MDPI

Article

Electoral and Public Opinion Forecasts with Social Media Data: A Meta-Analysis

Marko M. Skoric [1,*], Jing Liu [2] and Kokil Jaidka [3]

[1] Department of Media and Communication, City University of Hong Kong, Hong Kong 999077, China
[2] HKUST Business School, Hong Kong University of Science and Technology, Hong Kong 999077, China; echocityu86@gmail.com
[3] Department of Communications and New Media, National University of Singapore, Singapore 117416, Singapore; jaidka@nus.edu.sg
[*] Correspondence: mskoric@cityu.edu.hk

check for updates

Abstract: In recent years, many studies have used social media data to make estimates of electoral outcomes and public opinion. This paper reports the findings from a meta-analysis examining the predictive power of social media data by focusing on various sources of data and different methods of prediction; i.e., (1) sentiment analysis, and (2) analysis of structural features. Our results, based on the data from 74 published studies, show significant variance in the accuracy of predictions, which were on average behind the established benchmarks in traditional survey research. In terms of the approaches used, the study shows that machine learning-based estimates are generally superior to those derived from pre-existing lexica, and that a combination of structural features and sentiment analyses provides the most accurate predictions. Furthermore, our study shows some differences in the predictive power of social media data across different levels of political democracy and different electoral systems. We also note that since the accuracy of election and public opinion forecasts varies depending on which statistical estimates are used, the scientific community should aim to adopt a more standardized approach to analyzing and reporting social media data-derived predictions in the future.

Keywords: social media; public opinion; computational methods; meta-analysis

1. Introduction

Scholars have suggested that the social sciences are experiencing a momentous shift from data scarcity to data abundance [1,2], noting, however, that full access to data is likely to be reserved for the most powerful corporate, governmental, and academic institutions [3]. Indeed, as a growing share of our social interactions happen on digital platforms, our capacity to predict attitudes and behaviors should increase as well, given the scale, richness, and temporality of such data. Still, there is a growing debate in the research community about the usefulness of social media signals in understanding public opinion. Social scientists have identified the problems with using "repurposed" data afforded by social media platforms, which are often incomplete, noisy, unstructured, unrepresentative, and algorithmically confounded [4] compared to traditional surveys, which are carefully designed, representative, and structured.

Many recent studies have utilized social media data as a "social sensor" aiming to predict different economic, social, and political phenomena, ranging from election results to the box office success of Hollywood movies—with some achieving considerable success in this endeavor. Still, most of these studies have been primarily data-driven and largely atheoretical. Furthermore, instead of predicting the future, these studies have made predictions about the present [5] or the past, and have generally

not examined the validity of their predictions by comparing them with more robust types of data, such as surveys and government censuses. Researchers have also raised concerned about the lack of representativeness and authenticity of data harvested from social media platforms, noting the demographic biases, the curated nature of social media representations [6], biases introduced by sampling methods [7] and pre-processing steps [8], platform affordances [9], and false positives introduced by bot accounts [10]. Still, when compared to the more traditional methods, social media analytics promise to provide an entirely new level of insight into social phenomena, by allowing researchers unprecedented access to people's daily conversations and their networks of friendship and influence, across time and geographic boundaries.

Thus far, only one meta-analysis assessing the predictive power of social media data has been published, focusing solely on Twitter-based predictions of elections [11]. This paper, reviewing only nine studies examining the predictive utility of Twitter messages, concluded that although social media data provides some insights regarding electoral outcomes, it is unlikely to replace survey-based predictions in the near future. Given the lack of scientific consensus about the merits of social media-based predictions and the growing number of studies using data from diverse social media sources, a more comprehensive meta-review is needed. We thus aim to compare the merits of the most commonly used approaches in social media-based predictions and also examine the roles of contextual variables, including those related to the level of democracy and type of electoral system.

2. Background: Measuring Public Opinion

2.1. Survey vs. Social Media Approaches

Social media-based predictions have challenged the key underlying principles of traditional, survey based research—namely, probability-based sampling and the structured, solicited-nature of participant data.

First, probability-based sampling, as one of the foundations of scientific survey research, is based on the assumption that all opinions are worth the same and should be weighted accordingly [12], thereby representing "a possible compromise to measure the climate of public opinion" [13] (p. 45). However, this method does not take into account the differences in the levels of interpersonal influence that exist in the population and are highly important in the process of public opinion formation [14,15]. Traditional polling can occasionally reveal whether the people holding a given opinion can be thought of as constituting a cohesive group; however, it does not provide much information about the opinions of leaders who may have an important role in shaping opinions of many other citizens. Ceron et al. noted that the predictive capacity of social media analysis does not necessarily rely on user representativeness, suggesting that "if we assume that the politically active internet users act like opinion-makers who are able to influence (or to 'anticipate') the preference of a wider audience: consequently, it would be found that the preferences expressed through social media today would affect (predict) the opinion of the entire population tomorrow" [16] (p. 345). In other words, simply discounting social media data as being invalid due to its inability to represent a population misses important dynamics that might make the data useful for opinion mining.

Second, the definition of public opinion as an articulated attitude toward certain subjects or issues might simplify and therefore miss the implicit, underlying components (affect, values) of public opinion. For example, Bourdieu [12] contends that these implicit fallacies stem from the artificiality of the isolated, survey-interviewing situation in which opinions are typically measured, compared to the real-life situations in which they are formed and expressed. Further fallacies may come from the design of survey questions—it is well established that solicited, survey-based responses can be influenced by social desirability, which can lead to serious biases in the data [17]. For instance, Berinsky demonstrated that some individuals who harbor anti-integrationist sentiments are likely to hide their socially unacceptable opinions behind a "don't know" response, and suggested that for many sensitive

topics it would be very difficult to obtain good estimates of public sentiment, as survey participants typically refrain from giving socially undesirable responses [18].

Scholars have also noted that "the formation of public opinion does not occur through an interaction of disparate individuals who share equally in the process" [19] (p. 544); instead, through discussions and debates in which citizens usually participate unequally, public opinion is formed. Drawing on Blumer's notion [19], Anstead and O'Loughlin re-theorized public opinion in social media analysis as "more than the sum of discrete preferences and instead as an on-going product of conversation, embedded in social relationship" [20] (p. 215). Such a definition implies that citizen associations and public meetings [21]—and newer forms like social media [22]—should be considered "organs of public opinion." Accordingly, data derived from these settings should be worthy of consideration in public opinion research.

Social media data differs from the structured, self-reported survey data in three main ways. First, it provides a large volume of user-generated content from organic, non-reactive expressions, generally avoiding social desirability biases and non-opinions, although self-censorship of political expressions has also been noted [23]. Second, social media data can be used to profile not only the active users but also "lurkers" who may not openly volunteer an opinion, by examining the relational connections and network attributes of their accounts. Third, by looking at social media posts over time, we can examine opinion dynamics, public sentiment and information diffusion within a population [24].

Webster recognized the agentic power of internet users in constituting the online communication structure, arguing that there are two genres measuring public attention in the current digital media environment [25]. The first one is the market information regime (e.g., audience rating, survey), the medium "through which producers observe each other and market participants make sense of their world" [26] (p. 272). By measuring linear media activities (e.g., the size and composition of available audiences), the market information regime has long influenced the operation of media industries and governments. The second genre is the user information regime (e.g., linkages, pathways of web traffic), a more recent development which includes social media, search engines, and recommender systems. The user information regime is designed for ordinary users, although institutions may exploit this medium to supplement market information. Compared with the traditional genre, the user information regime measures the patterns and structure of online human behavior and social interaction (i.e., click, link, sort, retrieve, recommend, comment, and search).

The differences between two genres of public measures that Webster proposed correspond closely to the differences between traditional polls and social media-based predictions of public opinion [25]. For example, the features that the user information regime measures (e.g., linkage, pathways of web traffic) are analytical features of social network analysis. In contrast, the survey-based market information regime collects and examines the expressed attitude of the scattered and linear public in order to provide guidelines for market decisions. The network-based user information regime examines the communication structure of a very large number of internet users, offering researchers more insights into the dynamics of opinion leadership as well as "the actual lay of attitudes in the population" [27].

Still, social media data is criticized for not being representative of the general population, as it typically comes from the ranks of early adopters, teens, and better-educated citizens [28,29]. Political discussions on social media also tend to be dominated by a small number of frequent users [30]. However, dismissing social media data as being invalid due to its inability to represent a population misses capturing the dynamics of opinion formation. As opinions held and debates conducted by certain politically active groups pre-empt those that develop in broader society [31], it is likely that social media conversations by active users play a stronger role in shaping public opinion.

We thus argue that answering the ontological question of "whether public opinion exists or not" is largely dependent on how we address the epistemological question of "how to measure public opinion." Public opinion and its measures are socially constructed, defined by the researchers and shaped to fit their theoretical orientations and methodological training. The emergence of social

media and the possibility of large-scale analyses of user-generated content and the social structure of communication provide researchers with a viable alternative approach to study public opinion, which may avoid some the problems associated with survey-based studies.

2.2. Measuring Public Opinion with Social Media Data

2.2.1. Diversity in Methods

The first research gap lies in identifying which approach—whether sentiment-based or social network-based—yields the most accurate predictions of election outcomes from social media. Sentiment-based studies mine positive and negative opinion as a proxy for voting behavior—some using pre-existing lexica [32–34] while others training new models of sentiment specifically for political tweets [35]. A drawback is that there is little consensus about which methods have the best predictive performance because different studies have reported success with different approaches and implementations.

O'Conner et al. demonstrated that tweet sentiment, mined using the OpinionFinder sentiment lexicon [36,37], correlated well with Obama's support rate during the 2008 presidential election and his job approval in surveys in 2009. On the other hand, Chung and Mustafaraj and O'Connor et al. reported that sentiment analysis with lexica like Subjectivity Lexicon and SentiWordNet are not reliable for predictions due to the low coverage in their dataset [36–39]. Other reasons why lexica fail are due to (1) incorrect classification of the word in the lexicon, (2) the mismatch of parts-of-speech (POS) in tweets due to their poor grammar, (3) lack of word disambiguation, and (4) their reliance on word polarity rather than context inference. Gayo-Avello reported that a lexicon-based sentiment analysis overestimated Obama's vote share in 2008 U.S. presidential election, pointing out that political predictions based on sentiment analysis (1) miss the subtleties of political language, (2) exhibit poor recall, and (3) produce unbalanced results, making it unrealistic to expect that errors will cancel out after the data aggregation [40].

Studies which trained new sentiment models using supervised machine learning methods appear to have had better success [32,34,41–47]. For instance, Contractor and Faruquie trained a regression model based on the bigram features from 37 million tweets to gauge Obama and Romney's approval rates in the 2012 U.S. Presidential election [35]. Mejova et al. demonstrated that tweet sentiments correlated poorly with the presidential approval rates obtained in national polls; they found the most popular tweets to be mostly joking banter about the politicians, all negative, even for the politicians whose national poll numbers were improving [48]. Monti et al. trained a classification algorithm using Twitter Train Data and News Train Data and observed a strong correlation between offline inefficiency and online disaffection [44].

We expect that machine learning will be superior to dictionary-based sentiment classifiers because they capture implicit signals of preference beyond simple positive or negative words. In the scope of our meta-analysis, we are concerned not with how well they capture signals of emotion, but whether these signals correspond to actual vote share or poll outcomes. Accordingly, we hypothesize that:

Hypothesis 1 (H1). *Compared with lexicon-based sentiment approaches, studies using machine learning sentiment methods yield more accurate predictions of public opinion polls and electoral outcomes.*

Note that in all our analyses, we have equally considered both actual electoral outcomes and traditional survey-based polls as valid benchmarks for our meta-analysis (see Section 3.2.2). This allows us to obtain a sense of the efficacy of social media mining against traditional poll results. Structural metrics of social networks encompass both the relational connection referring to the actions of "following" or "friending," and the interactive connection that can be built through "mentions," "likes," "comments," "replies," "retweets," "shares," etc. For example, the mentions of a political party reflect the interaction edges in a social network, while likes and follows are structural edges; both will ideally contribute to boosting the network centrality. While interactional behaviors often denote the public's

active engagement with political parties/candidates, the relational connections are relatively static links established between the public and political candidates. A similar conceptualization was found in prior work by Livne [49], Vepsalainen et al. [50], and Metaxas et al. [51] to recognize the structural signals implicit in mentions and retweets, and to distinguish structural features from sentiment features.

Early studies focused mainly on interaction edges, e.g., the number of times political parties/candidates are mentioned by social media users, as a proxy to predict political parties'/candidates' offline political support [30]. MacWilliams use Facebook's PTAT ("People Talking About This") data—counting the interactions between the public and the candidates to measure "the effectiveness of each campaign in enlisting and engaging Facebook users as well as its potential to mobilize the vote on Election day" [52] (p. 581). Such "participation advantage" improved upon models that used only the partisan vote index and incumbency as predictors. Pimenta, Obradovic, and Dengel predicted candidates' vote shares in 2012 Republican primaries and opinion polls using the number of incoming links to blog posts, number of likes/reposts a Facebook post received, number of retweets a tweet post received, the number of comments and likes/dislikes a YouTube video received, and so on [53]. However, simplistic interaction-based approaches have largely been criticized because they fail robustness checks [8,54]. Studies using relational edges have demonstrated that the "likes" recorded on candidates' Facebook pages/Fan pages could be used to predict electoral outcomes [50,52,55,56]. On the other hand, studies based on centrality-based structural metrics can model both the relational features of social networks [57] and the interactive behavior. In terms of more traditional social network measures, Livne et al. used centrality metrics (i.e., closeness, In/Out-degree) [50] and the PageRank of political candidates in Twitter to predict their success in 2010 U.S. midterm elections [58]. Cameron et al. found that the number of "friends" a candidate has on Facebook and number of "followers" he/she has on Twitter could be used the predict the candidate's vote share and the winner in 2011 New Zealand general election [59].

All these results suggest that structural features are likely to be more robust indicators of political preference as compared to sentiment. Accordingly, it is hypothesized that:

Hypothesis 2 (H2). *Structural features outperform sentiment features in predicting public opinion polls and electoral outcomes.*

Hypothesis 3 (H3). *A combination of structural and sentiment features will outperform any singular type of social media feature in predicting public opinion polls and electoral outcomes.*

2.2.2. Diversity in Data Sources

A second research gap in the literature is that it is currently not known whether one platform or data source yields more accurate predictions than others. Because of the public nature of posts on Twitter, most studies have utilized Twitter data to predict public opinion, followed by the use of Facebook, forums, blogs, and YouTube. Given that each platform suffers from its own set of algorithmic confounds, privacy constraints, and post restrictions, it is unknown whether consolidating data from multiple platforms would have any advantage over predictions from a single data source. Most studies have utilized data from a single platform, while very few use data from multiple platforms [47]. It is also reasonable to expect that studies based on multiple data sources would be more likely to cover a broader cross-section of the electorate. Accordingly, we hypothesize that:

Hypothesis 4 (H4). *Compared with single data sources, studies using multiple data sources are more likely to yield accurate predictions of public opinion polls and electoral outcomes.*

2.2.3. Diversity in Political Contexts

The existing literature has not provided specific insights regarding the role of political systems in influencing the predictive power of social media data mining. The predictive power in a study conducted in semi-authoritarian Singapore was significantly lower than in studies done in established democracies [30,60]. It can be inferred that the context in which the elections take place also matters. Issues like media freedom, competitiveness of the election, and idiosyncrasies of electoral systems may lead to over- and under-estimations of voters' preferences. Studies suggest incorporating local context into election prediction, i.e., by controlling for incumbency [61], because often the incumbents are better placed to build networks and use media resources to sustain the campaign through to victory [59,62,63]. To further explore the predictive power of social media data in different political contexts, we include the factors that are related to the ability of citizens to freely express themselves and exercise their voting rights, such as the level of political democracy. We also examine the role of the electoral system by comparing the findings from the countries using proportional representation to those that do not. Our research questions are as follows:

RQ1: Does the predictive power of social media data vary across different: a) levels of political democracy; and b) types of electoral systems?

3. Methods

3.1. Literature Search

We aimed to collect all the studies that used social media data to gauge public opinion or political behavior. Benchmarked against election results or opinion polls, these studies mainly predicted: (1) vote share received by a political party or candidate in diverse types of elections (presidential, parliamentary, administrative); (2) seat share received by a political party in parliamentary elections; (3) winning parties/candidates in election; and (4) public support rate, disaffection, or presidential approval rate, as measured in opinion polls.

The data collection was finalized in August 2018, using the following keywords—Twitter, Facebook, YouTube, microblog, blog, forum, social media, social networking site, online discussion or political sentiment, election, public opinion, protest, dissent or opinion mining, predict, measure, forecast, approximate—to search within the following databases: ACM Digital Library, IEEE Xplore, AAAI, ScienceDirect, Web of Science, EBSCO, JSTOR, SCOPUS, Taylor and Francis, Wiley Online Library, and ProQuest. Prediction studies based on search engine data were excluded from this review.

After the initial search, a manual selection was performed to filter for relevance. Studies were included if they (a) utilized social media data to predict offline political behavior or opinion; and (b) measured one or more of the three criterion variables (i.e., political voting, protest, or public opinion) as a predicted variable. This resulted in a corpus of 96 articles published between 2007 and 2018, among which 22 did not report any statistics regarding the predictive power. Finally, 74 studies were included in the meta-review (See Supplementary Materials).

3.2. Coding

3.2.1. Predictors

Social media predictors are first categorized into two types: sentiment vs. structure. Sentiment predictors refer to the sentiment/preference extracted from the users' text-based social media posts. Sentiment scores are obtained by using (1) lexicon-based approaches, which use dictionaries of weighted words, in which accuracy is highly dependent on the quality and relevance of the lexical resources to the domain it is being applied to; and (2) (supervised) machine learning-based models, which predict the sentiment score of a piece of text based on the regression weights that it learns from example data (a corpus) labeled by a human.

On the other hand, structural predictors represent the social structure of conversations within the social media platform. They are different from sentiment measures because they do not focus on the author's positive or negative sentiments, but rather on how others engage with the author's posts and profile. Structural predictors include (1) counts of structural connections formed by social media users and political candidates/parties through interactions and relationships, such as "follows," "likes," "mentions," "retweet," "reply," etc., and (2) centrality metrics. Count-based measures comprise (1) interactional behaviors between political candidates and the public in social media, including counts of "mentions," "retweets," "shares," "likes," "replies," "comments," etc.; (2) relational edges between political candidates and the public in social media, including counts of "followers" or "friends." For blogs or forums, the interactive edge refers to the number of incoming links to a blog post; for YouTube videos, it includes setting videos as "favorites." Centrality metrics are sophisticated approaches dominant in social network analysis and used to examine an individual entity's importance in its communication network.

3.2.2. Predicted Election and Traditional Poll Results

The types of outcomes predicted are: (1) votes share that political candidates or parties received during the election; (2) winning party or candidate in the election; (3) seat share that political candidates or parties received in the election; and (4) public support or approval towards certain political parties or candidates, which are often obtained through polls. We have categorized the dependent variables as either predicting (1) election results, or (2) a traditional poll result.

3.2.3. Data Source

The data sources refer to the social media applications from which the predictors are generated, which are categorized into (1) Twitter; (2) Facebook; (3) Forums (Usenet, PTT, Uwant.com, etc.); (4) Blogs; (5) YouTube; and (6) other platforms.

3.2.4. Context Variables

To explore the influence of contextual variables on social media's predictive power in various political contexts, we included the following four variables in the analysis: democracy score, electoral system, press freedom, and internet penetration. The latter two are used as control covariates in the main hypothesis tests.

Democracy Score. Democracy scores are drawn from the Democracy Index compiled by the Economist Intelligence Unit. Based on the democracy score released in the year of the study, countries across the world are categorized into four types of political regimes: (1) full democracy, referring to any country whose democracy score ranged from 8 to 10, including the U.S., the U.K., Germany, the Netherlands, and New Zealand; (2) flawed democracy, referring to any country whose democracy score ranged from 6 to 8, including Italy, France as of 2012, Hong Kong, Taiwan, Brazil, India, and Indonesia; (3) hybrid regime, referring to any country whose democracy score ranged from 4 to 6, including Singapore, Venezuela, and Ecuador; (4) authoritarian regime, referring to any country whose democracy score ranged from 0 to 4, including Nigeria as of 2011.

Electoral System. Based whether where the country's elected body could proportionally reflect the electorate, the electoral system of the country was categorized as (1) proportional representation or (2) not proportional representation. The categorization is drawn from ACE Electoral Knowledge Network, a public website of electoral knowledge.

Press Freedom. The degree of press freedom was drawn from Freedom House's annual report. As the most comprehensive data set available on global media freedom, the report accessed the independence degree of media in various forms throughout 199 countries and territories.

Internet Penetration. The degree of internet penetration measures the percentage of individuals using internet in that country, with data drawn from the International Telecommunication Union (ITU) World Telecommunication/ICT Indicators Database, 2015.

4. Results

Since each study may test more than one prediction, we ended up with 460 estimates in total—232 mean average error (MAE) or other convertible forms (RMSE, absolute error, etc.), 205 R squared or coefficients, and 23 estimate reported race-based accuracy (percentage of successful predictions among all races). Considering that race-based prediction accuracy is affected by the type of electoral system and number of candidates in the election, we decided to focus on the remaining 437 estimates for our meta-analysis.

4.1. Comparing the Predictive Accuracy of Supervised vs. Unsupervised Sentiment Approaches

To answer H1, an ANCOVA test was performed while distinguishing between the two sentiment-based approaches, and controlling for internet penetration, democracy score, electoral system, and freedom of the press. Unexpectedly, lexicon-based sentiment reported a higher R^2 but also a higher MAE than machine-learning sentiment. Statistically significant effects were observed for both MAE and R^2 ($F(4, 223) = 4.042$, $p = 0.003$ for MAE; $F(4, 196) = 3.455$, $p = 0.009$ for R^2). With respect to H1, we infer that the hypothesis is not supported. However, we suggest that while lexicon-based sentiments are able to capture more of the noise in the data, machine-learning based models have higher precision at measuring vote shares, hence the lower average MAEs. An implicit assumption here is that sentiment-based approaches follow a binary outcome, and the errors across positive and negative estimates are likely to be symmetrical to each other and thus suitable for measurement with MAE values.

4.2. Comparing the Predictive Accuracy of Structure vs. Sentiment

Sentiment-based approaches were conflated together to answer H2, and a one-way ANCOVA was performed for the different types of predictors (sentiment, structure, a combination of sentiment and structure) while controlling for internet penetration, democracy score, electoral system, and freedom of press. First, comparing the estimated marginal means (covariate-adjusted means) showed that the highest predictive accuracy (lowest MAE %) occurred when using a combination of structure and sentiment as predictors (mean MAE = 2.929; mean $R^2 = 0.621$). This is corroborated in Table 1, where the combination of machine learning-based sentiment analysis and structural features as the predictors yielded the highest predictive accuracy (mean =2.141 for MAE; mean = 0.818 for R^2). The results suggest that machine learning-based sentiment analysis approaches are robust at predicting offline politics.

Table 1. Covariate-adjusted means of structure, lexicon, machine learning, and combination approaches.

Predictors	MAE (%)			R^2		
	Mean	SD	N	Mean	SD	N
Lexicon Sentiment	6.427 [a]	0.758	30	0.690 [b]	0.083	14
Machine Learning Sentiment	3.644 [a]	0.466	74	0.406 [b]	0.053	51
Structure	5.038 [a]	0.396	103	0.615 [b]	0.030	114
ML-Sentiment and Structure	2.141 [a]	0.944	20	0.818 [b]	0.162	4
Lexi-Sentiment and Structure	4.828 [a]	1.793	5	0.607 [b]	0.087	22

[a] Covariates appearing in the model are evaluated at the following values: Internet penetration (%) = 65.2504, electoral system = 0.49, press freedom= 64.8276, democracy score = 7.6360. [b] Covariates appearing in the model are evaluated at the following values: Internet penetration (%) = 67.7971, electoral system = 0.20, press freedom = 65.3854, democracy score = 7.2079.

Unlike the results in Table 1, we would expect the errors accrued due to structure and combination-based methods to be asymmetrical; thus, we are skeptical of using MAE to compare the different approaches. Furthermore, it can be observed that in Table 2, MAE has a large standard deviation compared to the R^2 values. Previous studies have also argued against using MAEs to

interpret social media predictions of elections, because they are less sensitive to individual studies with large residuals. Particularly in situations where win margins and final party rankings matter (especially in multi-party races), it is important to have a measure that is more sensitive to individual errors. Accordingly, it is recommended that the results from Table 2 and onwards should be interpreted using the R^2 estimates, which have manifested low variance and are thus more stable across the different studies. With reference to H2 and H3, the R^2 values suggest that structural approaches do substantially better than sentiment (H2 is supported), and combination methods have a big advantage over sentiment alone. However, the difference between the three predictors was not significant when predictive accuracy was measured with R^2 ($F(2, 198) = 0.210$, $p = 0.116$) (H3 is not supported).

Table 2. Covariate-adjusted means of sentiment, structure, and combination approaches.

Predictors	MAE (%)			R^2		
	Mean	SD	N	Mean	SD	N
Structure	5.027 [a]	0.404	103	0.605 [b]	0.031	114
Sentiment	4.397 [a]	0.401	104	0.492 [b]	0.044	65
Sentiment and Structure	2.929 [a]	0.842	25	0.621 [b]	0.063	26

[a] Covariates appearing in the model are evaluated at the following values: Internet penetration (%) = 65.2504, electoral system = 0.49, press freedom = 64.8276, democracy score = 7.6360. [b] Covariates appearing in the model are evaluated at the following values: Internet penetration (%) = 67.7971, electoral system = 0.20, press freedom = 65.3854, democracy score = 7.2079.

4.3. Comparing the Predictive Accuracy of Social Media for Different Kinds of Outcomes

To compare the predictive accuracy of social media facing different kinds of outcomes (vote share, seat share, public support, winner), ANCOVA tests were also conducted in the two datasets controlling for the four contextual variables. Once again, R^2 has a lower standard deviation and is more sensitive to measuring the statistically significant differences between the predictive accuracy of the difference models ($F(3, 197) = 2.690$, $p = 0.047$). No significant difference was detected among the models' MAEs ($F(2, 225) = 0.855$, $p = 0.427$).

Seat share explains 20% more variance on average as compared to public support. Comparing the highest average R^2 from Tables 2 and 3, we observe that social media-based metrics do not, on average, outperform seat share as a predictor of election outcomes. It is harder to interpret the performance of different metrics using MAEs, as seat share and public support have nearly the same average MAEs.

Table 3. Covariate-adjusted means for different types of election and poll-based predictions.

Type of Outcome	MAE (%)			R^2		
	Mean	SD	N	Mean	SD	N
Vote share	4.796 [a]	0.339	157	0.595 [b]	0.027	141
Seat share	3.943 [a]	1.670	6	0.716 [b]	0.127	7
Winner				0.102 [b]	0.224	2
Public support (in polls)	3.937 [a]	0.539	69	0.510 [b]	0.043	55

[a] Covariates appearing in the model are evaluated at the following values: Internet penetration (%) = 65.2504, electoral system = 0.49, press = 64.8276, democracy score = 7.6360. [b] Covariates appearing in the model are evaluated at the following values: internet penetration (%) = 67.7971, electoral system = 0.20, press freedom = 65.3854, democracy score = 7.2079.

4.4. Comparing the Predictive Accuracy of Social Media from Various Sources

To answer H4 and compare the predictive accuracy of social media data from different platforms, ANCOVA tests were conducted while controlling for the four contextual variables. Significant effects have been observed for both MAE and R^2 (($F(5, 222) = 3.016$, $p = 0.012$ for MAE; $F(5, 195) = 15.992$, $p = 0.000$ for R^2).

As demonstrated in Table 4, the highest predictive accuracy (as measured with R^2) occurs when social media data come from blogs. Once again, these estimates are more stable, with smaller standard deviations over a similar sample size as compared to the average MAEs. In comparing MAEs, blogs have the second highest predictive accuracy. With an $N = 2$, little can be said about the success of using multiple platforms despite the low MAE, unless more studies attempt to predict election outcomes after compiling social media data from multiple sources (H4 is not supported).

Table 4. Covariate-adjusted means of predictions across data sources.

Data Source	MAE (%)			R^2		
	Mean	SD	N	Mean	SD	N
Twitter	4.355 [a]	0.274	210	0.548 [b]	0.027	108
Facebook	7.149 [a]	1.205	11	0.707 [b]	0.050	33
Forum	6.510 [a]	2.320	3	0.774 [b]	0.121	5
Blog	1.366 [a]	2.325	3	0.981 [b]	0.067	18
YouTube	9.939 [a]	2.325	3	0.261 [b]	0.067	18
Multiple platforms	0.802 [a]	2.822	2	0.365 [b]	0.096	23

[a] Covariates appearing in the model are evaluated at the following values: Internet penetration (%) = 65.2504, electoral system = 0.49, press freedom = 64.8276, democracy score = 7.6360. [b] Covariates appearing in the model are evaluated at the following values: Internet penetration (%) = 67.7971, electoral system = 0.20, press freedom = 65.3854, democracy score = 7.2079.

4.5. Comparing Predictive Power across Different Levels of Political Democracy

To answer RQ1 (a) and compare social media data's predictive power across different levels of political democracy, an ANCOVA test of the predictive accuracy on political regimes (recoded based on the democracy score) was conducted in the two datasets controlling for the three social media-related variables. As shown in Table 5, a significant effect was observed among the R^2 ($F(3, 198) = 3.061$, $p = 0.029$). However, it is noteworthy that the Levene's test shows the assumption of equal variances may be violated ($F(3, 201) = 3.182$, $p = 0.025$). No significant effect was observed among the MAEs (($F(2, 226) = 0.844$, $p = 0.431$).

Table 5. Covariate-adjusted means predictive power across levels of political democracy.

Political Democracy	MAE (%)			R^2		
	Mean	SD	N	Mean	SD	N
Authoritarian Regime				0.597 [b]	0.069	30
Hybrid Regime	4.644 [a]	0.790	29	0.825 [b]	0.101	11
Flawed Democracy	4.069 [a]	0.438	91	0.599 [b]	0.055	40
Full Democracy	4.852 [a]	0.408	112	0.533 [b]	0.029	124

[a] Covariates appearing in the model are evaluated at the following values: data source = 1.21, predictor =31.66, predicted result = 1.38. [b] Covariates appearing in the model are evaluated at the following values: data source = 2.39, predictor = 31.57, predicted result = 1.39.

4.6. Comparing Predictive Power across Electoral Systems

To answer RQ1 (b) and compare social media data's predictive power across different electoral systems, an ANCOVA test of the predictive accuracy on electoral system was conducted in the two datasets controlling for the three social media-related variables. As shown in Table 6, no significant effect were observed in either of the two datasets (($F(1, 227) = 0.912$, $p = 0.340$ for MAE; $F(1, 200) = 0.670$, $p = 0.414$ for R^2). Again, the Levene's test shows the assumption of equal variances may be violated ($F(1, 203) = 0.3731$, $p = 0.055$).

Table 6. Covariate-adjusted means of predictive power across electoral systems.

Electoral Systems	MAE (%)			R^2		
	Mean	SD	N	Mean	SD	N
Non-proportional Representation	4.268 [a]	0.375	119	0.581 [b]	0.025	164
Proportional Representation	4.783 [a]	0.385	113	0.533 [b]	0.052	41

[a] Covariates appearing in the model are evaluated at the following values: data source = 1.21, predictor =31.66, predicted result = 1.38. [b] Covariates appearing in the model are evaluated at the following values: data source = 2.39, predictor = 31.57, predicted result = 1.39.

5. Discussion

In this meta-analysis, we compared the predictive power of social media analytics across different approaches, platforms, and contexts. One of the unexpected findings was that depending on whether studies report MAE-based estimates or R^2 estimates, a very different picture of the efficacy of different methods emerges. While R^2 based measures showed most stability and can be interpreted as higher recall or explainability of the data, MAE-based estimates can be used in cases where errors are symmetrical, e.g., in sentiment analyses, or in two-party races, wherein precision is of importance. Machine learning-based sentiment analysis tends to produce predictions with higher precision than lexicon-based approaches; however they typically explain less variance.

Our findings indicate the limitations of applying generic sentiment tools to mining political opinions, and applying dictionaries developed in the 1980s to analyze the present-day language of social media, which can falsely detected positive sentiments where there is sarcasm, and hence can lead to erroneous predictions [36]. In addition, lexica are designed for standard English, but many messages on the platforms like Twitter are written in informal versions of English, which include alternatively spelled words and emoticons. Informal language cues are potentially useful signals, which are usually ignored in traditional methods of sentiment analysis. On the other hand, a supervised learning approach, which trains sentiment models on a small set of hand-annotated political social media messages, yields much better predictions by inferring sentiments from otherwise neutral words used in context. Furthermore, studies have suggested that discarding negative posts and instead focusing on the positive tweets can help to filter out a large part of the noise from election-related content on social media [24].

Although individual sentiment is important, our results reflect the theoretical importance of interactions in the formation of public opinion. Combinations of structural features and machine learning-based sentiment analyses provide the most accurate predictions among all the individual approaches considered. This means that sentiment analyses work best when they are combined with structural features to model the diffusion of opinion in a social network, in terms of the reach of the authors, their interaction patterns, and their importance as influencers within their own communities of followers. Although most studies have relied on simple counts of relational (followers, friends) or interactive connections (retweet, reply, likes, etc.), these can possibly be gamed by astroturfing or by heavy users, spammers, and propagandists [64]. They may also show attention spikes because of news cycles. Instead, we recommend that more sophisticated measures of author importance; e.g., centrality and PageRank measures, should be adopted to provide more accurate measures of online communication structures. Structural features can also capture the density of online discussions. More decentralized networks have more active users and thus wider outreach to a larger potential voter base [24]. Structural features have been found to be useful to dampen the estimation effects associated with national parties that are over-represented on social media, or regional parties which may be popular online. It is also important to note that most studies reviewed rely on the data that are not demographically representative of the target general population; thus, applying appropriate weights could potentially improve the quality of predictions, when it is feasible.

Interestingly, blogs are found to outperform other platforms in predicting political outcomes—potentially because blogs provide politically focused data sources when compared to typical social

media posts. Furthermore, political blogs are often more similar to broadcast media outlets (some are indeed run by them) than to individual user profiles such as those on Facebook and Twitter, and they are typically run by key opinion leaders and other influential figures. This is similar to discussion forums which are populated by politically interested citizens and where discussion posts have been shown to be representative of opinion trends even in the absence (or lack of information about) demographic representativeness of the discussants [32].

Although we expected that methods combining multiple sources of data would have the best performances (and in terms of MAE, they do), the number of studies using multiple data sources or using blog data as predictors is not big enough to make any definitive conclusions. Furthermore, with the increasing instances of privacy breaches and the generally unmoderated, toxic nature of political discourse on public social media platforms, there is a noticeable shift of users towards more closed instant messaging app ecosystems characterized by end-to-end encryption. We thus see greater challenges ahead for opinion mining research, as social media platforms continue their fragmentation into those where user posts and interactions are still publicly visible (and with API support) and those where conversations happen in private, encrypted digital chambers. As everyday conversations among citizens become less visible (at least to most researchers) those profiles and messages that remain public are likely to be more strategically crafted and even less representative of the average citizen.

Another surprising finding is that predictions based on R^2 hold the most predictive power in hybrid regimes, which combine the characteristics of democratic and autocratic regimes. While elections may be held at regular intervals, citizens' knowledge of and participation in offline political activities may be heavily curtailed, and the press freedom and the freedom of expression online is often restricted. Still, in most hybrid regime-countries, social media platforms tend to be less controlled than the traditional media, which are usually under tight government control and sometimes even prevented from reporting opinion poll data [60]. Thus, social media posts may become more useful as "sensors" of public opinion, particularly as more traditional means may not be available. However, we acknowledge that because of a relatively small number of studies from hybrid regime countries in our sample making strong inferences would be premature at this point.

Although this study is one of the first systematic reviews of social media-based predictions, it is important to note some of its limitations. First, we have included both electoral outcomes and traditional polls as the benchmarks for comparison. This follows the assumption and general wisdom from many of the studies we examined, wherein traditional polls were expected to be more sensitive to public opinion than social media trends, which are often noisy and easy to manipulate. However, recent events, such as poll-based predictions for the 2016 General Election in the United States, give pause to even how traditional polls should be interpreted when there is a mismatch between the type of outcome predicted (i.e., salience and popularity) and the outcome necessary to form government (i.e., a majority in the electoral college). Second, since social media-based predictions are still in the early stages, there are insufficient data to produce reliable estimates across different analytical categories. Third, most of the studies rely on data from social network sites, especially Twitter, while the number of estimates for other social media platforms (e.g., blogs, forums) is quite limited, making a systematic and detailed comparison more difficult. Fourth, several studies did not report their data sizes [34,65], while the other studies reported a wide range of data sizes—ranging from thousands to hundreds of millions [66,67], making examination of effect of data sizes on social media data's predictive power of public opinion difficult. Fifth, the predictive power is reported in a range of formats—MAE, RMSE, correlation coefficients, regression beta, R^2, offset error, and race-based percentage, making a systematic comparison difficult. We thus need a more standardized way of reporting data collection methods and statistical estimates of predictive power. Lastly, we were unable to systematically explore the temporal dimension in opinion mining, which is one of the key advantages of social media data and has been shown to affect the quality of election forecasts, but has not been commonly utilized in the extant research [24].

6. Conclusions

Digital traces of human behavior now enable the testing of social science theories in new ways, and are also likely to pave the way for new theories in social science. Both in terms of its nature and scale, user-generated social media data is dramatically different from the existing market information regime data, which is constructed to serve a specific purpose and typically relies on smaller samples of unconnected citizens [25]. As new studies and new methods for social media-based predictions proliferate, it is important to consider both how data are generated, and how they are analyzed, and consolidate a validated framework for data analysis. As Kuhn argued, there is a close, intimate link between the scientific puzzle and the methodology designed to solve it [68]. Social media analytics will not replace survey-based public opinion studies, but they do offer additional data sources and tools for improving our understanding of public opinion and political behavior, while at the same time changing the norms around voter engagement, mobilization, and preference elicitation. Among other things, we see the potential of social media data to provide deeper insights regarding the opinion dynamics and the role of opinion leaders in the process of public opinion formation and change. Furthermore, the problems that plague our digital ecologies today, such as filter bubbles, disinformation, incivility, and hate speech, can all be better understood using social media data and computational methods, than with survey-based research. Still, we note a general lack of theoretically-informed work in the field—most studies report predictions without paying sufficient attention to the underlying mechanisms and processes. We are also concerned about the future availability of social media data, as citizens switch to more private, encrypted instant messaging apps and social media companies further restrict access to user data because of legal, commercial, and privacy concerns. Thus, policymakers should create legal and regulatory environments that will promote access to social media data for the research community in ways that guarantee strong privacy protection rights for the users.

References

1. King, G. Ensuring the data-rich future of the social sciences. *Science* **2011**, *331*, 719–721. [CrossRef] [PubMed]
2. Lazer, D.; Pentland, A.S.; Adamic, L.; Aral, S.; Barabasi, A.L.; Brewer, D.; Gutmann, M. Life in the network: The coming age of computational social science. *Science* **2009**, *323*, 721–723. [CrossRef]
3. boyd, D.; Crawford, K. Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Inf. Commun. Soc.* **2012**, *15*, 662–679. [CrossRef]
4. Salganik, M.J. *Bit by Bit: Social Research in the Digital Age*; Princeton University Press: Princeton, NJ, USA, 2018.
5. Choi, H.; Varian, H. Predicting the present with Google Trends. *Econ. Rec.* **2012**, *88*, 2–9. [CrossRef]
6. Manovich, L. Trending: The promises and the challenges of big social data. In *Debates in the Digital Humanities*; Gold, M.K., Ed.; The University of Minnesota Press: Minneapolis, MN, USA, 2011; pp. 460–475.
7. González-Bailón, S.; Wang, N.; Rivero, A.; Borge-Holthoefer, J.; Moreno, Y. Assessing the bias in samples of large online networks. *Soc. Netw.* **2014**, *38*, 16–27. [CrossRef]
8. Jungherr, A.; Jürgens, P.; Schoen, H. Why the Pirate party won the German election of 2009 or the trouble with predictions: A response to Tumasjan, A., Sprenger, T.O., Sander, P.G., & Welpe, I.M. "Predicting elections with Twitter: What 140 characters reveal about political sentiment.". *Soc. Sci. Comput. Rev.* **2012**, *30*, 229–234. [CrossRef]
9. Bucher, T.; Helmond, A. The Affordances of Social Media Platforms. In *The SAGE Handbook of Social Media*; Burgess, J., Marwick, A., Poell, T., Eds.; Sage Publications: New York, NY, USA, 2018; pp. 233–253.

10. Ratkiewicz, J.; Conover, M.; Meiss, M.R.; Gonçalves, B.; Flammini, A.; Menczer, F. Detecting and Tracking Political Abuse in Social Media. In Proceedings of the 20th international conference companion on World Wide Web, Hyderabad, India, 28 March–1 April 2011.

11. Gayo-Avello, D. A meta-analysis of state-of-the-art electoral prediction from Twitter data. *Soc. Sci. Comput. Rev.* **2013**, *31*, 649–679. [CrossRef]

12. Bourdieu, P. Public opinion does not exist. In *Communication and Class Struzgle*; Mattelart, A., Siegelaub., S., Eds.; International General/IMMRC: New York, NY, USA, 1979; pp. 124–310.

13. Lazarsfeld, P.F. Public opinion and the classical tradition. *Public Opin. Q.* **1957**, *21*, 39–53. [CrossRef]

14. Katz, E. The two-step flow of communication: An up-to-date report on an hypothesis. *Public Opin. Q.* **1957**, *21*, 61–78. [CrossRef]

15. Katz, E.; Lazarsfeld, P.F. *Personal Influence: The Part Played by People in the Flow of Mass Communications*; Free Press: New York, NY, USA, 1955.

16. Ceron, A.; Curini, L.; Iacus, S.M.; Porro, G. Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media Soc.* **2014**, *16*, 340–358. [CrossRef]

17. Espeland, W.N.; Sauder, M. Rankings and reactivity: How public measures recreate social worlds. *Am. J. Sociol.* **2007**, *113*, 1–40. [CrossRef]

18. Berinsky, A.J. The two faces of public opinion. *Am. J. Political Sci.* **1999**, *43*, 1209–1230. [CrossRef]

19. Blumer, H. Public opinion and public opinion polling. *Am. Sociol. Rev.* **1948**, *13*, 542–549. [CrossRef]

20. Anstead, N.; O'Loughlin, B. Social media analysis and public opinion: The 2010 UK general election. *J. Comput. Mediat. Commun.* **2015**, *37*, 204–220. [CrossRef]

21. Bryce, J. *The American Commonwealth*; Macmillan: London, UK, 1888.

22. Herbst, S. Public opinion infrastructures: Meanings, measures, media. *Political Commun.* **2001**, *18*, 451–464. [CrossRef]

23. Kwon, K.H.; Moon, S.I.; Stefanone, M.A. Unspeaking on Facebook? Testing network exposure effects on self-censorship of political expressions in social network sites. *Qual. Quant.* **2015**, *49*, 1417–1435. [CrossRef]

24. Jaidka, K.; Ahmed, S.; Skoric, M.M.; Hilbert, M. Predicting elections from social media: A three-country, three-method comparative study. *Asian J. Commun.* **2019**, *29*, 252–273. [CrossRef]

25. Webster, J.G. The duality of media: A structurational theory of public attention. *Commun. Theory* **2011**, *21*, 43–66. [CrossRef]

26. Anand, N.; Peterson, R.A. When market information constitutes fields: Sensemaking of markets in the commercial music industry. *Organ. Sci.* **2000**, *11*, 270–284. [CrossRef]

27. Allport, F.H. Toward a science of Public opinion. *Public Opin. Q.* **1937**, *1*, 7–23. [CrossRef]

28. Fox, S. Four in Ten Seniors Go Online. Available online: http://www.pewinternet.org/2010/01/13/four-in-ten-seniors-go-online (accessed on 5 February 2020).

29. Wei, L.; Hindman, D.B. Does the digital divide matter more? Comparing the effects of new media and old media use on the education-based knowledge gap. *Mass Commun. Soc.* **2011**, *14*, 216–235. [CrossRef]

30. Tumasjan, A.; Sprenger, T.O.; Sandner, P.G.; Welpe, I.M. Predicting elections with Twitter: What 140 characters reveal about political sentiment. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC, USA, 23–26 May 2010.

31. Zaller, J. The Nature and Origins of Mass Opinion. Cambridge University Press: Cambridge, UK, 1992.

32. González-Bailón, S.; Banchs, R.E.; Kaltenbrunner, A. Emotions, public opinion, and US presidential approval rates: A 5-year analysis of online political discussions. *Hum. Commun. Res.* **2012**, *38*, 121–143. [CrossRef]

33. Ibrahim, M.; Abdillah, O.; Wicaksono, A.F.; Adriani, M. Buzzer Detection and Sentiment Analysis for Predicting Presidential Election Results in a Twitter Nation. In Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW), Atlantic City, NJ, USA, 14–17 November 2015. [CrossRef]

34. Li, H.L.; Ng, V.T.; Shiu, S.C. Predicting short interval tracking polls with online social media. In Proceedings of the 17th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Whistler, UK, 27–29 June 2013. [CrossRef]

35. Contractor, D.; Faruquie, T.A. Understanding election candidate approval ratings using social media data. In Proceedings of the 22nd International Conference on World Wide Web companion, Rio de Janeiro, Brazil, 18–22 May 2013. [CrossRef]

36. O'Connor, B.; Balasubramanyan, R.; Routledge, B.R.; Smith, N.A. From tweets to polls: Linking text sentiment to public opinion time series. In Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, Washington, DC, USA, 23–26 May 2010.

37. Wilson, T.; Wiebe, J.; Hoffmann, P. Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, Vancouver, BC, Canada, 6–8 October 2005.

38. Chung, J.E.; Mustafaraj, E. Can collective sentiment expressed on twitter predict political elections? In Proceedings of the 25th AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 25–30 January 2011.

39. Esuli, A.; Sebastiani, F. SentiWordNet: A High-Coverage Lexical Resource for Opinion Mining. Available online: http://nmis.isti.cnr.it/sebastiani/Publications/2007TR02.pdf (accessed on 5 February 2020).

40. Gayo-Avello, D. Don't turn social media into another 'Literary Digest' poll. *Commun. ACM* **2011**, *54*, 121–128. [CrossRef]

41. Beauchamp, N. Predicting and interpolating state-level polls using Twitter textual data. *Am. J. Political Sci.* **2017**, *61*, 490–503. [CrossRef]

42. Dwi Prasetyo, N.; Hauff, C. Twitter-based election prediction in the developing world. In Proceedings of the 26th ACM Conference on Hypertext & Social Media, Guzelyurt, TRNC, Cyprus, 2–4 September 2015. [CrossRef]

43. Kalampokis, E.; Karamanou, A.; Tambouris, E.; Tarabanis, K.A. On predicting election results using Twitter and Linked Open data: The case of the UK 2010 election. *J. Univers. Comput. Sci.* **2015**, *23*, 280–303.

44. Monti, C.; Zignani, M.; Rozza, A.; Arvidsson, A.; Zappella, G.; Colleoni, E. Modelling political disaffection from twitter data. In Proceedings of the 2nd International Workshop on Issues of Sentiment Discovery and Opinion Mining, Chicago, IL, USA, 11 August 2013.

45. Sharma, P.; Moh, T.S. Prediction of Indian election using sentiment analysis on Hindi Twitter. In Proceedings of the 2016 IEEE International Conference on Big Data, Washington, DC, USA, 5–8 December 2016. [CrossRef]

46. Huberty, M.E. Multi-cycle forecasting of congressional elections with social media. In Proceedings of the 2nd workshop on Politics, Elections and Data, San Francisco, CA, USA, 28 October 2013. [CrossRef]

47. Xie, Z.; Liu, G.; Wu, J.; Wang, L.; Liu, C. Wisdom of fusion: Prediction of 2016 Taiwan election with heterogeneous big data. In Proceedings of the 13th International Conference on Service Systems and Service Management (ICSSSM), Kunming, China, 24–26 June 2016. [CrossRef]

48. Mejova, Y.; Srinivasan, P.; Boynton, B. GOP primary season on twitter: Popular political sentiment in social media. In Proceedings of the Sixth ACM International Conference on Web Search and Data Mining, Rome, Italy, 4–8 February 2013. [CrossRef]

49. Livne, A.; Simmons, M.P.; Adar, E.; Adamic, L.A. The party is over here: Structure and content in the 2010 election. In Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, Barcelona, Spain, 17–21 July 2011.

50. Vepsäläinen, T.; Li, H.; Suomi, R. Facebook likes and public opinion: Predicting the 2015 Finnish parliamentary elections. *Gov. Inf. Q.* **2017**, *34*, 524–532. [CrossRef]

51. Metaxas, P.T.; Mustafaraj, E.; Gayo-Avello, D. How (not) to predict elections. In Proceedings of the 2011 IEEE Third International Conference on Social Computing (SocialCom), Boston, MA, USA, 9–11 October 2011. [CrossRef]

52. MacWilliams, M.C. Forecasting congressional elections using facebook data. *PS Political Sci. Politics* **2015**, *48*, 579–583. [CrossRef]

53. Pimenta, F.; Obradovic, D.; Dengel, A. A comparative study of social media prediction potential in the 2012 US republican presidential preelections. In Proceedings of the 3rd International Conference on Cloud and Green Computing, Dublin, Ireland, 27–29 June 2013. [CrossRef]

54. Gayo-Avello, D.; Metaxas, P.T.; Mustafaraj, E.; Strohmaier, M.; Schoen, H.; Gloor, P.; Tarabanis, K. Understanding the predictive power of social media. *Internet Res.* **2013**, *23*, 544–559. [CrossRef]

55. Barclay, F.P.; Pichandy, C.; Venkat, A.; Sudhakaran, S. India 2014: Facebook 'like'as a predictor of election outcomes. *Asian J. Political Sci.* **2015**, *23*, 134–160. [CrossRef]

56. Williams, C.; Gulati, G. What is a social network worth? Facebook and vote share in the 2008 presidential primaries. In Proceedings of the 2008 Annual Meeting of the American Political Science Association, Boston, MA, USA, 6–27 June 2008.

57. Freeman, L.C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1979**, *1*, 215–239. [CrossRef]

58. Page, L.; Brin, S.; Motwani, R.; Winograd, T. The PageRank Citation Ranking: Bringing Order to the Web. Available online: http://homepages.dcc.ufmg.br/~{}nivio/cursos/ri11/sources/pagerank.pdf (accessed on 5 February 2020).

59. Cameron, M.P.; Barrett, P.; Stewardson, B. Can social media predict election results? Evidence from new zealand. *J. Political Mark.* **2016**, *15*, 416–432. [CrossRef]

60. Skoric, M.M.; Poor, N.D.; Achananuparp, P.; Lim, E.P.; Jiang, J. Tweets and votes: A study of the 2011 Singapore General Election. In Proceedings of the 45th Hawaii International Conference on System Sciences, Maui, HI, USA, 4–7 January 2012. [CrossRef]

61. DiGrazia, J.; McKelvey, K.; Bollen, J.; Rojas, F. More tweets, more votes: Social media as a quantitative indicator of political behavior. *PLoS ONE* **2013**, *8*, e79449. [CrossRef] [PubMed]

62. Campbell, J.E. Issues in presidential election forecasting: Election margins, incumbency, and model credibility. *PS: Political Sci. Politics* **2014**, *47*, 301–303. [CrossRef]

63. Krasno, J.S. *Challengers, Competition, and Reelection: Comparing House and Senate Elections*; Yale University Press: New Haven, CT, USA, 1994.

64. Metaxas, P.T.; Mustafaraj, E. Social media and the elections. *Science* **2012**, *338*, 472–473. [CrossRef] [PubMed]

65. Franch, F. Wisdom of the Crowds: 2010 UK election prediction with social media. *J. Inf. Technol. Politics* **2013**, *10*, 57–71. [CrossRef]

66. Gayo-Avello, D.; Metaxas, P.T.; Mustafaraj, E. Limits of electoral predictions using twitter. In Proceedings of the ICWSM, Barcelona, Spain, 17–21 July 2011.

67. Gaurav, M.; Srivastava, A.; Kumar, A.; Miller, S. Leveraging candidate popularity on Twitter to predict election outcome. In Proceedings of the 7th Workshop on Social Network Mining and Analysis, Chicago, IL, USA, 21–23 November 2013. [CrossRef]

68. Kuhn, T.S. *The Structure of Scientific Revolutions*; University of Chicago Press: Chicago, IL, USA, 1962.