerto Rico	1	Virgin Islands	1		

Note: Number of Tweets observing incidents in initial human-labeled training sample ($n = 2,610$) by State obtained via Twitter Streaming API during October 1–November 8, 2016.

Table 5: Primary/Caucus Machine Classifier Performance

Class	Precision	Recall	F-Measure	Support
Not a hit	.82	.93	.87	1065
Hit	.68	.42	.52	380
Average/Total	.78	.80	.78	1445

Table 6: General Election Machine "Hit" Classifier Performance

	Without Stemming			With Stemming			
Class	Precision	Recall	F-Measure	Precision	Recall	F-Measure	Support
Not a hit	.92	.93	.92	.91	.92	.92	3300
Hit	.78	.76	.77	.81	.72	.76	1127
Average/Total	.88	.89	.88	.88	.89	.88	4427

Table 7: Incident Type Frequency in Initial Sample of General Election "Hits"

Incident Category		Raw	Adjectives		
Description	Number	Count	0	1	2
Outside Influence	1	1	—	—	—
Disability/Accessibility Issue	2	2	1	0	1
Line Length, Waiting Time	3	752	61	77	614
Polling Place Event	4	700	51	327	322
Electoral System	5	199	46	137	16
Absentee or Early Voting Issue	6	733	70	269	394
Election Official	7	29	9	1	19
Voter Identification	8	14	2	7	5
Registration	9	156	27	107	22
Voter Fraud	10	5	1	4	0
Ballot and Voting Technology	11	20	15	4	1
Unspecified Other	12	3	—	—	—
Not an Incident		42	—	—	—

Note: Manual type classifications for 1,149 Tweet texts sampled from the 40,687 Tweet texts classified as "hits," using the coding scheme described in Appendix Section 4.5. Dashes indicate subtypes ("adjectives") that are not defined.

Table 8: General Election Binarized Classifier Performance

(a) SVM Classifier:

	Raw		Adjective					
			0		1		2	
Class	F	*M*	F	*M*	F	*M*	F	*M*
Outside Influence	—	2	*		*		*	
Disability/Accessibility Issue	1.00	21	—	9	—	0	—	12
Line Length, Waiting Time	.88	996	.26	82	.26	85	.82	829
Polling Place Event	.76	1411	.27	83	.52	449	.61	879
Electoral System	.62	552	.08	81	.67	438	.15	33
Absentee or Early Voting Issue	.84	1614	.33	116	.60	565	.66	933
Election Official	.07	50	—	17	—	7	.22	26
Voter Identification	.63	31	—	8	—	16	—	7
Registration	.86	440	.24	58	.78	308	.41	74
Voter Fraud	—	15	—	7	—	8	—	0
Ballot and Voting Technology	.29	48	.27	31	—	15	—	2
Unspecified Other	—	3	*		*		*	
Not an incident	.63	1044	*		*		*	

(b) Ensemble Classifier:

	Raw		Adjective					
			0		1		2	
Class	F	*M*	F	*M*	F	*M*	F	*M*
Line Length, Waiting Time	.91	996	.61	82	.21	85	.84	829
Polling Place Event	.78	1411	.08	83	.47	449	.63	879
Electoral System	.63	552	.05	81	.65	438	—	
Absentee or Early Voting Issue	.87	1614	.34	116	.60	565	.70	933
Registration	.85	440	.17	58	.84	308	.50	74
Not an incident	.66	1044	*		*		*	

Note: Overall number of labeled Tweets: 4,018. "F" is F-measure and *M* is support (the number of instances of the class). A dash indicates a class that is not used to determine (a) active-learning sampling or (b) a final classification. An asterisk indicates a class that is not defined. (a) binarized SVM performance at the end of the process of human labeling of types of incidents guided by active learning. (b) binarized ensemble classifier performance.

Table 9: Primary/Caucus Election Complaint Tweets by State

Unique Tweet Texts[a]				All Tweets[b]			
State	count	State	count	State	count	State	count
AK	162	MT	37	AK	175	MT	44
AL	9	NC	204	AL	29	NC	301
AR	55	ND	11	AR	65	ND	12
AZ	3,756	NE	111	AZ	4,065	NE	143
CA	11,903	NH	573	CA	13,064	NH	713
CO	3,916	NJ	41	CO	4,087	NJ	65
CT	677	NM	52	CT	796	NM	78
DC	229	NV	832	DC	309	NV	858
DE	19	NY	2,017	DE	29	NY	2,285
FL	440	OH	397	FL	594	OH	450
GA	143	OK	53	GA	225	OK	72
HI	256	OR	72	HI	319	OR	100
IA	2,543	PA	863	IA	3,113	PA	1,179
ID	149	RI	117	ID	191	RI	149
IL	3,366	SC	741	IL	3,680	SC	981
IN	379	SD	1	IN	550	SD	1
KS	520	TN	107	KS	526	TN	132
KY	742	TX	525	KY	922	TX	682
LA	30	UT	109	LA	35	UT	130
MA	208	VA	219	MA	265	VA	269
MD	321	VT	39	MD	407	VT	47
ME	100	WA	3,244	ME	132	WA	3,602
MI	238	WI	1,090	MI	310	WI	1,363
MN	1,097	WV	229	MN	1,356	WV	274
MO	77	WY	13	MO	103	WY	15
MS	39			MS	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 10: California Primary Complaint Tweets by County

County	count	County	count
Alameda	323	Placer	46
Alpine	2	Plumas	3
Amador	3	Riverside	97
Butte	20	Sacramento	287
Calaveras	0	San Benito	6
Colusa	1	San Bernardino	99
Contra Costa	106	San Diego	674
Del Norte	1	San Francisco	873
El Dorado	9	San Joaquin	43
Fresno	52	San Luis Obispo	61
Glenn	3	San Mateo	37
Humboldt	23	Santa Barbara	38
Imperial	6	Santa Clara	179
Inyo	1	Santa Cruz	42
Kern	24	Shasta	18
Kings	0	Siskiyou	2
Lake	15	Solano	12
Lassen	2	Sonoma	31
Los Angeles	3,271	Stanislaus	19
Madera	8	Sutter	5
Marin	21	Tehama	6
Mariposa	0	Trinity	1
Mendocino	2	Tulare	15
Merced	6	Tuolumne	3
Modoc	0	Ventura	48
Mono	1	Yolo	65
Monterey	41	Yuba	0
Napa	3	Bay Area	131
Nevada	2	Silicon Valley	28
Orange	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 11: Frequency of Incidents by Type in Sample of California Primary Complaint Tweets

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

Type	Unique Tweet Texts[a] Count	Percent	Omit Positive[b] Count	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 13: 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 14: 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 15: “Hit” Classification of California Primary Election-Day Tweets by County

County	count	County	count
Alameda	112	Riverside	34
Alpine	2	Sacramento	82
Amador	1	San Benito	1
Butte	10	San Bernardino	38
Contra Costa	21	San Diego	210
Del Norte	0	San Francisco	305
El Dorado	2	San Joaquin	12
Fresno	15	San Luis Obispo	4
Glenn	1	San Mateo	14
Humboldt	2	Santa Barbara	10
Imperial	2	Santa Clara	65
Kern	12	Santa Cruz	13
Lake	7	Shasta	5
Lassen	0	Siskiyou	0
Los Angeles	1,355	Solano	4
Madera	4	Sonoma	10
Marin	6	Stanislaus	11
Mariposa	0	Sutter	0
Mendocino	2	Tehama	1
Merced	3	Trinity	1
Monterey	11	Tulare	1
Napa	0	Ventura	20
Nevada	0	Yolo	11
Orange	64	Bay Area	46
Placer	16	Silicon Valley	13
Plumas	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 16: General Election Incident Observation Tweets by State

(a) Tweets with `place` information:

Unique Tweet Texts

State	count	State	count	State	count	State	count
AK	52	ID	80	MT	45	RI	88
AL	401	IL	1,279	NC	2,079	SC	402
AR	289	IN	848	ND	42	SD	45
AZ	823	KS	268	NE	151	TN	893
CA	4,522	KY	386	NH	106	TX	4,395
CO	590	LA	526	NJ	737	UT	239
CT	272	MA	1,043	NM	141	VA	1,185
DC	787	MD	978	NV	707	VT	59
DE	85	ME	132	NY	2,773	WA	745
FL	3,249	MI	792	OH	1,673	WI	401
GA	1,532	MN	479	OK	351	WV	154
HI	122	MO	556	OR	383	WY	16
IA	247	MS	198	PA	1,358		
PR	19	VI	2				

(b) Tweets with or without `place` information:

Unique Tweet Texts:

State	count	State	count	State	count	State	count
AK	133	ID	510	MT	289	RI	454
AL	1,775	IL	5,381	NC	8,274	SC	1,626
AR	1,277	IN	3,202	ND	175	SD	244
AZ	1,907	KS	1,559	NE	871	TN	3,872
CA	20,546	KY	1,373	NH	459	TX	19,922
CO	3,121	LA	2,348	NJ	2,806	UT	1,259
CT	1,274	MA	5,433	NM	722	VA	4,625
DC	6,047	MD	3,299	NV	2,489	VT	346
DE	380	ME	627	NY	15,182	WA	4,095
FL	12,552	MI	3,455	OH	6,342	WI	2,224
GA	6,069	MN	2,658	OK	1,855	WV	546
HI	273	MO	2,720	OR	2,317	WY	111
IA	1,220	MS	674	PA	5,372		
PR	121	VI	57				

Note: Number of unique Tweet texts classified as “hits” by State. (a) Counts using the 39,726 (of 40,678) Tweets for which a state could be identified from `place` or `location` information. (b) Counts using the 176,468 (of 315,180) Tweets for which a state could be identified from `place` or `location` information.

Table 17: Per Capita General Election Incident Observations by State

(a) Tweets with `place` information:

Unique Tweet Texts

State	rate	State	rate	State	rate	State	rate
AK	70.1	ID	47.5	MT	43.2	RI	83.3
AL	82.5	IL	99.9	NC	204.9	SC	81.0
AR	96.7	IN	127.8	ND	55.4	SD	52.0
AZ	118.7	KS	92.2	NE	79.2	TN	134.3
CA	115.2	KY	87.0	NH	79.4	TX	157.7
CO	106.5	LA	112.4	NJ	82.4	UT	78.3
CT	76.1	MA	153.1	NM	67.8	VA	140.9
DC	1155.4	MD	162.6	NV	240.5	VT	94.5
DE	89.3	ME	99.1	NY	140.4	WA	102.2
FL	157.6	MI	79.8	OH	144.0	WI	69.4
GA	148.6	MN	86.8	OK	89.5	WV	84.1
HI	85.4	MO	91.3	OR	93.6	WY	27.3
IA	78.8	MS	66.2	PA	106.2		

(b) Tweets with or without `place` information:

Unique Tweet Texts

State	rate	State	rate	State	rate	State	rate
AK	179.3	ID	303.0	MT	277.2	RI	429.8
AL	365.0	IL	420.3	NC	815.4	SC	327.7
AR	427.3	IN	482.7	ND	230.9	SD	281.9
AZ	275.1	KS	536.2	NE	456.7	TN	582.2
CA	523.5	KY	309.4	NH	343.9	TX	715.0
CO	563.3	LA	501.5	NJ	313.7	UT	412.6
CT	356.2	MA	797.6	NM	346.9	VA	549.8
DC	8877.4	MD	548.3	NV	846.6	VT	554.0
DE	399.1	ME	470.9	NY	768.9	WA	561.9
FL	609.0	MI	348.0	OH	546.0	WI	384.9
GA	588.6	MN	481.5	OK	472.8	WV	298.2
HI	191.1	MO	446.4	OR	566.0	WY	189.6
IA	389.2	MS	225.5	PA	420.2		

Note: Number per million persons (per state) of Tweets observing incidents based on (a) 39,726 and (b) 176,468 classified "hits" obtained via Twitter APIs for the general election period. 2016 state population data source: United States Census Bureau (2016).

Table 18: OLS Regression Models for State-level Per Capita Incidents

		Line Length		Voting		Absentee	
Model	Covariate	coef.	s.e.	coef.	s.e.	coef.	s.e.
(a)	(Intercept)	.248	.011	.241	.0067	.272	.0056
	Voter ID	.091	.014	.074	.0093	.016	.0076
	Early Voting	.297	.014	.190	.0087	.117	.0070
	ID × Early Voting	−.456	.019	−.331	.0119	−.257	.0094
(b)	(Intercept)	.150	.014	.148	.0086	.184	.0078
	Voter ID	.197	.019	.184	.0122	.131	.0105
	EV plus In-person Abs.	.366	.016	.273	.0099	.201	.0086
	ID × EV+In-person Abs.	−.496	.022	−.406	.0139	−.347	.0116
(c)	(Intercept)	.192	.012	.184	.0075	.209	.0063
	Voter ID	.038	.014	.034	.0090	−.025	.0074
	No Excuse Absentee	.346	.015	.250	.0091	.192	.0075
	ID × No Excuse Absentee	−.300	.020	−.228	.0122	−.165	.0096
(d)	(Intercept)	−6.8	.06	−5.7	.04	−5.3	.03
	White	12.3	.12	10.3	.09	9.5	.07
	Bachelor's plus	33.9	.25	28.1	.19	26.0	.14
	White × Bachelor's plus	−59.0	.55	−48.5	.41	−44.0	.31

Note: ordinary least squares regression coefficients and standard errors estimated using classified incident types among the 176,468 incident Tweet texts for which state information could be extracted. "Line Length" models: $n = 13602$. "Voting" models: $n = 25776$. "Absentee" models: $n = 41737$.

Table 19: Strata for Sampling Tweets to use in Initial Primary/Caucus Training Set

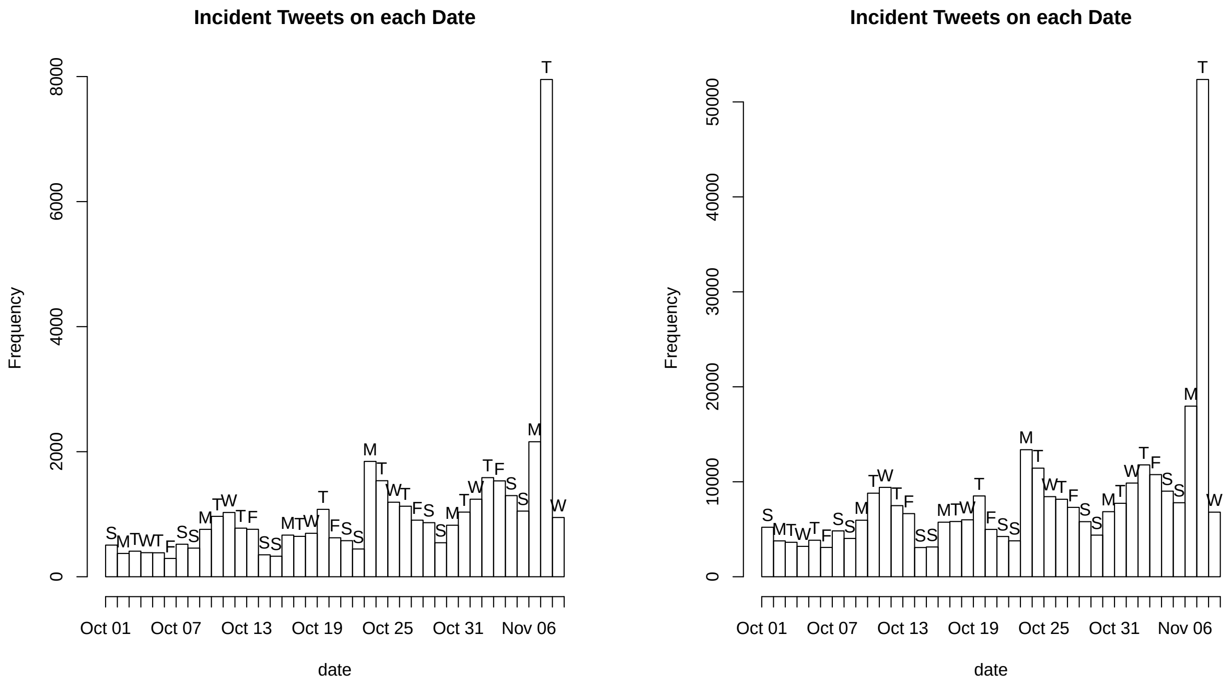
		population		sample	
state+	total[a]	not hit	hit	not hit	hit
AZ	9,890	1,607	478	62	109
CA	52,296	10,774	271	414	62
CT	3,537	712	24	27	6
CO	8,388	1,511	261	58	59
WAd	10,062	1,958	169	75	39
WAr	2,910	608	7	23	2
CAeo	3,041	558	72	21	16
COeo	177	68	6	3	2
WAeo	505	105	3	4	2

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 by Date

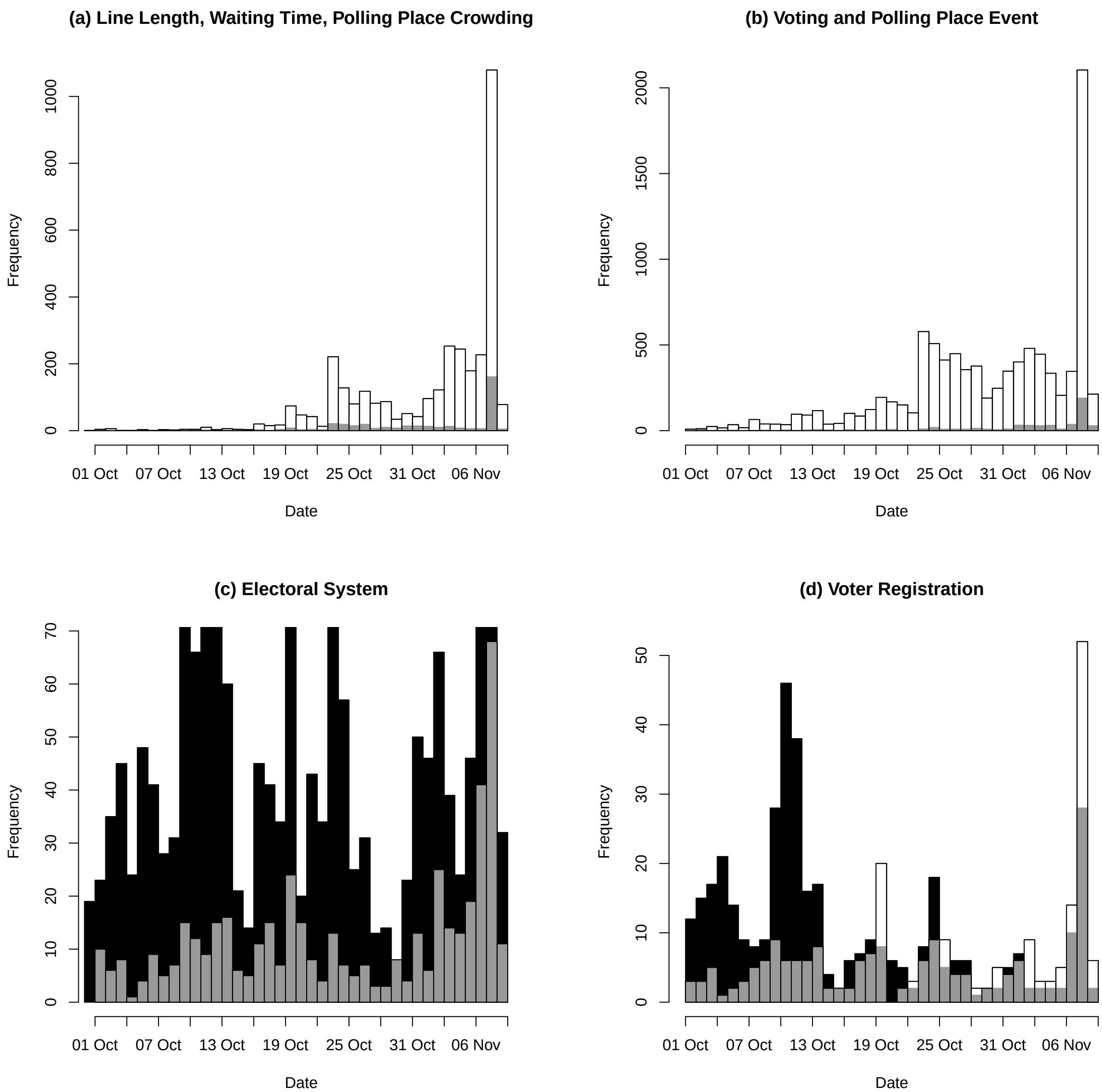
(a) Tweets with `place` information

(b) all Tweets



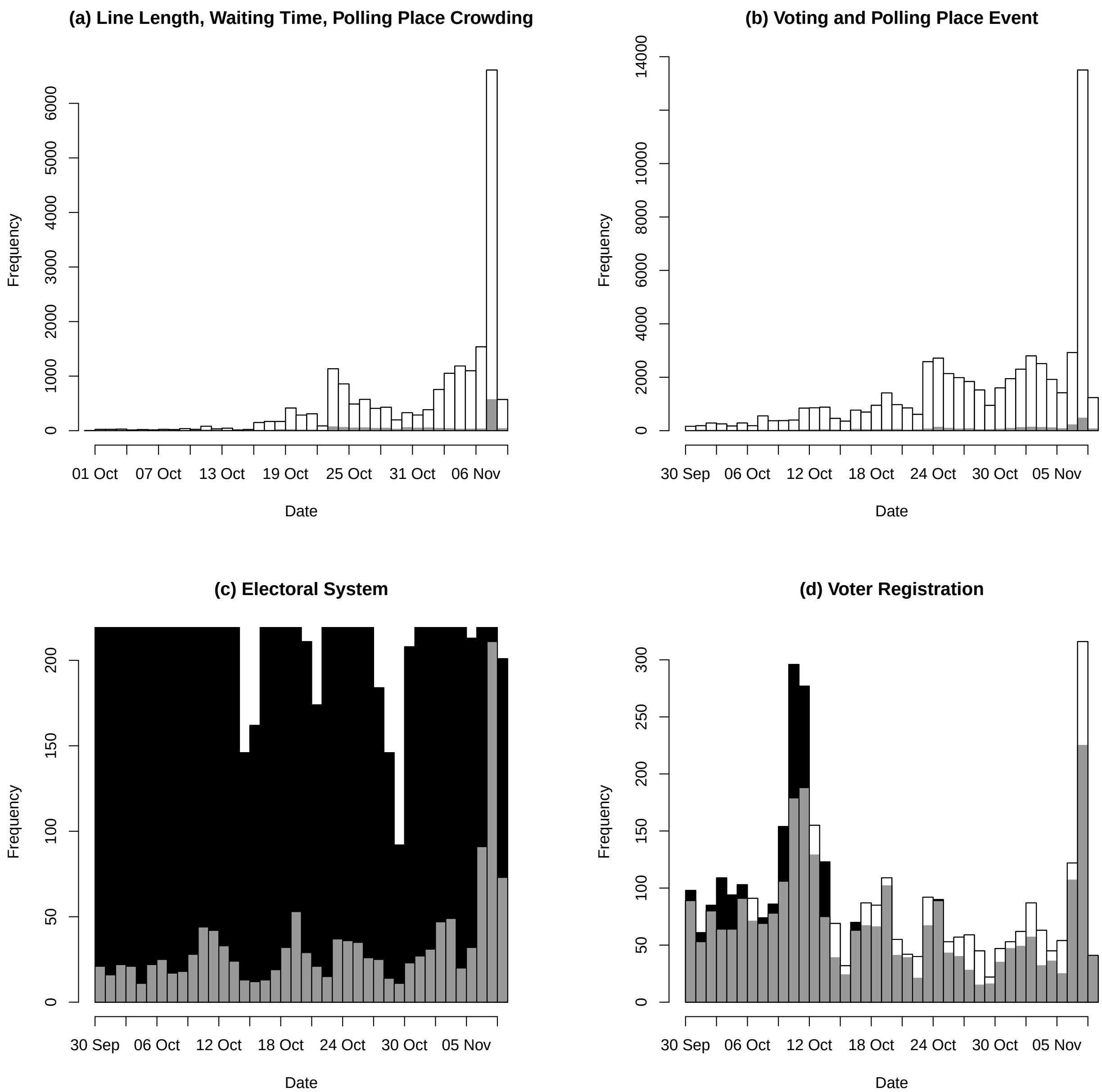
Note: (a) incidents by day for 40,678 Tweets with `place` information. (b) incidents by day for 315,180 Tweets with or without `place` information.

Figure 2: General Election Incident Observations by Type and Date



Note: using classifications for all Tweets that have `place` information. (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, election system problem; black, election system success. (d) white, voter registration problem; black, registration success.

Figure 3: General Election Incident Observations by Type and Date



Note: using classifications for all Tweets with or without `place` information. (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, election system problem; black, election system success. (d) white, voter registration problem; black, registration success.

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

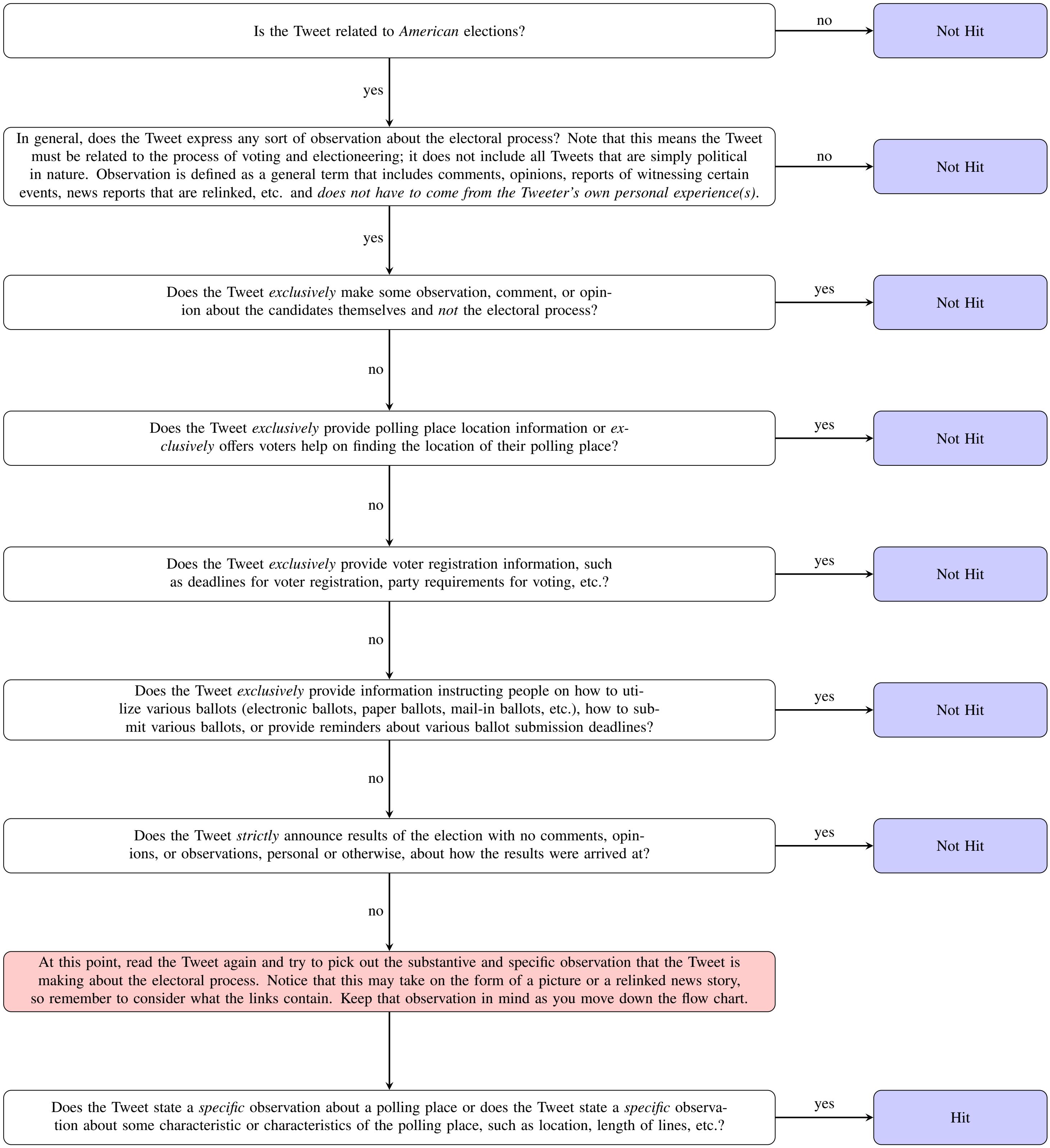


Figure 5: Primary/Caucus Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 2

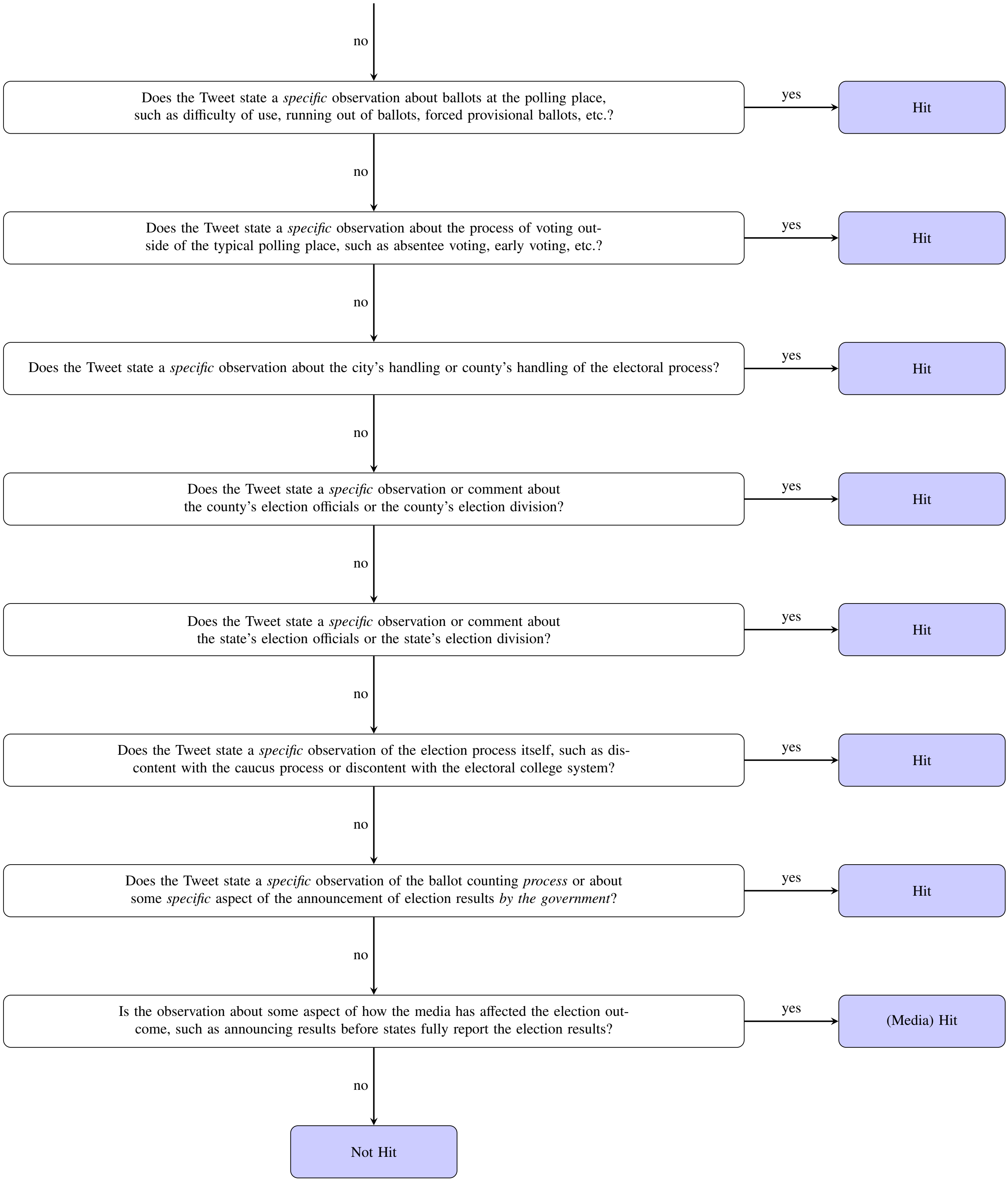


Figure 6: General Election Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 1

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.

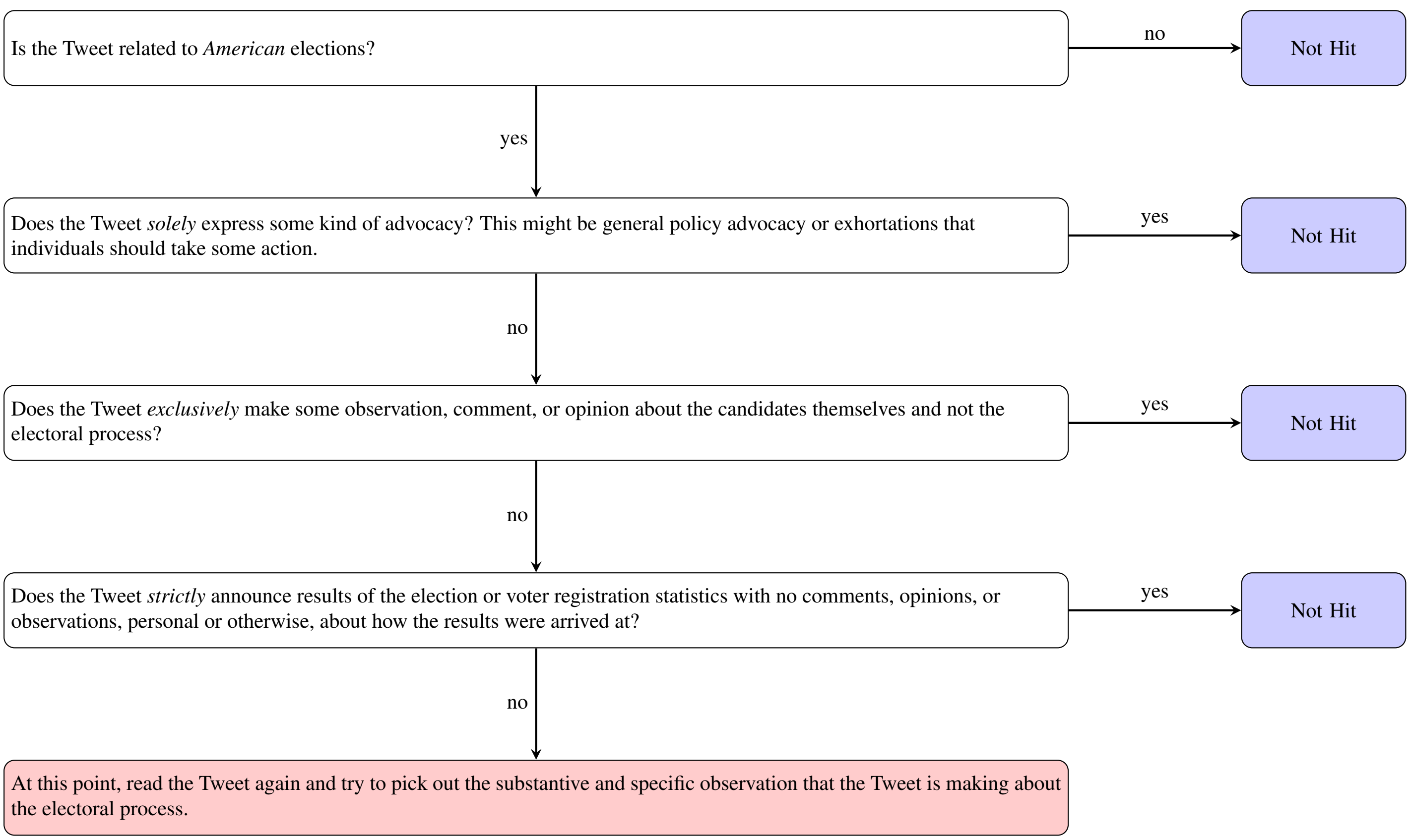


Figure 7: General Election Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 2

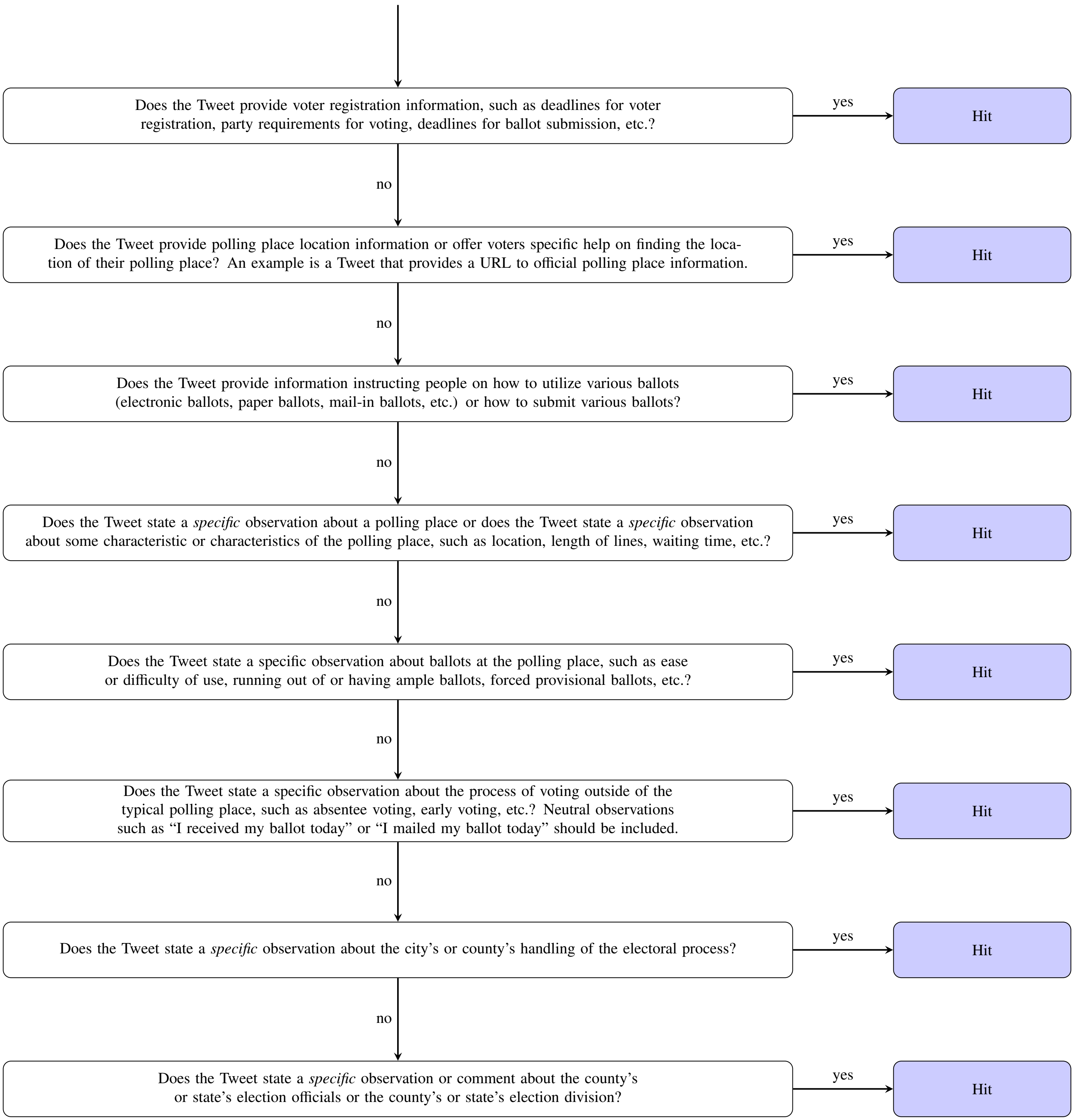


Figure 8: General Election Flowchart for Making Hits Decisions in American Twitter Election Comments, Part 3