How Accurate Are Survey Responses on Social Media and Politics?*

Andrew Guess† Princeton University

Kevin Munger‡ New York University

Jonathan Nagler§ New York University

Joshua Tucker¶ New York University

June 12, 2018

Abstract

How accurate are survey-based measures of social media use, in particular about political topics? We answer this question by linking original survey data collected during the U.S. 2016 election campaign with respondents' actual social media activity. We use supervised machine learning to classify whether this Twitter and Facebook account data is content related to politics. We then benchmark our survey measures on frequency of posting about politics and the number of political figures followed. We find that, on average, our self-reported survey measures tend to correlate with actual social media activity. At the same time, we also find a worrying amount of individual-level discrepancy and problems related to extreme outliers. Our recommendations are twofold. First, for survey questions about social media use to provide respondents with options covering a wider range of activity, especially in the long tail. Second, for survey questions to include specific content and anchors defining what it means for a post to be "about politics."

*This study has been approved by the New York University Institutional Review Board (IRB-12-9058, IRB-FY2017-150). Thanks to Samantha Luks at YouGov for facilitating survey data collection, Yvan Scher for research assistance, and Jona Ronen for development and technical assistance.

†Assistant Professor, Department of Politics and Woodrow Wilson School, Princeton University. Email: aguess@princeton.edu

‡Ph.D. Student, Wilf Family Department of Politics and SMaPP Lab, New York University. Email: km2713@nyu.edu

§Professor, Wilf Family Department of Politics and SMaPP Lab, New York University. Email: jonathan.nagler@nyu.edu

¶Professor, Wilf Family Department of Politics and SMaPP Lab, New York University. Email: joshua.tucker@nyu.edu

1 Introduction

In 2016, the Pew Research Center surveyed Americans about their social media use during the election campaign. One question asked: “In the past week, did you yourself use a social networking site to share news or information about the presidential campaign or candidates, such as by posting or replying to or commenting on a post?”[1] This is a vital question for understanding the changing ways in which people consume and interact with the news about political developments (e.g., Mutz and Young 2011; Duggan 2015; Duggan and Smith 2016; Gottfried and Shearer 2016). Yet, so far, we know very little about how accurately people’s responses to such a question reflect their actual behavior on social media.

This paper provides the first systematic attempt to benchmark such questions, focusing on a set of items developed to study Americans’ social media use during the 2016 U.S. presidential election campaign. It is important to know how well such questions perform when compared against a known “ground truth,” as well as whether there are systematic biases that could potentially be corrected. In the future, this will aid both studies of social media use as well as research on the effects of social media on politics (e.g., Bartels 1993; Bohdanova 2014). At the same time, we do not claim that the current paper is the final word on the subject; there are thorny measurement issues involved, and studies of social media are inevitably aiming at an ever-shifting target (Karpf 2012).

We combine survey data from a panel fielded over several waves in 2016 with data on respondents’ Twitter and Facebook posts, classified as political or not using supervised learning techniques to generate measures of real-world posting behavior: both in general and on political topics. We find that, on average, there is a correspondence between the average frequency of tweets and Facebook posts and the self-reporting by our respondents. However, individual-level discrepancies are common, and there remain persistent issues related to outliers and top-coding of response categories.

The next section briefly covers some related literature. We then provide an overview

[1] http://www.journalism.org/2016/02/04/the-2016-presidential-campaign-a-news-event-thats-hard-to-miss/

of our data sources, survey questions, and coding procedures. In Section 4, we present our results on tweeting and posting frequency. Section 5 discusses results on our survey measures of political follow networks. We conclude with a discussion of issues with the kinds of survey questions studied here.

2 Measuring Media Exposure and Social Networks

Social scientists have long been interested in the effects of exposure to information from the mass media and interpersonal interactions (Lazarsfeld, Berelson and Gaudet 1944; Iyengar and Kinder 1987; Mutz and Martin 2001). However, scholars have consistently found that their favorite tool in this context — self-reported survey data — is flawed, as individuals often misreport their true exposure to these information sources. This insight has taken creative research design, as objective measures of exposure to traditional mass media such as television and radio are rarely linked to individual-level survey data (Prior 2009, 2012; Scharkow 2016; Jerit et al. 2016).

The rapidly evolving and increasingly fragmented information environment adds urgency to the important task of benchmarking measures before widespread deployment (Niederdeppe 2014; Guess 2015; de Vreese and Neijens 2016; Taneja 2016). Simply put, can we trust responses to questions about what people do on social media sites? This paper provides the first comprehensive benchmarking test of self-reported social media exposure measures. By comparing self-reports of activity on Twitter and Facebook to observed activity, we are able to measure the accuracy of the self-reports that researchers have previously had to take at face value.

There is a long history of asking for self-reports about political activity, everything from "how often do you talk to a neighbor about politics" to whether or not the respondent voted or gave money to a candidate (Prior 2009; Anderson and Silver 1986; Schwarz 1999). Responses to a large set of behavior questions cannot be verified, and thus scholars have had to rely on self-reports for both simply tabulating the behavior and for trying to understand the correlates of the reported behavior (Romantan et al. 2008). However, social media data, by nature of its "digital footprints," gives us the opportunity to observe political behavior.[2] That means we can, in principle, see if self-reports of activity are valid. But, since observation of the actual behavior can be difficult, we want to know if the self-report corresponds to observed behavior, and if we can rely on surveys when direct observation of behavior is not possible.

The literature on the accuracy of behavioral self-reports suggests a core set of expectations (e.g., Stone et al. 1999). Questions about sensitive behaviors are known to produce systematic underreporting, but it is unlikely that most respondents consider their social media activity — by definition public or semi-public in nature — to fall into this category. As a result, we are more concerned about patterns of overreporting. There are a number of possible mechanisms for systematic inflation of self-reported behavioral tendencies, including imperfect recall (Revilla, Ochoa and Loewe 2017) or inattentive responding (McKibben and Silvia 2017), faulty calibration to population frequencies (Prior 2009), and social desirability (Krumpal 2013). Some of these errors may cancel out in the aggregate, while others (such as social desirability) lead in a predicted direction. This implies that survey measures could be substantially correlated with ground truth even in the presence of substantial biases at the individual level. Perhaps for this reason, scholars have sought to identify demographic or political correlates of over- or underreporting in surveys. Some hypothesized factors, such as the role of political interest and partisan strength in driving earnest or expressive respondent enthusiasm, are more plausible than others (Prior 2013); predictors such as age or gender may be associated with overreporting for some questions in certain contexts, but it is not clear whether effects of this kind should be expected to generalize.

Taken together, this discussion implies that the misreporting we uncover should tend to

[2] At the same time, the issue of what defines "political" behavior in the context of Facebook is just beginning to receive scholarly attention. Below, we discuss the issue of subjectivity of what is and is not political behavior.

inflate rather than deflate social media use frequency, and that this tendency to overreport may be associated with measures of political interest and partisan allegiance (Prior 2009). At the same time, we have some expectation that our respondents' self-selection into social media use in the first place may be related some of the same characteristics. Thus we are left with a puzzle: Individuals who we would expect to be overrepresented in social media discussions about politics may be the very types of people who are predisposed to overreport that activity, due to partisan misperceptions, flawed recall, or other cognitive biases. It is unclear whether this might cause more extreme errors in self-reporting or, instead, whether social media use could be a case in which some subjects' actual behaviors happen to match their generally inflated response patterns.

We also face a classic measurement issue in benchmarking posts about politics. We can only truly determine if respondents' reports of behavior of a specific type are accurate if the behavior we are asking about can be objectively measured. If we ask people whether or not they voted, there is a true answer. But when we ask if they posted 'about politics' there is an inherent subjectivity to the question. Even if all respondents had perfect recall and desired to be as accurate as possible, their varying conceptions the definition of a "political" post would present a challenge to any benchmarking effort. To use the example developed in detail in Settle (2018), consider a Facebook post in which someone says that they're having lunch at Chik-Fil-A. This appears to be a non-political post, but during a period around 2012, there was considerable controversy about Chik-Fil-A's funding anti-LGBT organizations. Some liberals announced a boycott of Chik-Fil-A, and some conservatives began a counter-boycott with associated posts on their Facebook feeds. Many other people continued to eat (or not) at Chik-Fil-A and post the occasional picture of a fried chicken sandwich.

If our coders do not interpret posts about where someone ate lunch as political, then the supervised machine learning model we use will not be likely to classify posts with a picture of a chicken sandwich at Chik-Fil-A as political. But a conservative activist who answers our survey question may think that their posting the sandwich is political, and we would thus conclude that they have over-reported their true rate of posting about politics. On the other hand, if our coders are attuned to the politics surrounding Chik-Fil-A, they may think the posting of a chicken sandwich is political. This would lead us to be likely to classify such postings as political, and we would conclude that respondents who do not view postings of their lunch decisions as political were under-reporting their rate of posting about politics.[3]

We note that the high degree of inter-coder reliability we achieved suggests that there is in fact generally a shared sense of what it means for a post to be "about politics." But this may be because of the homogeneity of our undergraduate set of coders. As a result, for the survey questions we study, we refrain from claiming that gaps between our baseline measures and self-reports of posting political content represent "lies." But they do represent the discrepancy between self-reports and what our coders believe to be posts about politics. We note that this particular issue could be resolved by asking respondents a more specific question that gives them guidance about what it means for a post to be "about politics": we could, for instance, spell out that we mean a post mentioning a campaign or candidate or public policy issue.

3 Data Overview

3.1 Panel Survey

The self-reports against which we benchmark objective measures comes from a panel survey conducted during the 2016 U.S. presidential election.[4] These surveys were designed explicitly to understand the relationship between social media use and changes in political beliefs and preferences; this was accomplished by pairing a panel survey design with information about

[3]The contours of this subjectivity have only begun to be explored by political scientists. Settle (2018) asks a sample of undergrads and a sample of MTurk coders to label Facebook posts as "political" or not. She reports that "Users who say they more frequently encounter political content on their News Feeds do in fact have a broader conception of content that could be considered to be political" (ch. 5). Settle's argument is that the more frequently people use Facebook, the more likely they are to identify posts as being political. We do not try to test this claim directly, but our data in Appendix C are consistent with it.

[4]The survey was designed by the authors and conducted by the polling firm YouGov.

respondents' social media accounts.

The U.S. presidential election survey took place over the course of three waves: the first was April 9–May 1, 2016 (3,500 respondents); the second was September 9–October 9 (2,635 respondents); and the third was October 25–November 7 (2,628 respondents). The first wave of the survey was estimated by YouGov to take 15 minutes to complete, and contains a wealth of demographic, social media use and political preference questions. This paper focuses on questions in which respondents self-report several aspects of their use of social media.

In wave 1, respondents were asked to select all of the social networking sites on which they had accounts (options provided: Twitter, Facebook, Instagram, LinkedIn, Snapchat, Other); if they indicated that they had a Facebook and/or Twitter account, we asked them about their frequency of posting on that platform both in general and specifically about politics. Response options ranged from "Never" to "Several times a day." We also asked respondents to describe the number and types of people they were connected to on the accounts they have — both in general (for Twitter only) and "elected officials, candidates for office, or other political figures."[5] Tables 1 and 2 show the raw responses in each category for these questions (for those who reported using the relevant social media platform). The modal response for the number of politicians followed on both Facebook and Twitter is zero; likewise, the modal respondent who uses either social media platform reports never posting about politics.

The sampling frame included members of the YouGov panel who had previously agreed to supply YouGov with their Twitter IDs. There were 2,163 respondents in the first wave (62% of the overall sample) who elected to enter some text when prompted for their Twitter ID. We were able to verify 1,816 of these with the Twitter API.[6] We scraped the tweets by these respondents on November 20, 2016.[7] We also retrieved the IDs of the accounts followed

[5]See Appendix D for full question wordings and response options.

[6]Some people simply provided something other than a valid Twitter ID to YouGov, and it is possible that some of these accounts were deleted by their owners or banned/suspended by Twitter.

[7]Because we only accessed each respondents' tweets at one point, the number of tweets per subject in our

Table 1: Self-Reported Social Media Posting Behavior (Counts and Proportions)

	Twitter	Twitter (politics)	Facebook	Facebook (politics)
Never	219 (0.09)	992 (0.44)	102 (0.04)	874 (0.33)
Less often	673 (0.29)	473 (0.21)	329 (0.12)	562 (0.21)
Every few weeks	361 (0.15)	257 (0.11)	362 (0.13)	373 (0.14)
1 to 2 days a week	325 (0.14)	178 (0.08)	391 (0.14)	245 (0.09)
3 to 6 days a week	195 (0.08)	95 (0.04)	356 (0.13)	152 (0.06)
About once a day	257 (0.11)	124 (0.05)	453 (0.17)	172 (0.07)
Several times a day	304 (0.13)	160 (0.07)	711 (0.26)	256 (0.10)

Table 2: Self-Reported Social Media Following Behavior (Counts and Proportions)

	# politicians followed: Twitter	# politicians followed: FaceBook
0	1047 (0.47)	1245 (0.49)
1-10	892 (0.40)	1067 (0.42)
More than 10	269 (0.12)	254 (0.10)

by these respondents. Ultimately, we were able to collect a non-zero number of tweets from 1,421 respondents. This raises understandable concerns about sample selection. For a full discussion of the attrition process in our data, including descriptive statistics suggesting that it did not seriously bias the final sample, see Appendix B. These results are similar to the finding by Munger et al. (2016) that a sample of U.K. Twitter users who shared their account information with YouGov is representative of U.K. Twitter users generally.

We also collected Facebook profile data from our respondents. Using a web application that we developed, we asked respondents if they would be willing to supply information about their own past Facebook activity and told them that we could get this information directly from their Facebook accounts if they agreed. This was done via a separately administered survey question post-election that linked respondents to the web app connecting to the Facebook API. 1,221 of our respondents agreed to let us retrieve their Facebook information in return for compensation in the form of $5 in YouGov "points." Specifically, we requested

dataset is capped at 3,200, the maximum allowed by the Twitter API.

their public profile information, Timeline posts (including text and links if available), page likes, and what Facebook saves as religious and political views. If a respondent chose to log into Facebook after the survey prompt, they were asked what specific pieces of information they were willing to share. They could approve sharing all of the types of information, selectively approve only some of these types of information, or approve nothing.[8]

3.2 Coding Social Media Networks

To create an objective measure about the networks of our respondents, we used friend/follower data. The first question to benchmark — how many people do you follow on Twitter? — was straightforward to measure: Twitter API data always contains this information.[9]

A more challenging task was to create a measure for the question of how many *politicians* respondents follow on Twitter/Facebook. It was straightforward to create follower account lists for all respondents for whom we had a valid ID, but we needed to match these accounts against those known to be associated with politicians.[10]

We used two different sources for the lists of accounts used by politicians. For the Facebook data, we used information from the Sunlight Foundation.[11] This database contains the Facebook username of the account (either a normal, personal account or a public-facing page) belonging to each sitting member of Congress. For the Twitter data, we accessed public Twitter "lists" of members of Congress. We combined one list maintained by C-SPAN with a list of accounts that Twitter had "verified" as authentic. In both cases, we supplemented the lists with the accounts for prominent politicians not currently serving in Congress.

[8]No data on News Feed content or exposure was shared with researchers, since such access is no longer allowed via the API (for a rare example of data collected with the earlier permissions, see Wells and Thorson 2017). Data access was temporary and lasted only 2 months after permission was granted. All respondents who agreed to share information consented to a privacy policy that specified, in part, "This application will not access the profile information of any friends, groups, or other information associated with your profile page."

[9]Unfortunately, the Facebook data our app gave us access to does not contain information on the number of friends users have, so we could not create an analogous measure.

[10]Our question suggested respondents should include *candidates* for office in their totals, but since the question was asked in spring 2016, we do not think this should have much influence on the reported figures.

[11]Accessed at `https://sunlightlabs.github.io/congress/legislators.html` on 5/3/2017.

Note that these are only accounts associated with national-level figures. We do not currently have access to a similar list of social media accounts used by state or local level officials (although our supplemental list does include several figures who have only ever held state-level office but who are nationally known, like Gary Johnson). As a result, our objective measure may under-report the true number of politicians followed by respondents who are particularly politically interested.

3.3 Coding Social Media Posts

The simplest measures of social media activity are how often respondents post to the respective platforms. However, we are not only interested in how often they post, but also in how often they post about politics. To examine the calibration of self-reported posts "about politics," we trained a machine learning model using a combination of supervised and semi-supervised methods.

The most straightforward approach would be to implement a keyword-generating algorithm like those described in King, Lam and Roberts (2017); Linder (2017); Munger et al. (2016). We could iteratively discover and validate keywords that define when a post is "about politics," but this requires us to treat keywords as the quantity of interest, when in fact the topic of the post is the quantity of interest.[12] We elected to use the keyword approach as the first step in our classification process.

Ultimately, our aim is to see how well people's self-perceptions map to objective measures, so we want to be able to classify a post as being "about politics" in the eyes of a human reader. We modify the procedure outlined in Benoit, Munger and Spirling (2017), who use the intuitions of human coders to infer a concept (textual complexity) that is difficult to define proscriptively. The quantity of interest is how people *perceive* texts as political or non-political, so we aim to measure that directly using human coders, then create a classifier

[12]This approach is perfectly appropriate for defining topics; although there will always conceptual slippage around the edges, the question of "is this text about gun control?" is far more concrete than "is this text about politics?".

that we can apply to the entire corpus of posts by our respondents. We performed this procedure separately for tweets and Facebook posts, and will note differences between the two.

We began with anchor terms that are nearly guaranteed to be found only in posts about politics: "Obama, Donald, Barack, Trump, election, Clinton, Hillary, Republican, Democrat." We coded all posts containing any of these terms as political. For a summary of the number of tweets and Facebook posts coded at each step, see Table 3. From these anchor texts, we found the 100 terms that were most likely to co-occur with our anchor terms in a single post and used these to identify 1,000 tweets and 1,000 Facebook posts to be hand-coded.[13] We then had six undergraduates hand-code these posts, with 3 coders assigned to each post.

In keeping with the above discussion, our instructions to the coders were intentionally (and explicitly) vague.[14] We coded each text according to the majority vote of the coders who evaluated that text. Table 3 displays a summary. In general, inter-coder reliability was somewhat higher for the Facebook posts than for the tweets, but this effect is explained by the fact that the distribution of labels was unbalanced for the former task: the coders labeled only 14% of the Facebook posts as political compared to 65% of the tweets.

Using this labeled data, we trained a Naive Bayes classifier to predict whether texts were political or not. To validate the model, we performed 50-fold cross-validation — 50 separate samples of the data used as training data to evaluate the accuracy of the model on the "held-out" data, a crucial machine learning step to prevent overfitting. The models had out-of-sample accuracies of 77% (tweets) and 89% (Facebook posts). The higher score for the latter model was expected because tweets were capped at 140 characters, giving the model less information to classify them with. However, accuracy is not the only criterion of interest in selecting the appropriate model. We also wanted to balance the false-negative and

[13]For details on coding, including pre-processing and cleaning steps, see Appendix A.

[14]Our instructions to coders: "Put a '1' if the tweet is 'about politics.' We're intentionally leaving this vague because we're actually interested in figuring out what different people mean when they say something is 'about politics,' so use your best judgment."

false-positive rates; failing to do so would introduce bias in the labels of the entire corpus. To select the ideal model (one with the highest accuracy conditional on balanced false-negative and -positive rates), we performed a grid search over two parameters: the sparsity of the terms introduced to the model and the threshold in the posterior estimates.[15]

Finally, we applied our optimized models to the relevant corpora of tweets and Facebook posts. Our estimates are that 27% of the tweets sent by subjects in our sample were political, compared to only 5% of Facebook posts. This is a large difference, larger than the difference between the percentage of tweets (4%) and Facebook posts (2%) containing any of our anchor terms.

However, neither of these "population estimates" of the distribution of political posts can be validated using our intentionally selected sample of human-coded posts as this sample contained only posts with our anchor terms, or terms likely to co-occur with our anchor terms. Thus we can not say how well our classifier would perform on tweets or posts without any of those terms. To validate our results, we took a random sample of 1,000 tweets and 1,000 Facebook posts from the entire population of posts, added 100 "gold standard" posts of each type that contained an anchor term, and paid workers on Amazon's Mechanical Turk to code them as political or not. Each of these posts had already been labeled by our model as political or not, and these codes allowed us to measure the performance of the model as it would be expected to perform on the entire corpus.[16]

Our Facebook model performed extremely well: 97% of the machine labels and human labels were the same. The Twitter model performed slightly less well, at 84% accuracy. This was lower than we found acceptable, particularly because the errors were almost all in the

[15] To vary the sparsity of the terms, we searched over models in which the minimum number of posts in which a given term appeared varied. In general, more complex models (including sparser terms) are lower-bias and higher-variance, but there is no *ex ante* ideal amount of complexity. The posterior threshold is the line at which we take the outputs of the Naive Bayes model (on the 0 to 1 space) and transform them into the binary political/non-political labels we use. If the threshold is too low (high), the rate of false positives to false negatives is too high (low).

[16] We used a coding framework where each post was coded by 2 crowd workers verified by the platform as "Master" coders, held to a higher standard of accuracy. The inter-coder reliability was high: 86% for tweets, 95% for Facebook posts. In the discussion in the paper, we discarded the posts for which there was coder disagreement.

	Tweets	FB Posts	Tweets (Model 2)
Total user texts	1,317,617	1,446,785	
Political texts with "anchor term"	53,537	34,167	
50-fold CV accuracy on human-coded texts	77%	89%	83%
False-negative rate	15%	6%	8%
False-positive rate	9%	8%	9%
% of hand-labeled texts coded as political	65%	14%	
Inter-coder reliability (Fleiss' κ)	0.67	0.75	
% of texts coded as political	27%	5%	20%

Table 3: Data overview of the social media texts in our sample. The first row records how many texts of each type were in our sample, and the second row restricts that number to texts containing one of the political "anchor terms" (Obama, Donald, Barack, Trump, election, Clinton, Hillary, Republican, Democrat). The third through fifth rows report model performance when cross-validated on the hand-labeled datasets. The sixth and seventh rows report the distribution and inter-coder reliability of the hand-labeled datasets (noting that this dataset was not a random sample of the overall "population"). The final row reports the respective "population estimates" of political posts among the texts described in row 1.

same direction: the machine labeled 34% of the posts as political, compared to only 20% of the human-labeled posts.

As a result, we decided to re-train the Twitter model on the expanded dataset of labeled tweets: combining the set labelled by undergraduates that contained anchor terms or terms co-occuring with anchor terms, and the randomly chosen set that had been labelled by MTurk workers. The results of this re-fit are presented in the far right column of Table 3. This model is superior on every dimension — higher accuracy and more evenly balanced false-positive and false-negative rates. When this model is applied to the entire sample, our estimate of the proportion of political tweets falls from 27% of the sample to 20%.[17]

[17] When we adjust our estimates by incorporating information about the classification error structure as suggested by Bachl and Scharkow (2017), the proportion of political tweets falls even lower, to 6%, and the proportion of political Facebook posts shrinks even further to 2%. See Appendix E for details.

4 Results: Social Media Posts

4.1 Twitter

First, we compare respondents' self-reported tweet frequency to the overall number of tweets they posted. Figure 1 displays the raw individual-level data by each response category, with overlaid box plots. We see that there is a correspondence: respondents are more likely to report a greater frequency of tweeting if they actually tweet more ($r = 0.47$).[18] We reiterate that the individual-level tweet posting data is limited by the Twitter API to approximately 3,200 per person, which can be seen in the figure. This and subsequent figures also display the marginal distribution of responses within each category. Here, the marginals indicate that a plurality of respondents (28.8%) said that they tweet "Less often" (where "less often" means less often than every few weeks).

Figure 2 illustrates a similar relationship between self-reported *political* tweet frequency and the number of such tweets posted by respondents with linked accounts ($r = 0.45$). In this case, we identify tweets as "political" according to the procedure outlined in Section 3.3. And while the relationship is about equally strong, we note that the plurality of users (43.5%) say they never tweet about politics.

For tweeting, then, there at least appears to be a rough correspondence between overall post volume and perceived frequency of posting. But are these perceptions well-calibrated to the wording of the survey response options, which specify frequencies such as "Every few weeks" and "About once a day"? To answer this question, we construct measures of tweet frequency by day. This is simply the number of tweets (all or political) posted in the month prior to when we queried the Twitter API for users' posts divided by the number of days (31).[19] When we plot this daily tweet measure against the same self-reported survey

[18]Pearson's correlations — here and in subsequent results — are computed by converting self-reported frequency to a numeric variable and treating it as if on an interval or ratio scale. For per-day measures, we translate response categories as follows: "Never" to 0, "Less often" to 1/50, "Every few weeks" to 1/21, "1 to 2 days a week" to 2/7, "3 to 6 days a week" to 6/7, "About once a day" to 1, and "Several times a day" to 3.

[19]We also created versions of this measure for the period of Sept. 9–Oct. 9, 2016, which coincided with

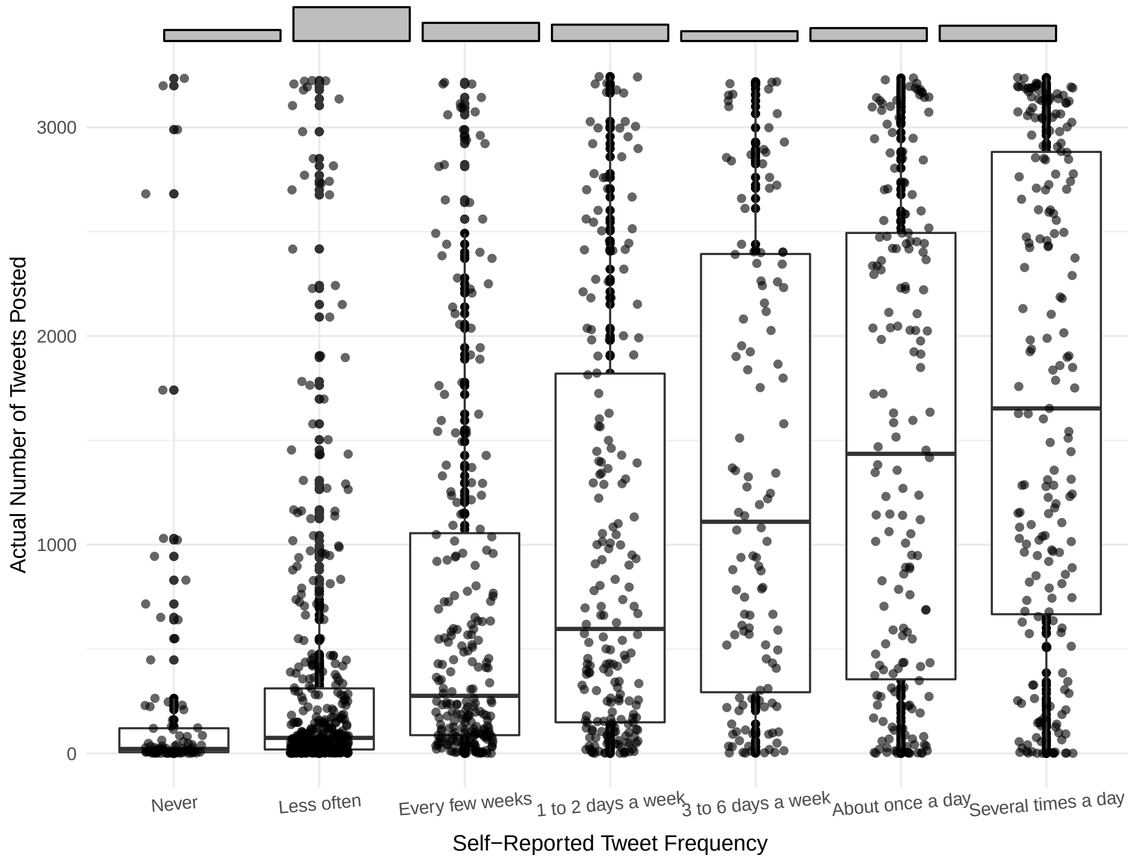

Figure 1: Total number of tweets posted (using linked data from respondents' Twitter accounts) plotted against self-reported tweet frequency.

responses (Figure 3), the correspondence is difficult to detect visually but still present, since we are effectively normalizing tweet volume (all: $r = 0.36$; political: $r = 0.43$). One notable feature of both the total and political versions of the graph is that there are a nontrivial number of tail-end outliers, especially in the "Several times a day" category. This likely reflects right censoring in our data: people who tweet 50 to 100 times a day, although a minority, are placed in the same category as people who tweet 2 to 3 times per day.

Since the average difference across survey response categories can be obscured by the outliers in the figures, we report the mean number of tweets per day in Table 4 below. Here, for both total tweets and political tweets, we generally see a positive relationship between

the second wave of survey data collection and was farther away from Election Day. The correlation between the two measures is 0.76 for tweets per day and 0.71 for political tweets per day, giving us confidence in their reliability.

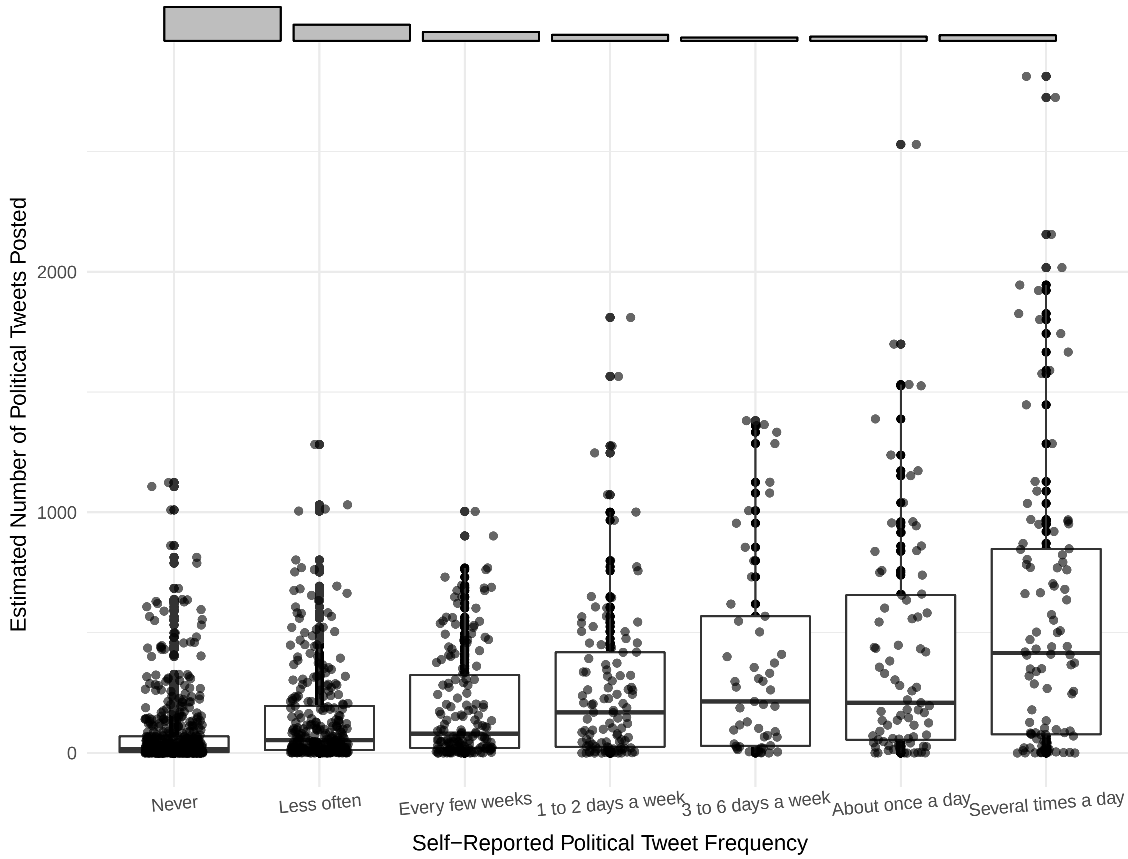

Figure 2: Total number of political tweets posted (using linked data from respondents' Twitter accounts) plotted against self-reported political tweet frequency. Tweets are categorized as "political" via the supervised learning technique discussed in Section 3.3.

a given tweet frequency in each survey category and the mean number of tweets posted by respondents. Strikingly, the calibration appears to be almost perfect in the case of political tweets: an average of 1.23 political tweets were posted per day by respondents who said they tweeted "About once a day." The survey categories representing frequencies less than one tweet a day are also associated with an average number of tweets posted per day of less than one. And there were an average of 6.24 tweets per day posted by those who said they tweet "Several times a day." It is important to keep in mind, however, that these calculations reflect specific decisions about how to define the total number of days counted in the denominator. Using only weekdays, for instance, would somewhat inflate these numbers. Of course these means are obscuring substantial variation in individual behavior. Note that

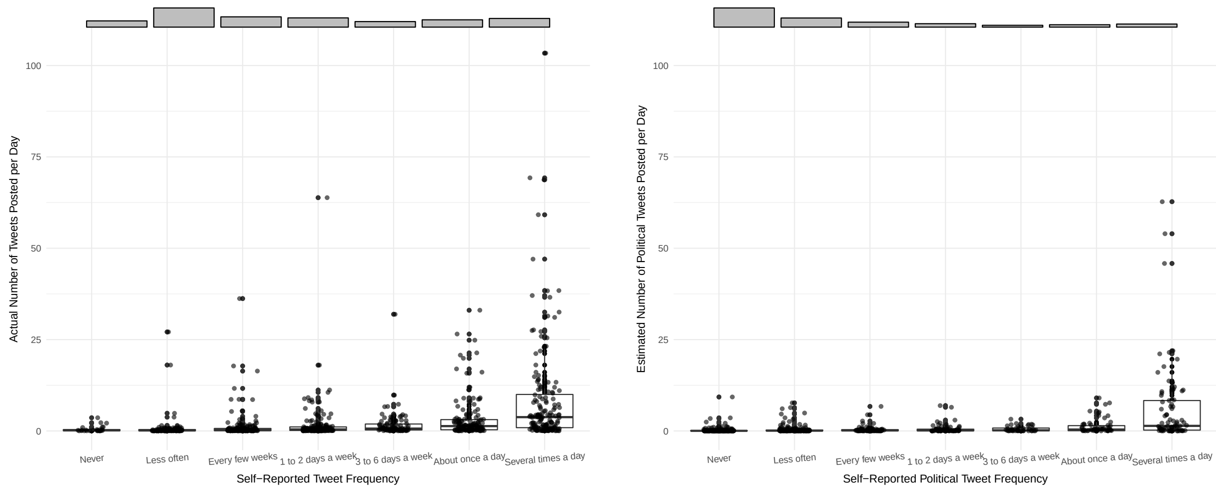

Figure 3: Number of total (left) or political (right) tweets per day posted (using linked data from respondents' Twitter accounts) plotted against self-reported total or political tweet frequency. Tweets are categorized as "political" via the supervised learning technique discussed in Section 3.3.

Survey response	**Mean Number of (all) Tweets Per Day**	**Mean Number of Political Tweets Per Day**
Never	0.53	0.16
Less often	0.63	0.40
Every few weeks	1.17	0.40
1 to 2 days a week	1.76	0.51
3 to 6 days a week	1.71	0.55
About once a day	3.24	1.23
Several times a day	8.60	6.24

Table 4: Mean number of tweets per day for each survey response category. Tweets are categorized as "political" via the supervised learning technique discussed in Section 3.3.

when we examine the column for the mean number of all tweets per day we see that the mean number of tweets per day even for people who claim they tweet only "1 to 2 days a week" is 1.76; thus the mean is higher than it should be if all respondents were accurately reporting. However, much of this is driven by outliers who report tweeting only 1 to 2 days per week, but are in fact tweeting as often as 10 times per day.[20]

[20]Respondents may have interpreted the question to mean 'how many times a day' they post things, rather than 'how many items they post' - and it may be that people post 'about once a day', but each time they post 5 items.

4.2 Facebook

Turning to Facebook, we see a similar pattern: a correspondence between self-reported post frequency and the mean number of total posts as determined by linked profile data. For all types of posts, the correlation is somewhat higher ($r = 0.38$) than for political posts ($r = 0.32$). In terms of general posts, we see in Figure 4 that a plurality of respondents who shared Facebook data (26.3%) said, perhaps unsurprisingly, that they do so "Several times a day." This drops to less than 10% when we look at posts related to politics in Figure 5 (where the most commonly chosen category was "Never," at 33.2%).

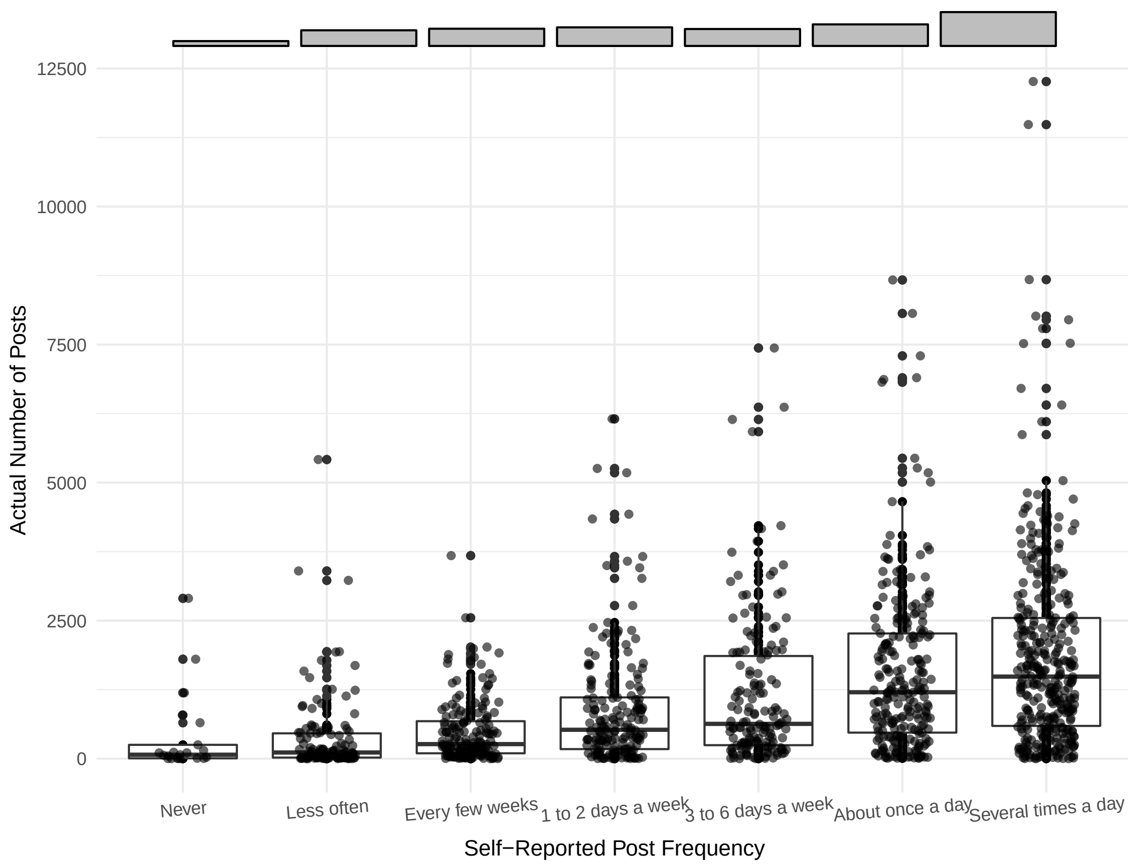

Figure 4: Total number of posts (using linked data from respondents' Facebook profiles) plotted against self-reported post frequency.

Per-day measures of Facebook posting frequency are also correlated with the corresponding survey measures (all: $r = 0.26$; political: $r = 0.20$), but they are heavily mismatched; the mean number of political posts per survey category is off by nearly two orders of magnitude.

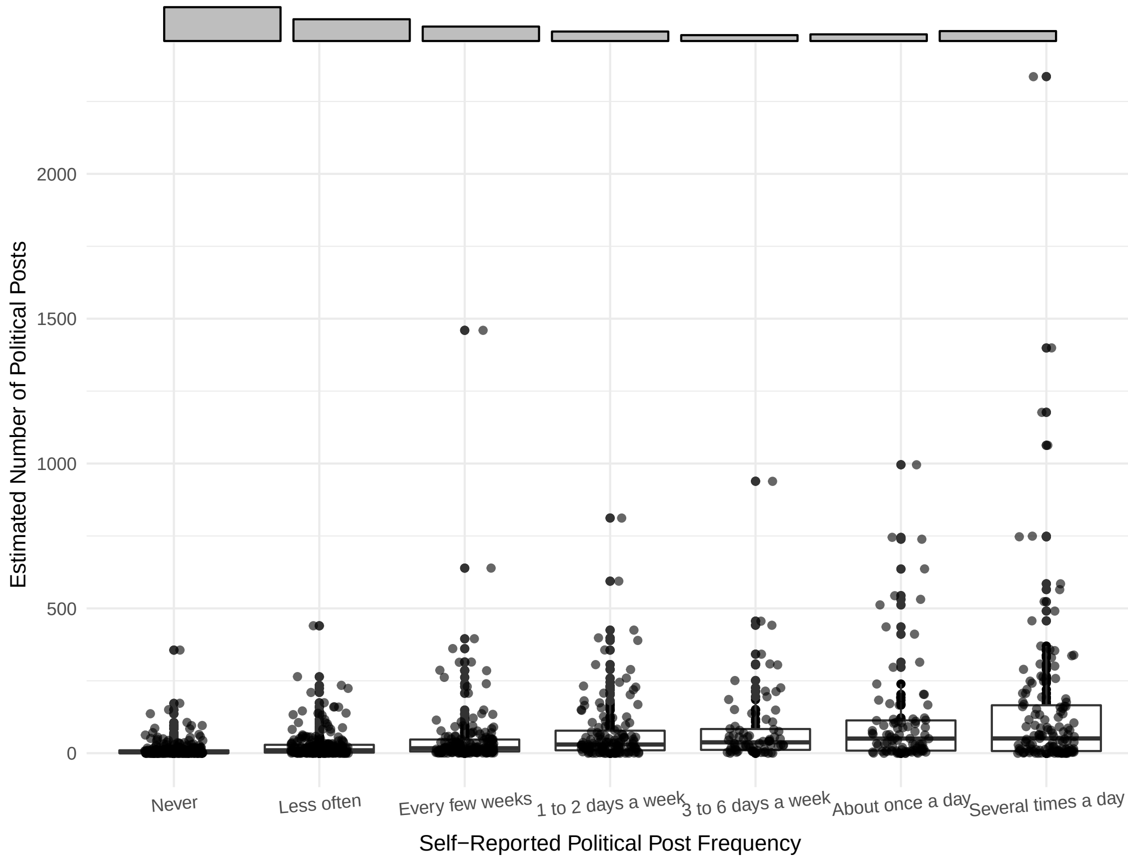

Figure 5: Total number of political posts (using linked data from respondents' Facebook profiles) plotted against self-reported political post frequency. Posts are categorized as "political" via the supervised learning technique discussed in Section 3.3.

This may be due to the fact that, lacking API limitations, we used the entire history of users' Facebook posts to construct the measures: We divided the total number of posts by the total number of days active on Facebook.

4.3 Characterizing Individual-level Discrepancies

While on average there is a correspondence between survey responses and actual social media posting behavior, it is critical to further explore discrepancies on the individual level. Do people tend to over- or underestimate their rates of crafting tweets and Facebook posts? Do certain characteristics systematically predict these discrepancies?

To investigate these questions, we calculate individual-level reporting discrepancies in

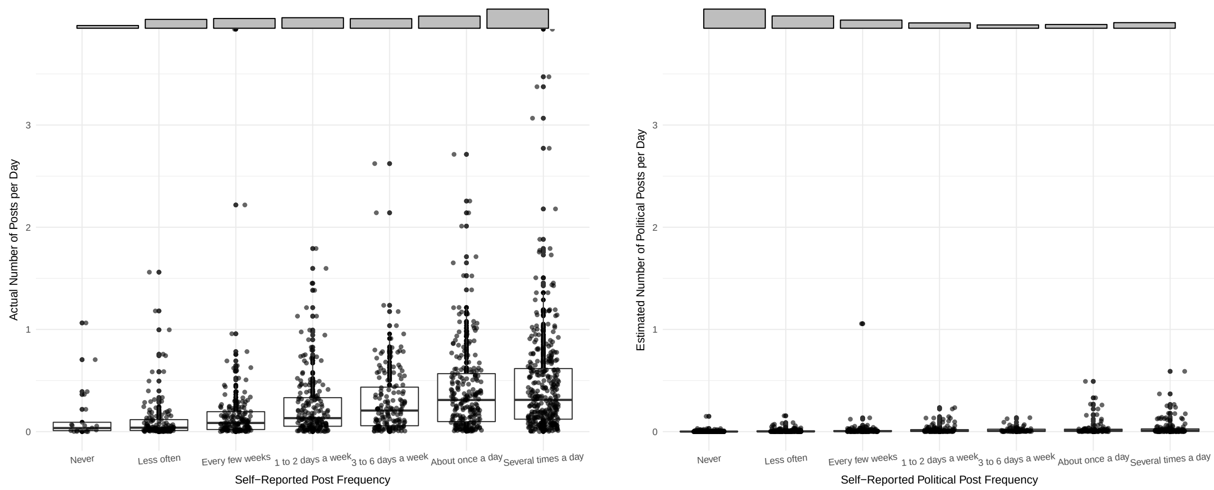

Figure 6: Number of total (left) or political (right) posts per day (using linked data from respondents' Facebook profiles) plotted against self-reported total or political post frequency. Posts are categorized as "political" via the supervised learning technique discussed in Section 3.3.

the following way. First, we convert the survey categories to numerical magnitudes in two different ways.[21] Then we subtract the daily tweet measures from each of those numerically translated survey measures and keep the quantity with the smallest absolute value (given the inherent imprecision in survey categories). Note that we are explicitly measuring *over-reporting*. We then run several linear regression models including a number of hypothesized predictors from the literature on self-reported media consumption, and a set of demographic characteristics. Most importantly, we include a measure of interest — whether or not a respondent voted in the 2016 primaries.[22] We also include both party identification and a measure of party strength: whether or not someone identifies as a "Strong Democrat" or "Strong Republican." The results are shown in Table 5.

There is a conceptual difference between the discrepancies in the rates of *overall* posting and posting *about politics*. The former is objective, while the latter is subjective; for some people, the definition of what makes a post "about politics" is more expansive than for others.

[21]In the first version, we equate "Never" to 0, "Less often" to 1/50, "Every few weeks" to 1/21, "1 to 2 days a week" to 2/7, "3 to 6 days a week" to 6/7, "About once a day" to 1, and "Several times a day" to 3. In the second version, "Less often" is 1/40, "1 to 2 days a week" is 1/7, and "3 to 6 days a week" is 3/7.

[22]The survey did not have a more direct measure of political interest.

One could interpret the results in columns 1 and 3 as "errors" while those in columns 2 and 4 could be some mix of errors and of measures of what describes the types of people with these more expansive conceptions of the political.

Patterns in reporting social media activity are quite different on Facebook and Twitter. In the first two columns, it is evident that none of our predictors are related to discrepancies in survey self-reports of overall tweet frequency or of tweeting about politics. The constant in both models is also insignificant, suggesting no general bias toward overreporting or underreporting among our respondents. Indeed, the adjusted R^2 of these models is below 0, indicating that they are quite poor at explaining the variation. A major explanation for this problem is that there is a massive left tail for the dependent variable: our transformed measure has a maximum of 3, a median near 0, and a minimum of approximately -100. The models in the first two columns have had their dependent variables log-transformed (symmetric around zero) to account for this fact.

This is due to the true variability in the way that people use Twitter. With some people posting hundreds of times a day and others posting twice a month, 6-point questions like the ones we employ are simply insufficient.

On Facebook, by contrast, a more familiar pattern emerges: Those who voted in the primaries, as well as people who are more partisan, are significantly more likely to overreport their posting frequency. People also seem to systematically underestimate the number of political posts they make to Facebook, as illustrated by the negative constant in the fourth column. Finally, age appears to be positively associated with errors in reporting Facebook posting frequency.

Appendix C expands this analysis, using survey responses about the use of Twitter to predict discrepancies in the reported/objective use of Facebook (and vice versa). This analysis is only possible on the subset of subjects who use *both* networks, but the results are informative: these questions are extremely powerful predictors. Also unexpected is the asymmetric effect of self-reported *viewing* the two services. People who report viewing

Table 5: Predictors of Individual-level Social Media Over-Reporting

	Tweets/day (All)	Tweets/day (Pol)	FB/day (All)	FB/day (Pol)
	(1)	(2)	(3)	(4)
Voted in primary	−0.069	0.036	0.223**	0.196***
	(0.106)	(0.103)	(0.090)	(0.075)
Party: Democrat	−0.066	0.043	0.166	0.229**
	(0.132)	(0.128)	(0.108)	(0.089)
Party: Republican	−0.108	0.009	0.123	0.116
	(0.260)	(0.252)	(0.229)	(0.192)
Party: Independent	−0.153	0.120	0.172	0.249
	(0.212)	(0.214)	(0.210)	(0.174)
Party: Not sure	−0.095	−0.085	−0.046	0.032
	(0.115)	(0.113)	(0.098)	(0.081)
Strong partisan	0.066	0.098	0.221**	0.294***
	(0.114)	(0.111)	(0.093)	(0.077)
Age	−0.003	−0.003	0.009***	0.009***
	(0.003)	(0.003)	(0.002)	(0.002)
High school graduate	0.181	0.147	0.025	0.084
	(0.148)	(0.146)	(0.126)	(0.104)
Some college	−0.035	−0.119	0.138	−0.028
	(0.153)	(0.151)	(0.126)	(0.105)
2-year college	−0.216	−0.087	0.382	0.194
	(0.368)	(0.375)	(0.247)	(0.202)
College graduate	0.054	0.067	−0.207	−0.115
	(0.169)	(0.165)	(0.137)	(0.114)
Postgrad	0.066	0.021	0.092	0.036
	(0.146)	(0.144)	(0.123)	(0.102)
Female	−0.019	−0.040	0.126*	−0.153***
	(0.084)	(0.083)	(0.072)	(0.059)
Nonwhite	0.117	0.084	−0.049	0.080
	(0.102)	(0.100)	(0.084)	(0.070)
Constant	−0.319	−0.606***	−0.050	−0.566***
	(0.225)	(0.221)	(0.193)	(0.160)
N	645	627	1,084	1,063
Adjusted R^2	-0.007	-0.006	0.028	0.048

*p < .1; **p < .05; ***p < .01
OLS models. Reference category for party is "Other."
Columns 1 and 2 take as their DV the log-transformed error,
to deal with the wide range of the data.

Survey response	Twitter	Facebook
0	0.72 (47.4%)	1.3 (48.5%)
1-10	3.4 (40.4%)	3.8 (41.6%)
More than 10	11.4 (12.2%)	7.0 (9.9%)

Table 6: Mean number of "elected officials, candidates for office, or other political figures" followed on Twitter or liked on Facebook within each corresponding survey response category. The numbers in parentheses report the marginal distribution of survey responses in each category.

Twitter more often have lower levels of overreporting *all four* behaviors. The effect of viewing Facebook is the opposite: those who report viewing Facebook more often have *higher* levels of overreporting all four behaviors. We discuss possible explanations in Appendix C.

5 Results: Friends and Political Sources

We extend our focus beyond tweets and posts to people's social networks. Here we are interested in a variant on the traditional question of how large people's networks are, asking specifically how many "elected officials, candidates for office, or other political figures" they follow on Twitter or like on Facebook. In Table 6, we show the mean number of political accounts followed on Twitter and mean number of politicians' official accounts liked on Facebook against the corresponding survey responses. Similar to the previous results, we find a correspondence, and note that most respondents appear to follow either no or relatively few political figures on social media. The calibration for Twitter political follow networks seems to be especially accurate on average.

6 Discussion

We take advantage of a unique linked dataset which allows us to validate self-reports of political activity. The good news is that self-reports are correlated with actual behavior. People who follow political actors on Twitter are more likely to say they do than are respondents who do not follow political actors on Twitter. However, at the individual level, there are substantial discrepancies in reporting, discrepancies that covary with demographic variables of interest (on Facebook).

Our data also allows us to make comparisons about how people behave across different social media platforms. The most striking difference is that a much higher percentage of tweets in our dataset (20%) were coded as political than were Facebook posts (5%). This result is somewhat surprising; although Twitter users are less representative of the population than Facebook users (in particular they tend to be more highly educated), users of each platform report roughly equivalent rates of seeing (and posting) political content (Duggan and Smith 2016). Pew's most recent survey reports that 8% of Twitter users and 6% of Facebook users claim the highest possible category of political posting — that "a lot" of what they post is related to politics. This is at odds with our finding of very different proportions of content being about politics.[23]

We have demonstrated five important points in this paper. First, self-reports of social media use are meaningful; they are (perhaps surprisingly) accurate and correlated with our objective measure at the aggregate level. Second, there are substantial discrepancies between objective and self-reported posting behavior at the individual level. The nature of these discrepancies is difficult to summarize, and it is not clear whether the predictors we observe, such as age and political interest for Facebook posting, will always be correlated with overreporting.

Third, questions about the number of political actors a given respondent follows or likes appear to be particularly well-calibrated; this is a particularly explicit question, suggesting that people find it easier to answer. Fourth, contrary to what self-reports suggest and given the caveats above, a much higher proportion of posts on Twitter are about politics than are posts on Facebook.

[23] As mentioned above, adjusting these proportions as suggested by Bachl and Scharkow (2017) gives estimates of 6% political tweets and 2% political Facebook posts. Although the gap in percentage points terms is lower, the percentage gap is similar to the unadjusted gap. See Appendix E for details.

Finally, we have shown that researchers interested in measuring online behavior should consider modifications to standard survey question wordings. In the first place, they should be more liberal in their top-coding of survey responses and allow for the behavior of heavy users. We recommend that an adjustment be made to all survey questions about social media use to reflect the nearly exponential distribution of behavior we observe with our objective measures. There is simply more information on the tails of the distribution than is being recorded by current survey questions.

Additionally, given persistent questions about differing subjective perceptions of what constitutes "political" content, it would be useful to experiment with anchors that explicitly provide respondents with guidance on the kinds of political topics researchers wish to focus on. This could be as simple as "posts about policy proposals and government affairs" or "debates about collective responses to societal problems," and could additionally encompass examples illustrating the scope of such discussions.

The opportunity to compare self-reported behavior with observed behavior is rare (Guess 2015; Scharkow 2016; Revilla, Ochoa and Loewe 2017). Such cases provide us with a reference standard that is extremely valuable in validating the use of self-reports, as well as in providing appropriate cautions.

References

Anderson, Barbara A. and Brian D. Silver. 1986. "Measurement and Mismeasurement of the Validity of the Self-reported Vote." *American Journal of Political Science* 30:771–785.

Bachl, Marko and Michael Scharkow. 2017. "Correcting Measurement Error in Content Analysis." *Communication Methods and Measures* 11(2):87–104.

Bartels, Larry M. 1993. "Messages Received: The Political Impact of Media Exposure." *American Political Science Review* 87(2):267–285.

Benoit, Kenneth, Kevin Munger and Arthur Spirling. 2017. Measuring and Explaining Political Sophistication Through Textual Complexity. Technical report Working Paper.

Bohdanova, Tetyana. 2014. "Unexpected revolution: the role of social media in Ukraine's Euromaidan uprising." *European View* pp. 1–10.

de Vreese, Claes H and Peter Neijens. 2016. "Measuring media exposure in a changing communications environment.".

Duggan, Maeve. 2015. *The Demographics of Social Media Users.* Pew.

Duggan, Maeve and Aaron Smith. 2016. "The political environment on social media." *Pew Research Center* 25.

Gottfried, Jeffrey and Elisa Shearer. 2016. "News use across social media platforms 2016." *Pew Research Center* .

Guess, Andrew M. 2015. "Measure for Measure: An Experimental Test of Online Political Media Exposure." *Political Analysis* 23(1):59–75.

Iyengar, Shanto and Donald R Kinder. 1987. "News that matters: Agenda-setting and priming in a television age." *News that Matters: Agenda-Setting and Priming in a Television Age* .

Jerit, Jennifer, Jason Barabas, William Pollock, Susan Banducci, Daniel Stevens and Martijn Schoonvelde. 2016. "Manipulated vs. Measured: Using an Experimental Benchmark to Investigate the Performance of Self-Reported Media Exposure." *Communication Methods and Measures* 10(2-3):99–114.
URL: *https://doi.org/10.1080/19312458.2016.1150444*

Karpf, David. 2012. "Social science research methods in Internet time." *Information, Communication & Society* 15(5):639–661.

King, Gary, Patrick Lam and Margaret E Roberts. 2017. "Computer-Assisted Keyword and Document Set Discovery from Unstructured Text." *American Journal of Political Science* .

Krumpal, Ivar. 2013. "Determinants of social desirability bias in sensitive surveys: a literature review." *Quality & Quantity* 47(4):2025–2047.

Lazarsfeld, Paul F., Bernard Berelson and Hazel Gaudet. 1944. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign.* New York: Columbia U. Press.

Linder, Fridolin. 2017. "Improved Data Collection from Online Sources Using Query Expansion and Active Learning.".

McKibben, William Bradley and Paul J Silvia. 2017. "Evaluating the distorting effects of inattentive responding and social desirability on self-report scales in creativity and the arts." *The Journal of Creative Behavior* 51(1):57–69.

Munger, Kevin, Patrick Egan, Jonathan Nagler, Jonathan Ronen and Joshua A Tucker. 2016. Learning (and Unlearning) from the Media and Political Parties: Evidence from the 2015 UK Election. Technical report.

Mutz, D.C. and L. Young. 2011. "Communication and Public Opinion." *Public Opinion Quarterly* 75(5):1018–1044.

Mutz, Diana C and Paul S Martin. 2001. "Facilitating Communication Across Lines of Political Difference: The Role of Mass Media." *American Political Science Review* 95(1):97.

Niederdeppe, Jeff. 2014. "Conceptual, empirical, and practical issues in developing valid measures of public communication campaign exposure." *Communication Methods and Measures* 8(2):138–161.

Prior, Markus. 2009. "Improving media effects research through better measurement of news exposure." *The Journal of Politics* 71(03):893–908.

Prior, Markus. 2012. "Who watches presidential debates? Measurement problems in campaign effects research." *Public Opinion Quarterly* 76(2):350–363.

Prior, Markus. 2013. "Media and political polarization." *Annual Review of Political Science* 16:101–127.

Revilla, Melanie, Carlos Ochoa and Germn Loewe. 2017. "Using Passive Data From a Meter to Complement Survey Data in Order to Study Online Behavior." *Social Science Computer Review* 35(4):521–536.
URL: *https://doi.org/10.1177/0894439316638457*

Romantan, Anca, Robert Hornik, Vincent Price, Joseph Cappella and K Viswanath. 2008. "A comparative analysis of the performance of alternative measures of exposure." *Communication Methods and Measures* 2(1-2):80–99.

Scharkow, Michael. 2016. "The Accuracy of Self-Reported Internet Use—A Validation Study Using Client Log Data." *Communication Methods and Measures* 10(1):13–27.

Schwarz, Norbert. 1999. "Self-reports: How the questions shape the answers." *American Psychologist* 54(2):pp. 93–105.

Settle, Jaime. 2018. *Frenemies: How Social Media Polarizes America.* Cambridge University Press.

Stone, A.A., C.A. Bachrach, J.B. Jobe, H.S. Kurtzman and V.S. Cain. 1999. *The Science of Self-report: Implications for Research and Practice.* Taylor & Francis.
URL: *https://books.google.com/books?id=k9p4AgAAQBAJ*

Taneja, Harsh. 2016. "Using commercial audience measurement data in academic research." *Communication Methods and Measures* 10(2-3):176–178.

Wells, Chris and Kjerstin Thorson. 2017. "Combining big data and survey techniques to model effects of political content flows in Facebook." *Social Science Computer Review* 35(1):33–52.

Appendix A: Process for Coding Posts

For both the Facebook and Twitter models of political posts, we performed the following steps:

- Took labels from undergrads and assigned the modal (of 3) label to each text
- Created a corpus of texts from these 1,000 labeled posts
- Removed stopwords (both upper and lower case, from the R package `quanteda`), supplemented by terms that are artifacts of the social media text ("https", "t", "RT")
- Removed punctuation, URLs, numbers, and symbols (excluding Twitter-relevant symbols "#" and "@")
- Removed sparse terms (at this stage, terms needed to appear in 5 separate posts)
- Performed 10-fold cross-validation to optimize our Naive Bayes model, selecting the posterior cutoff point for assigned binary labels
- Read in "anchor posts" that contained an anchor term (Obama, Donald, Barack, Trump, election, Clinton, Hillary, Republican, Democrat), labeled these posts as political
- Created a corpus of texts from the 1,000 labeled posts, the labeled "anchor posts", and the rest of the unlabeled posts
- Performed the same pre-processing as above
- Removed sparse terms (at this stage, terms needed to appear in 1,000 separate posts)
- Used trained Naive Bayes model to predict labels for the unlabeled posts

Appendix B: Demographics of Respondents

The respondents who shared their social media information may not be a representative sample of all social media users, limiting the external validity of our findings. Indeed, we find that we were able to scrape tweets from only 1,421 of the 2,338 self-reported Twitter users in the sample.

There were three stages of attrition: people who claimed to use Twitter but did not give us a username (203 people); people who gave us a username that we were unable to match with an existing account (347 people); and people who gave us an existing account for which we were unable to retrieve any tweets (395 people).[24]

The first form of attrition is most relevant to the concern that people who choose to share their account information might be different from other social media users, and indeed, this is the only group that differs from the final sample in noteworthy ways. Panel A of Table B1 displays this information — this sample of 203 respondents is notably younger, much more likely to report never tweeting, and less likely to report tweeting daily. It seems like this sample represents (from the set of Twitter users) more casual Twitter users, many of whom may just have an account in order to read others' tweets. More importantly, this sample represents a small (8.6%) portion of the overall sample of Twitter users, suggesting that the overwhelming norm among survey respondents is in fact to share their Twitter account.

The second kind of attrition is most likely to be random, the result of a typo or mistaken memory, and indeed the subjects who provided invalid user IDs are similar to the final pool.

The third kind of attrition could be caused by several factors. If users provided us with a valid ID that they had set to "protected," we could not scrape their accounts. This was the case for 149 of the 395 subjects. Another possibility is these respondents simply never tweeted during the collection period — an additional 98 reported that they tweeted "never" or "less often" than every few weeks. Overall, these respondents were also very similar to the ultimate sample.

Panel B displays the analogous information for Facebook users. Here, there is only one source of attrition: whether respondents gave us access to their Facebook accounts. This is a somewhat more involved concession; this access allows us to see everything they posted and liked, information that would not otherwise be public. In contrast, for all Twitter users without private accounts (the vast majority), their tweets and likes are already public, so the only information they gave us was to identify themselves.

[24] Although this was not intended, the branching survey we designed allowed respondents who reported not using Twitter to still be prompted to provide their Twitter username. 28 respondents did so; for 12 of these, we were able to scrape their tweets. These respondents are excluded from all of the results in this paper.

As a result, only 1,221 of the 2,711 self-reported Facebook users in our sample authorized us to access their accounts. We see in Panel B that these populations are indeed somewhat different: the respondents who shared their information were more likely to be Democrats, were more educated, and were heavier Facebook users.

We thus have to acknowledge that there is a tradeoff between the generalizability of our efforts to measure the latent trait of "accuracy" of self-reported social media use and the need to get informed consent from all of the respondents. Unlike the case of, say, pre-election polling, we have very little information about the specific demographic characteristics that would be necessary to counter this selection bias problem through weighting. We thus welcome further efforts to address this thorny issue of weighing selection and informed consent.

Table B1: Descriptive Statistics of Relevant Populations

Panel A: Twitter Users

	Report Twitter Use	Provided Username	Valid Username	Have Tweets
% Democrat	41	46	45	44
% Republican	22	20	23	21
Mean Ideology (5-point)	2.9	3.1	3.0	3.0
Mean Age	43	51	51	49
% HS or less	20	27	24	25
% Postgraduate	12	14	17	14
Self-Reported				
% Tweet Daily	18	22	21	26
% Never Tweet	22	10	12	7
% Tweet Politics Daily	10	14	10	13
% Never Tweet Politics	57	45	43	42
N	203	347	395	1,421

Panel B: Facebook Users

	Report Facebook	Provided Facebook Access
% Democrat	41	45
% Republican	26	20
Mean Ideology (5-point)	2.9	2.8
Mean Age	50	49
% HS or less	31	25
% Postgraduate	12	15
Self-Reported		
% Post Daily	40	47
% Never Post	5	2
% Post Politics Daily	14	19
% Never Post Politics	37	29
N	1,490	1,221

Note: Each column represents a disjoint set of respondents. Each column displays the information for users who fit that criterion but not the one to the right. For example, the 347 users in column 3 of Panel A provided a Twitter username but did not provide a valid Twitter username.

Appendix C: Extended Models of Individual Discrepancies in Reported Social Media Use

In addition to the demographic information we use to predict individual-level discrepancies between reported and observed social media posting presented in Table 5, we have access to several questions about related aspects of social media use. The regressions presented below demonstrate their predictive power.

Each of these regressions contains all of the predictors included in the models in Table 5; we only display the results for the variables that achieve significance in at least one regression. The most robustly significant variables in the initial regressions are still significant here; age seems especially related to discrepancies in Facebook activity, with older people much more likely to overreport their rate of posting.

The next several variables are responses to our questions about frequency of posting and posting about politics on the *alternative* platform. That is, in columns 1 and 2 (3 and 4), these are the questions about posting on Facebook (Twitter). Here we find strikingly asymmetric results.

Reporting posting on Facebook more often is a significant predictor of less overreporting of sending tweets (column 1). The results are essentially the same in column 2, for sending political tweets; however, reporting posting about politics on Facebook more often is associated with *more* overreporting of sending political tweets.

The converse effect is significant in the *opposite* direction: reporting posting on Twitter more often is a significant predictor of *more* overreporting of posting on Facebook (column 3). The same trend can be observed in self-reported rates of tweeting about politics.

This asymmetry can also be observed in the final two rows: the frequency with which respondents report *viewing* Facebook and Twitter. People who report viewing Twitter more often have lower levels of overreporting *all four* behaviors. The effect of viewing Facebook is the opposite: those who report viewing Facebook more often have *higher* levels of overreporting all four behaviors.

The fact of a relationship between these questions and the dependent variable is unsurprising; we are simply capturing something about how a given respondent answers questions about their media use. Some people likely systematically over-estimate, others underestimate, due to a personal idiosyncrasy.

The platform asymmetry is surprising, however, and we do not have a conclusive explanation for it. Our conjecture is based on the relative algorithms of the two sites. Twitter power users take up a larger percentage of a given Twitter user's feed because of the relatively light algorithmic filtering used by Twitter. Simply tweeting frequently is in fact a necessary part of becoming a Twitter power user. The Facebook News Feed is far more active, and might show a smaller percentage of posts by a given Facebook power user. The structural incentives for becoming a Facebook power user are more focused on making popular, shareable posts that are more highly ranked by the News Feed algorithm.

Table C1: Additional Effect of Using Reported Use of Alternative Social Media Site

	Tweets	Tweets (Pol)	FB Posts	FB Posts (Pol)
	(1)	(2)	(3)	(4)
Strong Partisan	0.112	0.108	0.178^{**}	0.236^{***}
	(0.117)	(0.111)	(0.090)	(0.073)
Age	-0.005^{*}	-0.004	0.014^{***}	0.011^{***}
	(0.003)	(0.003)	(0.002)	(0.002)
Gender	-0.094	-0.075	0.154^{**}	-0.034
	(0.090)	(0.085)	(0.071)	(0.057)
Alt SM Post Frequency	-0.077^{**}	-0.152^{***}	0.168^{***}	-0.065^{***}
	(0.034)	(0.032)	(0.026)	(0.021)
Alt SM Post Pol Frequency	0.017	0.091^{***}	0.123^{***}	0.319^{***}
	(0.025)	(0.024)	(0.024)	(0.019)
Frequency View Twitter	-0.083^{***}	-0.126^{***}	-0.099^{***}	-0.059^{***}
	(0.025)	(0.024)	(0.025)	(0.020)
Frequency View Facebook	0.090^{*}	0.107^{**}	0.389^{***}	0.150^{***}
	(0.047)	(0.044)	(0.033)	(0.027)
Constant	-0.080	-0.134	-2.511^{***}	-1.451^{***}
	(0.324)	(0.305)	(0.270)	(0.219)
N	595	592	881	877
Adjusted R^2	0.024	0.083	0.275	0.310

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$
OLS models. Each model also controls for all of the variables modeled in Table 3, none of which are significant in the present models. Columns 1 and 2 take as their DV the log-transformed error, to deal with the wide range of the data.

Appendix D: Survey Question Wordings and Response Categories

- "How often do you view [Facebook/Twitter]?"
- "How often do you post things on [Facebook/Twitter]?"
- "How often do you post things **about politics** on [Facebook/Twitter]?"
  - Several times a day
  - About once a day
  - 3 to 6 days a week
  - 1 to 2 days a week
  - Every few weeks
  - Less often
  - Never
  - Don't Know
- "How many people do you follow on Twitter?"
  - 1-100
  - 101-250
  - 251-500
  - More than 500
  - Don't know
- "How many elected officials, candidates for office, or other political figures do you follow or like on [Twitter/Facebook]?
  - 0
  - 1-10
  - More than 10
  - Don't know

Appendix E: Error-corrected Category Proportions

The proportions of our Twitter and Facebook samples that we report as being about politics — like all content analyses — are subject to misclassification error. Standard practice is simply to quantify this error: report intercoder reliability statistics for human coding procedures in addition to measures of accuracy and error for supervised learning approaches. But, once classification is complete, the error is often set aside for the subsequent analysis. Here, we apply the matrix back-calculation technique recommended by Bachl and Scharkow (2017) to leverage what we know about the error structure of the Naive Bayes automated coding procedure to produce error-corrected estimates of the proportion of tweets and Facebook posts that are about politics. This does not solve the issue of lacking a universally agreed-upon ground truth regarding the political content of our social media data, but it potentially addresses errors in extending the process that generated the hand-coded labeled set to the entire corpus.

Bachl and Scharkow (2017) show that if a misclassification matrix can be constructed, either by referencing external standards or by assumption, then multiplying the inverse of that matrix by the vector of observed proportions will back out the "true" error-corrected proportions:

$$\begin{bmatrix} \theta_{\text{pol}|\text{pol}} & \theta_{\text{pol}|\neg\text{pol}} \\ \theta_{\neg\text{pol}|\text{pol}} & \theta_{\neg\text{pol}|\neg\text{pol}} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \theta^{obs}_{\text{pol}} \\ \theta^{obs}_{\neg\text{pol}} \end{bmatrix} = \begin{bmatrix} \theta_{\text{pol}} \\ \theta_{\neg\text{pol}} \end{bmatrix}. \tag{1}$$

As they note, the Bachl and Scharkow method can be adapted to our automated coding setting by substituting the confusion matrix for the misclassification matrix. We do so below for both the Twitter and Facebook classifiers. This procedure implicitly (and incorrectly) assumes that hand-coded labels are the "truth," with sources of error arising solely from the automated classification step.

First, the re-estimated Twitter model has the following confusion matrix (normalized so that columns sum to 1):

$$C_T = \begin{bmatrix} 0.79 & 0.16 \\ 0.21 & 0.84 \end{bmatrix}$$

Since the observed proportion of political tweets using that model is 0.20, we calculate $C_T^{-1} \cdot \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix} = \begin{bmatrix} 0.06 \\ 0.94 \end{bmatrix}$. In other words, the error-corrected proportion is computed to be 6%, or 14 percentage points lower than the naive estimate.

For the Facebook model, the calculation is:

$$\begin{bmatrix} 0.40 & 0.04 \\ 0.60 & 0.96 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0.05 \\ 0.95 \end{bmatrix} = \begin{bmatrix} 0.02 \\ 0.98 \end{bmatrix}.$$

The correction again lowers the estimated proportion, this time from 5% to approximately 2%.