How the Supply of Fake News Affected Consumer Behavior during the 2016 US Election

Anil R. Doshi*, Sharat Raghavan†, Rebecca Weiss‡, Eric Petitt§

October 18, 2018

Abstract

We characterize the effect of fake news on online browsing during the 2016 US presidential election. We estimate that weekday increases of 10 fake news articles—that were confirmed to be false by third-party services—increased the incidence of fake news site visits by 3.0%. To address endogeneity, we employ two approaches that attempt to isolate exogenous variation in fake news supply. We also estimate that weekday 10-article increases in fake news increase the odds of visiting one or more fake news sites by 3.7%. Overall, this evidence demonstrates the effectiveness of fake news production in reaching a diverse set of consumers.

JEL Classification: L82, P16

*UCL School of Management, Level 38, One Canada Square, London E14 5AA, UK. anil.doshi@ucl.ac.uk. Corresponding author.

†Haas School of Business, University of California at Berkeley.

‡Mozilla Corporation and Stanford University; rweiss@mozilla.com.

§Mozilla Corporation; epetitt@mozilla.com. We thank the Mozilla Corporation for providing data for this project, DomainTools for helping us obtain domain registration data, and Hunt Allcott and Matthew Gentzkow for generously sharing their data with us. We acknowledge Matt Backus, Ivan Png, and Thomas Wollmann—as well as participants in the seminars where this paper was presented—for helpful discussions at various stages of the project.

1 Introduction

How did the production of unverifiable or verifiably false news content affect online news consumption in the run-up to the 2016 US presidential election? Although misinformation campaigns are hardly novel in news about science (see e.g. Bushman and Anderson, 2001; Lewandowsky et al., 2012) or politics (for examples, see Allcott and Gentzkow, 2017), the low production cost of online content and the pervasiveness of social media may magnify such a campaign's effectiveness. For instance, Pew Research reports that over two thirds of Americans receive news through social media channels; some do so exclusively (Gottfried and Shearer, 2016). Misinformation during the 2016 US election was therefore probably distinct from prior iterations and has elicited concern about how "fake news" influences elections. Recent studies have analyzed the spread of fake news (Del Vicario et al., 2016) and how social media affected exposure to it (Allcott and Gentzkow, 2017; Fourney et al., 2017; Guess et al., 2017), but there is still uncertainty about the extent to which users responded to fake news. That particular mechanism has attracted the attention of US Congressional committees seeking to discover how non-US–linked Facebook advertising influenced voters (Shane and Isaac, 2017).

This paper examines the relationship between the production of fake news and the browsing behavior of Internet users. Despite the considerable public discussion of fake news, it could be that consumer news preferences are inelastic or limited to ideological "echo chambers". In either case, consumption may not have been substantially affected by changes in the production of fake news. However, if the supply of fake news directly affected browsing behavior after accounting for individual fixed characteristics, then we should want to assess the effectiveness of fake news producers in generating views of similar content. An estimate of this metric would be a step toward understanding the economics of fake news and would help us to understand how filters, or measures that restrict dissemination of news on social networks or Internet platforms, ultimately affect consumption (Kumar et al., 2016).

Our estimation strategy starts with a novel data set from the Mozilla Corporation, which develops the Firefox Web browser. As part of a recent research initiative, US Internet users were recruited to download a browser add-on that tracked their browsing history from September 20 to December 16, 2017.[1] The resulting database covers the period around the US presidential election, which saw the proliferation of many fake news articles. During the 49 days prior to the election, 1,418 users had joined the study and accumulated 2,278,825 website visits. Of those visits, only 22,533 (about 1% of the total) were to fake news sites. Even so, 54% of the study's users viewed at least one fake news site during the period. Moreover, daily user average visits to fake news sites was skewed in our sample—with a mean of 0.74 and a median of 0.03.[2]

This data set allows us to answer a key question: How do daily changes in the supply of fake news articles affect browsing behavior? We use a measure of fake news supply borrowed from Allcott and Gentzkow (2017; hereafter AG), who assembled a database of fake news articles that were confirmed to be false by three independent fact-checking services (Snopes, Politifact, and Buzzfeed). We find that fake news browsing behavior is responsive to changes in the supply of fake news articles; in particular, a weekday 10-article increase in fake news increases the consumption of such news by 3.0%. Our reduced-form analysis considers other factors that may expose users to fake news; examples include polling numbers, Twitter posts, and TV mentions. We also conduct additional tests using placebos and alternative dependent variables to strengthen our identification. For instance, when we relate the AG supply measure to news sites of a more mainstream nature (those that host what we call "credible" news, as distinguished from fake news) or to sites that are

[1]For more details on this project, see https://github.com/mozilla/miracle. User data are anonymized.

[2]Allcott and Gentzkow (2017) use several approaches to estimate that the average American adult viewed one or two fake news articles during the election; our findings are generally in line with those estimates (see Section 3 for additional details).

unrelated to news, we do not find similar effects on browsing behavior. To help rule out reverse causality, we regress fake news site visits on the following day's fake news supply; the estimate is close to zero and not statistically significant.

However, this analysis does not address other possible sources of endogeneity in our supply measure. In order to address the question of simultaneity in our supply and consumption measures, we conduct two tests. First, we interact fake news supply with a user's number of visits to social media websites—thus (arguably) capturing exogenous and idiosyncratic exposure to fake news. The assumption here is that social media visits are not systematically driven by fake news consumption and hence that social media visits lead to idiosyncratic exposure to fake news by each user on each day. Second, we use a machine learning prediction model to isolate variance in the fake news supply that could well be exogenous to user consumption. This approach relies on the implausibility of those fake news creators who are motivated by ideological concerns or short-term advertising revenue coordinating their levels of production with consumer behavior. The results of both these tests are consistent with earlier findings, and again neither placebo tests nor using alternative dependent variables affects the results.

Finally, we look at two other questions about consumer behavior. First, do increases in the supply of fake news increase the likelihood that a given consumer will visit at least one fake news site? In this regard we find evidence that a weekday 10-article increase in the supply of fake news also increases the odds, by 3.7%, of visiting at least one fake news site. Second, we ask whether users used fake news as a substitute for credible news as the former's supply increased. Here, our models do not deliver enough statistical significance to make any claim one way or the other. Overall, we conclude that the efforts of fake news producers not only increased consumption among those who might be so inclined but also had the effect of "spreading" fake news to those who would otherwise have been less likely to consume it.

This paper follows the line of research that looks at online information and misinformation as well as individual responses to them. Kumar et al. (2016) show how instances of misinformation (on Wikipedia) can proliferate online. Other work explores how the proliferation of information and misinformation differs on social media (Del Vicario et al., 2016). In the context of online social networks, scholars have also studied how network composition (Bakshy et al., 2015) and the timing of content (Gabielkov et al., 2016) affect the consumption of different sources of credible news and fake news. Effects of the production, diffusion, and consumption of untruthful news have a measurable societal impact; Enikolopov et al. (2011) show that exposure to independent news sources—rather than state-sponsored (and presumably biased) news—reduced voting for the incumbent party, increased voting for opposition parties, and led to lower overall voter turnout.

More recently, researchers have focused on fake news and the 2016 US presidential election. Allcott and Gentzkow (2017) provide a basis for understanding the production of fake new and offer evidence on how it was shared on social media. Fourney et al. (2017) examine the relationship between social media and fake news; that paper establishes a correlation between fake news consumption and aggregate voting patterns. Vosoughi et al. (2018) examine fake news from 2016 and 2017 and find that it spread much more rapidly than did credible news. Guess et al. (2017) investigate the differential consumption of fake news based on prior browsing history.

More broadly, our findings—which accord with those of Cagé et al. (2016), who investigate the production and consumption of general online news—are among the first to document a relationship between fake news and day-to-day browsing behavior. Prior research has demonstrated the importance, to societal outcomes, of consuming misinformation; it follows that understanding the impact of its production is necessary if we are to comprehend the scope of the problem that fake news poses during elections.

Our test is agnostic to the channel through which any given fake news article might be distributed. There is a preponderance of attention being paid—in the press and by government—to how social media sites affect the spread of fake news. Our question is complementary to the research on diffusion channels. Thus we ask: Given a network for distributing news and information, what effect do *changes* in the supply of fake news have on the consumption of fake news–related content? In other words, we are interested in the net effect that supply has on consumption irrespective of how users come upon these types of articles. Our primary goal is to estimate the overall impact of fake news on the US electorate, so we are less concerned about the particular means of distribution.

2 Background

News amounts to a market for political information: consumers seek information about the political world around them, and news outlets produce content to satisfy those demands (Hamilton, 2004). Online news, in particular, exemplifies this market mechanism given its substantially lower fixed and marginal costs of production and distribution. As online news outlets have proliferated, consumers have been given greater choice in their sources of political information.

It is well documented that consumers generally seek political information favoring their partisan identities (Gentzkow and Shapiro, 2006; Stroud, 2008; Sears and Freedman, 1967; Taber and Lodge, 2006; Knobloch-Westerwick, 2014). Since political identity is considered to be fairly consistent over time (Abramowitz and Saunders, 2006; Greene, 2004; Levendusky, 2009; Huddy, 2001), it seems likely that consumer news preferences are similarly inelastic notwithstanding an increasingly diverse market. However, the growth of social media as a means for distributing online content (including news) has encouraged consumers to "recommend" news within their social networks, where endorsement from trusted network ties is more likely—than is the original source's partisan leaning—to translate into clicking on an article (Messing and Westwood, 2014). This literature suggests that political identity affects consumer preferences for news articles and that social media networks can amplify the impact of that identity on the demand for specific types of news content, especially content that caters to one's political in-group.

Some formal models from media economics suggest that, as media consumers grow more heterogeneous in their information preferences, producers place less emphasis on accuracy (Mullainathan and Shleifer, 2005). This trend is likely exacerbated in the case of online news producers, who face greater competition in the form of consumers' increased switching behavior. Preliminary evidence suggests also that access to broadband Internet can lead to greater polarization (Lelkes et al., 2017); the implication is that, as more news consumption moves online, the incentives for minor producers to produce high-quality, veridical news content may be no match for the profitability of fake news—that is, because fake news producers need not vet sources or substantiate claims and so their "news" content is much cheaper to produce.

Thus the economic literature suggests that there are strong incentives to produce fake news: in comparison with credible news, it is not only cheaper and hence more profitable to produce but also easier to distribute given both the facility offered by social media networks and the psychological factors affecting consumer choice. If the mere presence—in the online news market—of fake news increases its likelihood of consumption, then the effects on society could be substantially negative. An extensive literature looks at the relationship between exposure to news and a variety of societal outcomes. For instance, DellaVigna and Kaplan (2007) show that the presence of Fox News in local news distribution can increase presidential vote share by half a percentage point; this increase is larger than the vote share for several key electoral college victories in the 2016 presidential election. Exposure to newspaper endorsements also has been shown to persuade voters (Gerber et al., 2009; Chiang and Knight, 2011). DellaVigna and Gentzkow (2010), Dilliplane (2011), and Huddy et al. (2015) review similar relationships between exposure to political information and such other indicators of civic engagement as campaign participation, political donations, and voter registration.

It follows that estimating the functional relationship between the supply of fake news and the likelihood of consumer exposure is highly relevant in today's political climate. It may be that minimal exposure to fake news can interact with strong political beliefs or robust partisan identity. Thus early stages of political attitude formation could impose higher evidentiary standards on later stages (Erisen et al., 2014; Lord et al., 1979), which suggests that counteracting the beliefs formed through fake news exposure may require an incommensurate amount of factual evidence. Classical "motivated reasoning" theory (Kunda, 1990) has been documented in the literature addressing the formation of political attitudes and opinions; that is, individuals strongly dislike challenges to their closely held political beliefs (Taber and Lodge, 2006). People may well become even more entrenched in their views when confronted with information that contradicts their world view (Redlawsk, 2002)—a phenomenon suggesting that, once a person whose political identity is strong has formed a particular political belief, that belief cannot be dislodged even by undeniable factual evidence (Friesen et al., 2015). This dynamic is clearly problematic when one considers that falsehoods are more often believed when they concern the political *opposition* (Weeks and Garrett, 2014).

In short, this theoretical narrative highlights the importance of understanding the relationship between the supply and the consumption of fake news—namely, whether or not its mere presence is sufficient to increase such consumption.

3 Data

Our individual-level data come from the Mozilla Corporation, and we collect additional time-varying covariates from a number of sources. These data sources are described in this section.

3.1 Mozilla Context Graph

Mozilla Corporation develops the Firefox Web browser, which has been downloaded more than a billion times. The firm states that it has more than half a billion users—a market share of between 9% and 16%. Firefox was created in 2002 by many of the same developers who worked on Netscape Navigator. In 2016, it launched the Context Graph initiative to build a "recommendation" system for the Web. As part of this project, Mozilla randomly recruited US users for an opt-in program to download a browser extension that would track their browsing behavior on desktop computers. A blacklist of the sites that would not be tracked is listed on the project's GitHub page.[3] Users are identifiable only by a unique identifier and are therefore anonymous in our analyses. The study was conducted from September 20 to December 16, 2017; we focus on browsing activity before the election date of November 8, 2016.

Mozilla randomly recruited participants using its Shield tool, which allows the company to ask "general release" users to participate in various studies. The US-based users who opted in had a browser add-on downloaded to their computer in the background and without additional intervention. Initial sign-up rates for the project were low, so Mozilla conducted two drives to recruit additional random users. The opt-in (participation) rate for users who received a notice was between 1% and 1.5%. Of course, the resulting sample was biased to

[3] More details on the project and the blacklisted sites can be found at https://github.com/mozilla/miracle. Most of the untracked sites contained sexual content.

the extent that users' opt-in decisions were associated with user proclivities that differed from those of the general population. Absent more detailed demographic information about the users, we gather summary statistics. Allcott and Gentzkow (2017) estimate that the average adult saw and remembered between one and two fake news articles during the election. Users in the Mozilla study visited 0.73 fake sites a day, on average. However, using the median of 0.03, our results suggest that users visited approximately 1.5 fake sites during the 49 days of our sample period—in line with the AG estimate. Note that their estimate is based on the *recall* of fake news, which we do not capture. We surmise, then, that our opt-in users are not markedly different from the average US-based Internet user.[4]

Users could join the study at any time—and, indeed, individuals downloaded the add-on up until the last day of the Context Graph initiative. Figure 1 plots the trend of user arrivals and shows that Mozilla's late-September recruitment drive was effective. Participants who were inactive (i.e., who visited no sites during any span of seven consecutive days) were excluded from the sample. On the eve of the presidential election, the Context Graph project was tracking 1,418 users.

Participants visited an average of 70.7 unique sites per day during the study period. Figure 2 plots the average number of site visits (of all users) on each day. Average daily usage was volatile during the first four days of the period, at which time the sample contained only 30 users.[5] Consistent consumption prevailed during the rest of our study period, with troughs regularly occurring on weekends.

In addition to collecting counts of total sites visited, we collect three subsets of site visits from the users: visits to fake news sites, to credible news sites, and to social media

[4]In Section 4.8 we compare Mozilla-based browsing behavior with that derived from Pocket, another source of online browsing activity and one that includes not only desktop but also mobile browsing. We find comparable trends, which provides additional corroboration of our sample's generalizability.

[5]Further inspection of these 30 users revealed that the initially volatile average site visit behavior was mainly driven by a single user, whose average number of site visits during the study period was the sample's third highest.

sites. We define each category as follows. In characterizing Web domains as "fake news sites", we use the union of Zimdars' (2016) publicly available OpenSources website and AG's database of fake news articles (to be detailed in what follows).[6] OpenSources tags sites that "entirely fabricate information, disseminate deceptive content, or grossly distort actual news reports"; there were 610 domains so classified as of January 31, 2017.[7] Although the most popular fake news articles were those maligning Hillary Clinton's candidacy, other articles were critical of Donald Trump (Allcott and Gentzkow, 2017; Guess et al., 2017). We also reference a list of news domains that OpenSources categorizes as "credible"—in other words, news and information that is circulated in a "manner consistent with traditional and ethical practices in journalism."[8] That list of credible news sites includes organizations from across the political spectrum (e.g., foxnews.com and msnbc.com). We supplement this list with websites of the top 50 US newspapers (by circulation). Finally, we define social media visits as those made to any of these six domains: Facebook, Twitter, LinkedIn, Instagram, Snapchat, and Pinterest.

The individual-level, time-varying summary statistics from both the Context Graph project and our sample are reported in Table 1. For our pooled models, we have a sample of 1,418 users who participated in the Context Graph project by the end of the sample period, or 32,020 user-days. The summary statistics for all users is presented in Panel A. Users visited an average of 0.69 and 1.08 fake and credible news sites each day, respectively. Average daily visits to social media sites were 4.18.

Of the 1,418 users, 625 did not browse any fake news sites and 28 users appeared in the data for only one day. For our fixed effects models (which excludes those 653 users that

[6]Some fake news articles were hosted by general content publishing platforms or other . Because these domains host a substantial amount of content unrelated to fake news, we exclude them from our counts of user fake site visits. See the Data Appendix for more information on how the domain lists were assembled.

[7]See http://www.opensources.co/ for more information.

[8]OpenSources no longer maintains this list of credible news sources. However, it can be seen by viewing (at Github) changes made to the database on April 2, 2017.

exhibit no variation with respect to the dependent variable), we are left with a sample of 765 users; thus we assemble an unbalanced panel comprising 19,325 user-days and on which our fixed-effects regression analyses are performed. This sample is the focus of Panel B in the table, which shows that these users visited an average of 1.15 fake news sites each day, as compared with the mean of 1.50 daily visits made to credible news sites. The mean of daily social site visits is 5.97. These numbers reflect substantial dispersion: median fake site visits is zero. It is thus clear that some users consumed a significant amount of fake news content, which underscores the importance of using a fixed-effects estimation.

3.2 Fake News Database

In order to measure participants' sensitivity to fake news, we use AG's database of fake news articles. The authors collect fake news articles across three fact-checking sites and articles that were shared extensively on Facebook. The database contains 536 fake news articles that were published during our study period from September 20 to November 7, 2016, the day before the election. Their list is certainly not exhaustive, but each article is confirmed by at least one of three fact-checking services (Snopes, Politifact, and Buzzfeed) as being fake and the list captures a large portion of the most popular topics on fake news outlets.[9] The production and copying of fake news articles was widespread across blogs and alternative news sites, so we would expect the content of many of the articles in the AG database to appear on other websites. Of the 375 domains in the complete AG database (with articles from May 15, 2016, to November 11, 2016), 130—or 35%—matched those on the OpenSources list.

[9] We refer readers to Allcott and Gentzkow (2017) for further details on the fake news database.

3.3 Other Independent Variables

Finally, we gather supplemental data that may tell us more about users' browsing choices. The first of these sources is the GDELT Project, which scans online media and news sites for trends; we download the number of times each presidential candidate was mentioned online and on television. The second source is MIT's Electome project, which is part of the Laboratory for Social Machines at MIT's Media Lab. During the presidential campaign, Electome separately tracked the proportion of Twitter discussions about candidates and about topics. To calculate Donald Trump's tweeting frequency during the study period, we use the Trump Twitter Archive site. We collect polling data on the two candidates from FiveThirtyEight, which published a "poll of polls" prior to the election. Variables that vary by date are summarized in Table 2.

3.4 Implications of Measures for Estimates

Here we provide some additional context for the data just described. First, our measure for consumption (the dependent variable) is the number of actual visits to sites categorized as fake news. This means that our data do not reflect instances of users reading an article or headline (on a social media site, for example) *without* clicking through to the hosting site—even though they might well have, in effect, consumed the content in question. Second, we capture visit counts to top-level domains but not the specific article link visited by the user on that site. In other words, if a user clicked directly on a news article on cnn.com, we would only observe that as being a visit to the cnn.com domain, rather than a visit to the specific article link or any subdomain. Third, we do not have data on length of time spent on the domain or on its number of links. The direction in which our estimates would be biased by these data deficiencies is not clear ex ante: users might stay for only a few seconds, and without actually consuming any content; or they might linger for several minutes while reading several articles. We are thus limited in our ability to make more detailed conclusions about user engagement.

Fourth, our measure for the supply of fake news content is derived from the AG database, in which all the articles were deemed to be false by Snopes, Politifact, and/or Buzzfeed. In other words, the articles in our supply measure can be seen as *exemplars* of fake news during the election in that they achieved enough notoriety for fact-checking sites to assess their veracity. The main trade-off when using this narrow supply measure is that we undercount the actual number of fake news articles. Yet our supply measure may not be problematic in that its very narrowness should more accurately capture articles that are confirmed to be false—that is, rather than being simply polemical or unverified.

4 Analysis

4.1 Descriptive Observations

We start by graphing the data in order to identify any browsing behavior trends in the run-up to the election. Figure 3 plots the daily percentage (for the full sample) of visits that are to credible news sites. We observe a flat trend in news consumption until Election Day.[10] As expected, the demand for news spiked on Election Day and persisted at a high level for several days; it then reverted to the previous trend. The takeaway from Figure 3 is that consumption of credible news was consistent overall—that is, except for the spike corresponding to Election Day itself.

Figure 4 replicates Figure 3 but with visits to sites hosting fake news instead of credible news. The trend line here tells a different story: fake news consumption per user increased

[10]Recall from Section 3.1 that there is some volatility in the earlier part of the sample period, when the study tracked a smaller number of users. When we run the primary specification from column (3) of Table 3 while *excluding* the first week of data (results not shown), the results are consistent with those reported in column (3); this outcome suggests that our results are not driven by the browsing behavior observed during that initial period.

steadily during the period leading up to Election Day and then declined thereafter. However, the magnitudes are relatively small when we aggregate across all users, which suggests that most online activity concerns matters other than news.

A graph of the daily proportion of users who visited at least one fake news site is given as Figure 5. Prior to the election, the proportion of users who visited at least one such site ranged from 8.8% to 15.1% on weekdays. After the election, however, that proportion steadily declined and ranged from 6.7% to 13.7%. Taken together, Figures 4 and 5 indicate that fake news declined—in both its reach and extent of consumption by online news readers—following the presidential election.

4.2 Empirical Analysis

The phenomena we describe lead us to construct a stylized model of browsing behavior. In this our purpose is to estimate how the market's supply of fake news affects a user's incidence of visiting sites that are devoted to such news. This model posits that a user's visits to fake news sites are a function of several factors. Formally, we have

$$
\begin{aligned}
&\mathrm{E}[\mathit{FakeSites}_{it}] \\
&\qquad = \exp\big\{\alpha_1(\mathit{FNSupply}_t) + \beta_1(\mathit{Social}_{it}) + \beta_2(\mathit{Mentions}_t) + \beta_3(X_t) \\
&\qquad\qquad + \mathit{Weekend}_t[\alpha_2(\mathit{FNSupply}_t) + \beta_4(\mathit{Social}_{it}) + \beta_5(\mathit{Mentions}_t) + \beta_6(X_t)] \\
&\qquad\qquad + \gamma_i + \omega_t\big\},
\end{aligned}
$$

where $FakeSites_{it}$ is a count of how many fake news sites user i visited on day t. This behavior depends on: $FNSupply_t$, the number of fake news links generated on day t (in our measure, we divide the raw count by 10 for the sake of estimates that are more readily interpretable; thus a single-unit increase in our measure is equivalent to an increase of 10 fake news articles); $Social_{it}$, the number of visits to social media sites (which may affect a person's exposure to shared fake news articles); $Mentions_t$, which measures online news mentions of the candidates; X_t, which includes other time-varying political factors that might affect news consumption; γ_i, an individual-user fixed effect that absorbs unobserved heterogeneity in a user's characteristics and browsing habits (e.g., political affiliation, education, tastes for types of news); and ω_t, which includes calendar-week and day-of-week dummies to control for secular changes over time.[11]

It is crucial that the drivers of both news production and news consumption differ on weekdays versus weekends (Boczkowski, 2010). Our initial evidence from Figure 2 corroborates that observation. This is why we implement a fully interacted model—that is, one in which $Weekend_t$interacts with each of our time-varying covariates.[12] The resulting specification allows us to separately estimate the weekday effect (α_1) and the weekend effect ($\alpha_1 + \alpha_2$) of the supply of fake news on the number of fake news sites visited. We are interested in the change in incidence of visiting fake news sites (for which we use a count variable); therefore, we use a conditional fixed-effects Poisson model with robust standard errors clustered at the user level.

4.3 Results

In Table 3 we present our results from the Poisson model just described. We operationalize $Mentions_t$ by including measures of the count (in thousands) of daily online mentions of Donald Trump and Hillary Clinton; these counts are acquired from GDELT.

[11]We include calendar-week and day-of-week fixed effects because using only calendar-*date* fixed effects would absorb all the variation in our supply measure.

[12]The main effect of $Weekend_t$ will be absorbed by the day-of-week dummies, which are included in ω_t.

4.3.1 Base Model

In column (1) of the table we give results from the pooled Poisson regression that excludes user fixed effects, where standard errors are clustered at the user level. This regression allows us to include the full Context Graph sample of 1,418 users; however, we cannot control for their fixed characteristics because we have excluded user fixed effects. Our first result is that a 10-article increase in fake news publications during weekdays is associated with a 4.3% increase in the sample's incidence of visits to fake news sites ($\exp(\beta) = \exp(0.042) = 1.043$, $SE = 0.016$). We observe a similar magnitude effect for weekend days ($\exp(0.042-0.005) = \exp(0.037) = 1.038, SE = 0.043$), although it is not statistically significant at the 10% level. Because weekend results of subsequent models are similarly inconclusive, we shall focus our discussion on the *weekday* impact of the fake news supply.

In addition to the pooled regression, we run a zero-inflated Poisson model and present results in column (2). The zero-inflated Poisson accounts for the high proportion of zero fake site visits by separately modeling whether there are zero visits or at least one and then modeling the count. This specification also does not include user fixed effects, allowing for the consideration of all the data. The logit stage predicting zero values was modeled using the same regressors as the Poisson model (results for the initial stage logit model are not shown for brevity). In this model, a 10-article increase in fake news publications increases fake news site visits by 2.7% ($\beta = 0.027$, $SE = 0.016$), which reflects a magnitude in line with the results from the fixed effects models that follow.

Our subsequent regressions include user fixed effects, so they omit users who did not visit a fake news site or who generated only one day's worth of data during the sample period. Column (3) of the table shows the results for our base model, from which we conclude that a 10-article increase in fake news publications on a given weekday increases a user's incidence of fake news site visits by 3.0% ($\beta = 0.029$, $SE = 0.014$). How large is this magnitude? We answer this question in two ways. First, it would take 236 (calculated as $\exp(x\beta/10) = 2$) fake news articles to double a user's fake news site visits. Second, a 100-article increase in the number of fake news articles would increase such visits by 34.1% (calculated as $\exp(10\beta)$). Since the sample's average user visited 1.25 fake news sites each weekday, it follows that a 34.1% increase is equivalent to an increase of 0.43 sites. This implied click-through rate of 0.43% is similar to the rates observed for display advertising, which (according to industry estimates; see Volovich, 2016) range from 0.1% to 0.9%.

Our base model reveals also that an increase in the number of social media website visits during weekdays increased the number of fake news site visits and that this effect is stronger during weekends: each additional social network website visit increased the number of fake news site visits by 2.1% ($\beta = 0.021$, $SE = 0.007$) and by 3.8% ($\beta = 0.021 + 0.017 = 0.038$, $SE = 0.011$) on weekdays and weekends, respectively. This finding accords with anecdotal and empirical evidence that users rely on social media for news (Gottfried and Shearer, 2016 and Boczkowski, 2010 report that 62% of US adults rely on social media for news) and that social media serves to "amplify" fake news (Vosoughi et al., 2018). We remark that our estimate is likely to be biased downward because our dependent variable does *not* capture (a) any fake news consumption on social media that was not accompanied by a visit to the underlying domain or (b) the consumption of false content that resided on social media only.

With regard to online discussion of the candidates, weekday online mentions of Hillary Clinton increased visits to fake news sites ($\beta = 0.025$, $SE = 0.018$) but online mentions of Donald Trump reduced such visits ($\beta = -0.024$, $SE = 0.017$); however, neither of these results is statistically significant at the 10% level.

In the following section, we include additional possible alternative explanations to the base model discussed above. These regressions include (a) Twitter measures that proxy for candidate and topic news, (b) changes in the political environment, and (c) offline coverage of the candidates. The addition of these covariates in columns (4)–(8) does not substantially affect the relationship between fake news supply and fake news site visits. Our estimate of the coefficient of fake news supply on browsing ranges from 0.024 (Column (6), $SE = 0.013$) to 0.034 (Column (7), $SE = 0.014$).

4.3.2 Alternative Explanations

In addition to online news mentions, we examine whether the relative level of social media discussion of the two candidates affected fake news site visits. We test for this possibility by using data from MIT's Electome project to measure the popularity of candidates and issues on Twitter. Although not all Internet users are active on Twitter (24% of them are, according to Greenwood et al., 2016), it is plausible that Twitter content reflects contemporaneous discussion of salient news topics. The MIT data give the share of Twitter discussion that concerns a particular candidate; daily changes in this measure are included in the regression whose results appear in column (4) of Table 3. The effect of those changes on users' visits to fake news sites during weekdays is both positive and statistically significant at the 10% level (albeit marginally so; $\beta = 0.005$, $SE = 0.003$).

Visits to fake news sites may also be affected by the salience of certain news topics. To consider the discussion of particular political issues, rather than of the candidates, we include the top two news topics on Twitter at the time; according to Electome, these were immigration and the economy. Results from these regressions are presented in column (5) of the table. Here we see no evidence that discussion of topics prominent during the campaign influenced browsing behavior. In particular: neither immigration ($\beta = -0.001$, $SE = 0.001$) nor the economy ($\beta = 0.0002$, $SE = 0.0008$) have an economically or statistically significant effect on weekday browsing behavior, which is rather surprising in light of the contentious tenor of most discussions on these topics. A candidate will often link the issues of immigration and the economy, so the difference in their respective proportions of Twitter mentions is not relevant. It is worth noting that the estimate of our independent variable of interest, $FNSupply_t$, is consistent with our other specifications in terms of magnitude ($\beta = 0.027$, $SE = 0.018$); however, this estimate does not quite reach the level of statistical significance.

With regard to the candidates' social media activity, we consider the impact of posts by Donald Trump—who averaged 15 tweets a day over the seven weeks preceding the election. We assess how Trump's own tweets, which garnered substantial news attention, affected fake news site browsing. As shown by column (6) in Table 3, the number of daily Trump tweets had a positive (but not statistically significant at the 10% level) effect on fake news site browsing ($\beta = 0.004$, $SE = 0.003$). Although Trump's tweets are unlikely to constitute an exogenous source of variation in our model, this result is intriguing nonetheless because many commentators believe that Trump's tweets could have played a role in promoting false or unverifiable news reports. We do not find support for this claim in our results.

In addition to discussion of the election, the issues, and the candidates, other political factors may have affected fake news browsing. One possibility that may explain our results is that the supply of fake news simply reflected how close the race was. In that case, our results would reflect not supply–demand calculus but instead other factors related to predictions about each candidates' odds of winning. Column (7) of the table accounts for this dynamic by including a variable that measures the spread between the poll numbers for Clinton and Trump as posted on the FiveThirtyEight political website, which aggregated polling data from major polling organizations to produce a daily prediction for each candidate's share of the popular vote. We use the spread—that is, the difference between Clinton's and Trump's predicted vote shares—as a covariate. The estimated coefficient is negative ($\beta = -0.055$), which does indicate that fake news consumption increased as the race tightened (i.e., as the spread decreased). Yet the value derived is not statistically significant ($SE = 0.061$), and its inclusion does not affect the coefficient for the supply of fake news.

Finally, if the pathway to consumption of either credible or fake news involves exposure to traditional media sources (viz., television), then controlling for only online mentions of the candidates would be inadequate for our purposes. That is, an exclusive focus on Internet-based measures could well result in our missing an important factor that reduces (or even eliminates) the observed effect of fake news supply. We address this concern by including, in column (8) of the table, TV mentions of each candidate as gathered from the GDELT database. The effect's direction for each candidate is consistent with the online mentions measure used in column (2), but again the estimates are not statistically significant (for Trump: $\beta = -0.013$, $SE = 0.027$; for Clinton: $\beta = 0.010$, $SE = 0.040$).

4.4 Placebo Tests and Alternative Dependent Variables

We consider three alternative specifications that serve as a "placebo" test and as tests of alternative dependent variables for the results in Table 3. These tests, whose results are reported in Table 4, include time-shifting the focal independent variable and regressing site visits other than those devoted to fake news on consumption behavior. In column (1), we regress fake news visits on one-day-forward fake news supply. News cycles are essentially contemporaneous, so we should not expect a user's browsing to be affected by the *next* day's fake news (absent spurious relationships in the data). Moreover, a finding of no effect would constitute evidence against reverse causality, whereby higher fake news consumption drives higher fake news production. Indeed, we find that the coefficient for our forward measure of fake news supply during weekdays is near zero and statistically insignificant ($\beta = 0.001$,$SE = 0.014$).

In columns (2) and (3) of the table, we regress different types of site-browsing behavior on our fake news supply measure. Column (2) tests the relationship between fake news supply and consumption of credible news (recall from Section 3.1 that the latter are mainstream and local news sites—that adhere to basic reporting standards and are considered credible by OpenSources—and the top websites for newspapers by circulation). On the one hand, not finding a relationship between fake news supply and the number of visits to *credible* news sites would tend to validate the relationship between fake news supply and the number of visits to *fake* news sites. On the other hand, finding such a relationship would amount to initial evidence that fake news production and consumers' browsing of credible news are complements (if the relationship were positive) or substitutes (if it were negative). Column (2)reports the results when we use credible news site visits as the dependent variable. The magnitude of the weekday estimate is positive, but it is smaller ($\beta = 0.011$) than in the fake news site visits regressions and is not statistically significant ($SE = 0.011$). Therefore, we cannot identify a relationship between fake news supply and credible news site visits. Column (3) repeats this analysis, but here the dependent variable is the number of sites visited *other* than credible news sites, fake news sites, or social media sites. Again we find that the magnitude of the weekday estimate is much closer to zero ($\beta = 0.002$) than in the fake news site visits regressions and is also statistically insignificant ($SE = 0.005$). These near-zero results give us some assurance regarding our identification of the relationship between the supply of fake news and its consumption.

4.5 Identification and Isolating Exogenous Variation in the Fake News Supply

One of our model's underlying assumptions is that the supply of fake news is exogenous to its demand. According to Allcott and Gentzkow (2017), there are two primary motivations for the production of fake news: ideological and economic. We argue that ideologically motivated producers of fake news were more likely producing content to *shape* demand than in *response* to it. Furthermore, we can subdivide the economic actors into (a) "traditional" content producers, who are probably responsive to demand when making their supply decisions, and (b) "pop-up" sites that opportunistically seek to generate advertising revenue around the time of the election by consistently releasing fake news. Press coverage during and after the election noted that some of the latter producers were purveying both liberal and conservative news (Silverman, 2017), that many were based in Eastern Europe or other foreign locations, and that most were producing a high volume of daily articles so as to generate increased ad revenue (Silverman and Alexandar, 2016; Subramanian, 2017). These observations suggest that such pop-up sites could be responsive to consumer behavior with regard to the type—but not the levels—of content produced (Boczkowski and Mitchelstein, 2013). In that event, ideologically driven or pop-up sites would be (as compared with traditional sites) more exogenous to demand, especially in the weeks leading up to the election. So far, then, our identifying assumption is that our fake news measure mainly captures pop-up and ideologically motivated sites, since little production of fake news is due to traditional content producers. However, if this assumption is suspect then there could be simultaneity between our dependent variable (*visits* to fake news sites) and our independent variable of interest (the *supply* of fake news). So if these two variables reflect a traditional supply–demand relationship, then we could end up identifying shifts in demand instead of shifts in supply. The placebo test—whose results are presented in column (1) of Table 4—partially addresses this concern, but we it would be prudent to find ways of isolating exogenous variation in our measure of fake news supply. For this purpose we adopt two approaches: by identifying idiosyncratic exposure to fake news based on an individual's social media browsing habits; and by limiting our measure of fake news supply to the sites most likely to be ideologically motivated or of the pop-up type.

4.5.1 Social Media Visits as Idiosyncratic Exposure to Fake News

Taking advantage of the facts that (a) social media drove users to fake news (Allcott and Gentzkow, 2017) and (b) fake news is much more likely than credible news to spread over social networks (Vosoughi et al., 2018), we interact fake news supply with individual-level social media site visits and report the regression results in Table 6. If a person's social media behavior on a given day is effectively idiosyncratic (i.e., if individuals visit social media for reasons that are not systematically correlated with a day's supply of fake news), then this interacted measure of fake news supply is more likely to reflect exogenous supply to an individual on that day with the inclusion of user fixed effects than the count of articles used in our base models. Survey evidence from Pew Research Center (2016) indicate that 81% of respondants said they did not share news about the election on social media, suggesting that most social media users are not being driven to those sites in order to share news. Thus, fake news daily supply interacted with individual daily visits to social media sites renders our aggregate fake news supply measure into a measure of individual-level exposure to the fake news supply—exposure that varies over time and across individuals based on their social media behavior.

The intuitive interpretation of the coefficient for the interaction term is fake news supply *weighted* by a user's social media habits. As seen in column (1) of the table, the weekday estimate of social media–adjusted fake news supply has a positive effect on weekday fake news site visits ($\beta = 0.0009,\ SE = 0.0005$). To facilitate interpretation of the coefficient, we compare the effect of changes in levels of fake news article production on an individual who made the median number ($= 1$) of daily social media site visits with the effect on an individual whose number of social media visits was at the 95th percentile ($= 19$). For the former, a production increase from 10 to 50 fake news articles resulted in an increase of 0.01 in the number of fake site visits; the corresponding increase for the latter was 0.21.

As we did when generating Table 4, we run placebo tests and tests of alternative dependent variables on this other measure of fake news supply. We use: (a) a forward measure of our fake news supply interaction as the independent variable, as shown in column (2); and for alternative dependent variables, (b) visits to credible news sites (column (3)) and (c) visits to other sites (column (4)). The weekday coefficients are estimated to be near zero and none is statistically significant. The consistency of these findings with the main results and the nil results in the placebo and alternative dependent variable tests corroborate our main result of the relationship between fake news supply and fake news site visits.

4.5.2 Identifying Exogenous Producers of Fake News

Our assumption that all fake news production is exogenous may be too strong, since some sites may coordinate their production based on news cycles or on other variables not reflected in our specifications. In that case, we should refine our supply measure to include only those sites most likely to be exogenous to consumption. We do so by leveraging institutional facts about the production and proliferation of fake news during the election. In this approach, we define more explicitly the conditions under which site content could be exogenous to demand by observing which domains continued to host content after the election.

To identify which news was more likely to be produced by exogenous sources, we look at news article domains. The 948 articles in the AG database were hosted on 375 unique domains. When we inspected those domains in June 2017, 102 of them (about 27%) had been shut down. Clicking on their URLs resulted in output such as "Buy this domain"—if the browser was able to find the domain at all. Another relevant fact is that, of the 102 sites that were no longer live, 90 were established during or after 2015. These two observations are consistent with the notion that many of the fake news sites popped up around the election with the aim of producing "clickbait" to generate short-term revenue. We therefore assume that if a site was no longer live, then it probably was a pop-up site set up to capture advertising revenue or to influence opinions during the election and was then shut down.

This assumption about a site's objectives is by definition an ex post assessment because we gather data from June 2017, well after the election. We therefore devise a measure that is defined ex ante by following Guzman and Stern (2015), who use certain characteristics—including company registration—to predict whether a startup would later achieve a successful exit. Similarly, we use characteristics included in a domain's registration to predict whether the site would eventually be shut down. Acquiring a domain requires that one register it with the Internet Corporation for Assigned Names and Numbers (ICANN), a nonprofit entity that manages domain names. At the time of registration, certain details must be provided; these include the name of the domain and the desired extension (e.g., ".com", ".net", ".org"), the names and contact information of the registrant, and the site, billing, and technical administrators. The date of registration is also part of the record. We worked with DomainTools, an online security company, to acquire the domain registration information for each of the fake news sites. Using that information, we created a set of 362 features (plus an intercept), which included: (a) the creation date, year, year-squared, and individual year dummies; (b) whether a billing contact was named, (c) registered name types (individual, institution, private), (d) the domain extension, (e) state and country of registration; and (f) indicators for the appearance of certain words or phrases in the domain name (e.g., "patriot", "trump", "america"). This set of features also included interactions among (i) the creation year, year squared, and 2016 registration and (ii) other indicators. Each of these defined features (along with the weights assigned to each by the variable selection algorithm, described next) are listed in the Data Appendix.

We use our dependent variable or outcome (i.e., the domain's eventual shutdown) and the independent variables or features (derived from the domain's registration information) in a predictive logit model. Because there are nearly as many features (362) as observations (375), we adopt an approach to variable selection and model fitting that is commonly used in the field of machine learning (Friedman et al., 2010). Machine learning methods are increasingly being leveraged in social science research (Athey and Imbens, 2017), with techniques such as LASSO often used in applications such as instrumental variable selection (Gilchrist and Sands, 2015). We used the LASSO method of variable selection and, to ensure the predictors' out-of-sample validity, we used 10-fold cross-validation: 10 iterations are run, where 90% of the data is used to estimate the coefficients and 10% is withheld as testing data.

This procedure yields coefficients (or weights) for each of the features, and we use those weights to predict the likelihood of a site being shut down (see the Data Appendix for more information on the features and the weights by the LASSO procedure). We place no conditions or constraints on the algorithm for assigning weights to features. Of the 363 features, the algorithm assigned a non-negative weight to 146 (of which 96 had an absolute weight exceeding 0.001). Although the weights lack any meaningful economic interpretation, it is noteworthy that both the date of registration and our dummy for 2016 registration have positive weights; these outcomes corroborate our initial observations relating a site's date of registration to the likelihood of it being shut down.

As for the final predictions, values close to 1 are indicative of domains predicted to be shut down. Examples include sonsoflibertymedia.com (predicted value 0.818), buzzfeedusa.com (0.895), and 365usanews.com (0.969). At the other extreme, values close to 0 indicate domains predicted to remain live; examples include liberalamerica.org (0.001), informationliberation.com (0.015), and libertynews.com (0.034). The correlation coefficient between our predictive measure and the binary measure of whether the site actually was down is 0.79. If we "discretize" the predicted value as being 1 (site down) when the continuous measure is greater than 0.5 and as 0 otherwise, then our procedure properly classifies 332 of the 375 domains (89%), misclassifies 16 live sites as being down (out of 273), and misclassifies 27 down sites as being live (out of 102). For the 16 false positives, the predictions range from 0.515 to 0.750 with a median of 0.634; for the 27 false negatives, the predictions range from 0.107 to 0.498 with a median of 0.305. That the values tend toward 0.5 for the 43 misclassified domains indicates that the classifier did not grossly err in those cases.

We then associate each news article with the predicted probability that its host domain will be shut down, using that value to create a predicted daily fake news article count as our measure for $FNSupply_t$. With this new count, we replicate our main findings reported in column (2) of Table 3. The results, which are presented in column (1) of Table 6, are the same as our main findings in terms of direction. In fact, the magnitude of the effect is substantially larger: a 10-article increase in the predicted fake news article count increases fake news site visits by 10.0% ($\beta = 0.096$, $SE = 0.049$). Columns (2)–(4) replicate the placebo and alternative dependent variable tests from Table 4 but using a one-day-forward measure of our predicted fake news supply (column (2)), and replacing the dependent variable with counts of credible news site visits (column (3)) or visits to other sites (column (4)). As before, the estimates for the respective independent variable of interest are closer to zero and not statistically significant at the 10% level.

We also consider other measures based on adjusting fake news counts by motivation (to save space, these results are not reported). First, we dichotomize the variable, assigning a value of 1 when the continuous measure exceeds 0.5 and otherwise assigning a value of 0.. Second, we relax the definition of a site being down to include 50 sites that were technically live but were evidently not being maintained; for example, there were no recent news posts. Using this looser definition of a down site, we rerun the prediction model to derive a "loose definition" predicted daily fake news article count for use in our regression models. In both instances, the direction of the effect was consistent with our main results: weekday magnitudes were 10.1% ($\beta = 0.097$, $SE = 0.051$) using the dichotomized predicted measure, and 6.0% ($\beta = 0.058$, $SE = 0.036$) using the looser definition.

4.6 Other Effects of Fake News Supply on Consumption

We now extend the analysis by asking two additional questions of practical importance. First, how did changes in the supply of fake news affect the likelihood of visiting a fake news site? Second, did increases in the supply of fake news substitute for or complement the consumption of credible news?

4.6.1 Probability of Consuming Any Fake News

Our estimation strategy has until now been to consider the effect of the supply of fake news on the visits to fake news sites. What may also have implications for responses is whether the supply of fake news had an effect on a user's propensity to consume any fake news at all. In other words: Did the supply of fake news lead non-consumers of fake news to become consumers of fake news? To answer this question, we define a new dependent variable for fake news site visits: a dummy set to 1 if the user visited at least one fake news site that day (and set to 0 otherwise). Here we use a fixed-effects logit model with bootstrapped standard errors (SEs are clustered at the user level in the pooled model). In Table 7, columns (1) and (2) report the results of (respectively) the pooled logit model and the fixed-effects logit model. Focusing on the fixed-effects estimates in column (2) reveals that, during weekdays, a 10-article increase in the supply of fake news articles increased the odds of visiting at least one fake news site by 3.7% ($\beta = 0.036$, $SE = 0.021$). Column(3) of the table replicates the placebo test in which we predict visiting at least one fake news site using supply from the one-day-forward period; the weekday effect is substantially lower—near zero—and statistically insignificant ($\beta = 0.001$, $SE = 0.016$). In column (4) we create an analogue of our dichotomized measure but for news site visits; that is, the indicator variable is set to 1 if at least one credible news site was visited and set to 0 otherwise. Our estimate for the effect of fake news supply on visits to credible news sites is near zero and statistically insignificant ($\beta = 0.008$, $SE = 0.015$).[13]

These results provide evidence that the level of fake news production did positively affect the likelihood that a user visited at least one fake news site. The implications are that the supply of fake news not only drove the intensity of consumption for some individuals but also increased the likelihood of exposure to any fake news sites. Fake news supply thus has some influence on the "flipping" of a day from one on which no fake news sites were visited to one on which at least one such site was visited.

4.6.2 Did Fake News Supply Cause Consumers to Substitute Fake News for Real News?

Finally, we investigate whether the supply of fake news induced consumers to substitute consumption of fake news for their previous consumption of credible news. Our first evidence concerning this question comes from column (2) of Table 4. In that regression, finding a negative relationship between fake news supply and visits to credible news sites would lead us to surmise that increasing the supply of fake news had the effect of reducing the consumption of credible news. However, we do not find a negative relationship; the

[13]For our logit models we do not replicate the alternative dependent variable test using other (non-news) site visits because, when that variable is dichotomized, there is almost no variance in the measure; that is, users typically visited at least one site nearly every day.

estimate is positive and fairly close to zero ($\beta = 0.011$, SE $= 0.011$).

We also compute a proportion-dependent variable calculated as the portion of all news sites visited on a given day that were fake news sites. Using this dependent variable, we run fixed-effects OLS regressions with standard errors clustered at the user level. For these models, however, the F-tests (not reported) indicate that none is statistically significant. In light of these two sets of results, we cannot draw any conclusions about whether consumers replaced their credible news consumption with fake news when more of the latter was produced.

4.7 Implications of Estimates in the Week before the Election

To place our estimates in the context of the 2016 presidential election, we provide the following calculation. During the first six weeks of the study period (i.e., between 7 weeks and 1 week before the election), the mean and median number of verified fake news articles produced on weekdays were 9.9 and 4.5, respectively. In comparison, the mean and median in the study's seventh week (i.e., the week leading up to the election) were 27 and 21, respectively. So in the five weekdays prior to the election, both the mean and median number of fake news articles increased by approximately 17 articles each day. Our estimates then imply that, in each of these five weekdays: (a) the incidence of visiting fake news sites increased by 5.0% (using the estimate from column (3) of Table 3 and the stated medians); and (b) the odds of visiting at least one fake news site increased by 6.2% (using the estimate from column (2) of Table 7 and the stated medians). This statistical exercise illustrates the incentives of fake news producers to write and disseminate articles.

4.8 Limitation (to Desktop Behavior) of Firefox Browsing Data

One downside of the Firefox Context Graph project is that it applies only to users browsing the Internet via the desktop version of the Firefox browser. An obvious concern is that considerable Internet browsing and social media activity occurs using the mobile Web and smartphone applications. If there are any systematic differences between desktop and mobile browsing, then our results may not capture the true effect of fake news supply on consumption. We evaluate this possibility by obtaining data from Pocket, a service that allows users to save (or "pocket") websites that are of interest to them. As of February 2017, this service had more than 10 million active users. It is important for this discussion that Pocket has a large mobile user base. At the time of our data collection effort with this company, it stated that almost half of saved websites were pocketed via a mobile device and that nearly three fourths of Pocket users later viewed these saved pages on a mobile device. The data we acquired from Pocket were daily counts of the number of fake and credible news websites that were pocketed each day (i.e., we did not acquire individual-level data).

To compare user data from the two services, we first aggregate our individual-level, Firefox-based fake and credible news site visits to obtain average daily visits for each. We then normalize our measures in units of standard deviations so that we can make comparisons across the two services. Next, we look at correlations and visual evidence of a relationship between the two sets of two measures. The normalized Firefox-based *fake* news site visits per person and the corresponding normalized Pocket-based visits have a correlation coefficient of 0.41. The relationship is plotted in Figure 6 and shows a general consistency in consumer behavior across the two services. Similarly, the correlation coefficient is 0.69 for the normalized Firefox-based *credible* news site visits per person and the corresponding normalized credible Pocket-based visits. The graph presented as Figure 7 likewise illustrates that users' behavior patterns are consistent across the two services.

This comparison of these general patterns leads us to believe that Firefox desktop browsing behavior is a useful proxy for overall news site visiting behavior that includes mobile consumption—that is, because the former reflects trends observed in the latter.

5 Conclusion

We find a link between the production and consumption of fake news during the 2016 US presidential election period. Not only did consumers increase their consumption of such news, they were also more likely to visit at least one fake news site. From an economic perspective, our findings imply that the production of fake news would, on average, yield traffic rates similar to those generated by traditional display advertising. Our findings support the view that producers of fake news were effective in attracting a diverse set of viewers. As policymakers and private firms wrestle with how (and whether) to manage the diffusion of unverifiable political content, our estimates provide some evidence that restricting supply reduces consumption in predictable ways.

Although the promulgation of fake news is widely viewed as potentially corrosive to democratic processes, we cannot say for certain that opinions of users were changed by their consumption of fake news. It is certainly possible that users are able to recognize misinformation and will reject it. Yet it is clear that the efforts of many fake news producers were effective at altering the US electorate's diet of news and information.

It has long been the case that embellished or exaggerated news stories appear during national elections, and the 2016 US presidential election was no exception. What may have been different in this cycle, however, is that many users now rely on social media or other non-mainstream sites to access news. And if fake news—with its low production and distribution costs—is a form of "cheap talk", then our results identify the economic incentives that persuaded many agents to produce it. Simply put, our findings suggest that producers of verifiably false news were effective at directing users to similar sites.

References

Abramowitz, A. I. and K. L. Saunders (2006). Exploring the bases of partisanship in the american electorate: Social identity vs. ideology. *Political Research Quarterly 59*(2), 175–187.

Allcott, H. and M. Gentzkow (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives 31*(2), 211–236.

Athey, S. and G. Imbens (2017). The state of applied econometrics-causality and policy evaluation. *Journal of Economic Perspectives 31*(2), 3–32.

Bakshy, E., S. Messing, and L. A. Adamic (2015). Exposure to ideologically diverse news and opinion on facebook. *Science 348*(6239), 1130–1132.

Boczkowski, P. J. (2010). *News at Work: Imitation in an Age of Information Abundance*. University of Chicago Press.

Boczkowski, P. J. and E. Mitchelstein (2013). *The News Gap: When the Information Preferences of the Media and the Public Diverge*. Cambridge: The MIT Press.

Bushman, B. J. and C. A. Anderson (2001). Media violence and the american public: Scientific facts versus media misinformation. *American Psychologist 56*(6-7), 477.

Cagé, J., N. Hervé, and M.-L. Viaud (2016). The production of information in an online world. *Working Paper*.

Chiang, C.-F. and B. Knight (2011). Media bias and influence: Evidence from newspaper endorsements. *The Review of Economic Studies 78*(3), 795–820.

Del Vicario, M., A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences 113*(3), 554–559.

DellaVigna, S. and M. Gentzkow (2010). Persuasion: empirical evidence. *Annu. Rev. Econ. 2*(1), 643–669.

DellaVigna, S. and E. Kaplan (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics 122*(3), 1187–1234.

Dilliplane, S. (2011). All the news you want to hear: The impact of partisan news exposure on political participation. *Public Opinion Quarterly 75*(2), 287–316.

Enikolopov, R., M. Petrova, and E. Zhuravskaya (2011, 12). Media and political persuasion: Evidence from russia. *American Economic Review 101*(7), 3253–3285.

Erisen, C., M. Lodge, and C. S. Taber (2014). Affective contagion in effortful political thinking. *Political Psychology 35*(2), 187–206.

Fourney, A., M. Z. Racz, G. Ranade, M. Mobius, and E. Horvitz (2017). Geographic and temporal trends in fake news consumption during the 2016 us presidential election.

Friedman, J., T. Hastie, and R. Tibshirani (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software 33*(1), 1.

Friesen, J. P., T. H. Campbell, and A. C. Kay (2015). The psychological advantage of unfalsifiability: the appeal of untestable religious and political ideologies. *Journal of personality and social psychology 108*(3), 515.

Gabielkov, M., A. Ramachandran, A. Chaintreau, and A. Legout (2016, 6). Social clicks: What and who gets read on twitter? In *ACM SIGMETRICS / IFIP Performance 2016*.

Gentzkow, M. and J. M. Shapiro (2006). Media bias and reputation. *Journal of political Economy 114*(2), 280–316.

Gerber, A. S., D. Karlan, and D. Bergan (2009). Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics 1*(2), 35–52.

Gilchrist, D. S. and E. G. Sands (2015). Something to talk about: Social spillovers in movie consumption. *Journal of Political Economy 1*.

Gottfried, J., M. Barthel, E. Shearer, and A. Mitchell (2016). The 2016 presidential campaign – a news event that's hard to miss.

Gottfried, J. and E. Shearer (2016). News use across social media platforms 2016. *Pew Research Center 26*.

Greene, S. (2004). Social identity theory and party identification. *Social Science Quarterly 85*(1), 136–153.

Greenwood, S., A. Perrin, and M. Duggan (2016). Social media update 2016. pew research center.

Guess, A., B. Nyhan, and J. Reifler (2017). Selective exposure to misinformation: evidence from the consumption of fake news during the 2016 u.s. presidential campaign.

Guzman, J. and S. Stern (2015). Innovation economics. where is silicon valley? *Science 347*(6222), 606–9.

Hamilton, J. (2004). *All the news that's fit to sell: How the market transforms information into news*. Princeton University Press.

Huddy, L. (2001). From social to political identity: A critical examination of social identity theory. *Political psychology 22*(1), 127–156.

Huddy, L., L. Mason, and L. Aarøe (2015). Expressive partisanship: Campaign involvement, political emotion, and partisan identity. *American Political Science Review 109*(1), 1–17.

Knobloch-Westerwick, S. (2014). *Choice and preference in media use: Advances in selective exposure theory and research.* Routledge.

Kumar, S., R. West, and J. Leskovec (2016). Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *Proceedings of the 25th International World Wide Web Conference.*

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin 108*(3), 480.

Lelkes, Y., G. Sood, and S. Iyengar (2017). The hostile audience: The effect of access to broadband internet on partisan affect. *American Journal of Political Science 61*(1), 5–20.

Levendusky, M. (2009). *The partisan sort: How liberals became Democrats and conservatives became Republicans.* University of Chicago Press.

Lewandowsky, S., U. K. Ecker, C. M. Seifert, N. Schwarz, and J. Cook (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest 13*(3), 106–131.

Lord, C. G., L. Ross, and M. R. Lepper (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology 37*(11), 2098.

Messing, S. and S. J. Westwood (2014). Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communication Research 41*(8), 1042–1063.

Mullainathan, S. and A. Shleifer (2005). The market for news. *The American Economic Review 95*(4), 1031–1053.

Redlawsk, D. P. (2002). Hot cognition or cool consideration? testing the effects of motivated reasoning on political decision making. *Journal of Politics 64*(4), 1021–1044.

Sears, D. O. and J. L. Freedman (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly 31*(2), 194–213.

Shane, S. and M. Isaac (2017). Facebook to turn over russia-linked ads to congress. *The New York Times*.

Silverman, C. (2017). This is how your hyperpartisan political news gets made. *BuzzFeed News*.

Silverman, C. and L. Alexandar (2016). How teens in the balkans are duping trump supporters with fake news. *BuzzFeed News*.

Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior 30*(3), 341–366.

Subramanian, S. (2017). Fake news factory: Inside the macedonian fake news complex. *Wired*.

Taber, C. S. and M. Lodge (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science 50*(3), 755–769.

Volovich, K. (2016). What's a good clickthrough rate? new benchmark data for google adwords.

Vosoughi, S., D. Roy, and S. Aral (2018). The spread of true and false news online. *Science 359*, 1146–1151.

Weeks, B. E. and R. K. Garrett (2014). Electoral consequences of political rumors: Motivated reasoning, candidate rumors, and vote choice during the 2008 us presidential election. *International Journal of Public Opinion Research 26*(4), 401–422.

Zimdars, M. (2016). False, misleading, clickbait-y, and satirical 'news' sources.

Table 1: Individual-Level Summary Statistics

	Mean	S.D.	Median	Min	Max
Panel A: Full Context Graph Sample					
fake site visits	0.69	4.01	0.00	0.00	187.00
fake site (0/1)	0.13	0.33	0.00	0.00	1.00
news site visits	1.08	3.98	0.00	0.00	139.00
other site visits	65.22	103.32	36.00	0.00	3,139.00
social media visits	4.18	31.79	0.00	0.00	1,478.00
Panel B: Sample for FE Models					
fake site visits	1.15	5.11	0.00	0.00	187.00
fake site (0/1)	0.21	0.41	0.00	0.00	1.00
news site visits	1.50	4.85	0.00	0.00	139.00
other site visits	72.69	96.06	43.00	0.00	2,290.00
social media visits	5.97	40.61	1.00	0.00	1,478.00
Panel C: No Fake News Consumers					
fake site visits	0.00	0.00	0.00	0.00	0.00
fake site (0/1)	0.00	0.00	0.00	0.00	0.00
news site visits	0.43	1.88	0.00	0.00	37.00
other site visits	53.86	112.64	27.00	0.00	3,139.00
social media visits	1.43	5.19	0.00	0.00	130.00

Note: Panel A includes $n = 32{,}020$ user-days for 1,418 users, and Panel B includes $n = 19{,}331$ user-days for 766 users. Panel C includes $n = 12{,}661$ user-days for 624 users; it excludes 28 users who appear in the data for only one day over the entire study period.

Table 2: Daily-Level Summary Statistics

	Mean	S.D.	Median	Min	Max
FN supply	1.16	1.56	0.60	0.00	8.00
DJT online mentions	32.96	9.70	31.51	15.87	62.03
HRC online mentions	27.08	9.11	27.04	13.49	51.66
change in pct DJT mentions	0.24	7.24	0.49	-19.84	17.76
change in pct immigration mentions	6.89	37.20	8.03	-60.14	161.86
change in pct economy mentions	6.83	37.07	3.33	-60.00	121.21
Trump tweet count	15.08	16.27	10.00	2.00	87.00
poll spread	4.57	1.83	5.10	1.50	7.10
DJT TV mentions	9.84	4.02	9.69	0.57	19.45
HRC TV mentions	6.26	2.86	6.26	0.25	13.51
predicted FN supply	0.24	0.40	0.12	0.00	2.27

Note: $n = 49$ days. The FN supply measure scales the article counts by a factor of 10. Count measures of online and television mentions for both candidates (i.e., Trump and Clinton online mentions as well as Trump and Clinton TV mentions) are scaled by 1,000.

Table 3: Effect of Fake News Supply on Visits to Fake News Sites

DV: fake site visits	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
FN supply	0.042	0.027	0.029	0.028	0.027	0.024	0.034	0.030
	(0.016)	(0.016)	(0.014)	(0.014)	(0.018)	(0.013)	(0.014)	(0.013)
social media visits	0.002	-0.003	0.021	0.021	0.021	0.021	0.021	0.021
	(0.001)	(0.001)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
DJT online mentions	-0.025	-0.024	-0.024	-0.027	-0.025	-0.025	-0.029	
	(0.015)	(0.016)	(0.017)	(0.016)	(0.018)	(0.016)	(0.016)	
HRC online mentions	0.029	0.028	0.025	0.031	0.026	0.019	0.031	
	(0.016)	(0.017)	(0.018)	(0.017)	(0.019)	(0.021)	(0.017)	
change in pct DJT mentions				0.005				
				(0.003)				
change in pct immigration mentions					-0.001			
					(0.001)			
change in pct economy mentions					0.000			
					(0.001)			
Trump tweet count						0.004		
						(0.003)		
poll spread							-0.055	
							(0.061)	
DJT TV mentions								-0.013
								(0.027)
HRC TV mentions								0.010
								(0.041)
Weekend Interactions (weekend ×)								
FN supply	-0.005	0.013	-0.012	0.001	-0.040	-0.021	-0.027	-0.063
	(0.042)	(0.042)	(0.042)	(0.043)	(0.053)	(0.043)	(0.043)	(0.044)
social media visits	0.002	-0.002	0.017	0.017	0.016	0.017	0.017	0.017
	(0.001)	(0.002)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
DJT online mentions	-0.007	-0.003	0.001	-0.001	0.004	0.005	-0.008	
	(0.015)	(0.015)	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	
HRC online mentions	0.054	0.041	0.037	0.032	0.025	0.023	0.037	
	(0.021)	(0.020)	(0.021)	(0.023)	(0.027)	(0.025)	(0.025)	
change in pct DJT mentions				0.004				
				(0.010)				
change in pct immigration mentions					0.002			
					(0.001)			
change in pct economy mentions					-0.001			
					(0.003)			
Trump tweet count						-0.001		
						(0.013)		
poll spread							-0.044	
							(0.045)	
DJT TV mentions								0.032
								(0.026)
HRC TV mentions								0.046
								(0.035)
Constant	-2.241	0.552						
	(0.631)	(0.464)						
Calendar week dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day of week dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zero Inflated	No	Yes	No	No	No	No	No	No
User FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Obs	32020	32020	19325	19325	19325	19325	19325	19325
Users	1418	1418	765	765	765	765	765	765
Log likelihood	-67549.5	-34029.8	-18949.5	-18936.3	-18940.7	-18937.6	-18937.5	-18959.1

Note: Each column presents results from a Poisson model with robust standard errors (in parentheses) clustered at the user level. The dependent variable is a count of the number of fake news websites that user i visited on day t.

Table 4: Placebo and Alternative Dependent Variable Tests: Effect of Fake News Supply on Visits to Fake News Sites

DV:	(1) fake site visits	(2) news site visits	(3) other site visits
FN supply (fwd)	0.001		
	(0.014)		
FN supply		0.011	0.002
		(0.011)	(0.005)
social media visits	0.022	0.030	0.022
	(0.007)	(0.003)	(0.003)
DJT online mentions	-0.030	0.015	0.003
	(0.015)	(0.008)	(0.004)
HRC online mentions	0.035	-0.015	-0.005
	(0.016)	(0.008)	(0.004)
Weekend Interactions (weekend $\times$)			
FN supply (fwd)	-0.031		
	(0.051)		
FN supply		-0.100	-0.028
		(0.056)	(0.021)
social media visits	0.017	0.007	0.011
	(0.005)	(0.004)	(0.004)
DJT online mentions	-0.006	0.004	-0.010
	(0.012)	(0.013)	(0.007)
HRC online mentions	0.015	-0.009	0.005
	(0.023)	(0.019)	(0.010)
Calendar week dummies	Yes	Yes	Yes
Day of week dummies	Yes	Yes	Yes
User FE	Yes	Yes	Yes
Obs	18141	17172	19325
Users	726	659	765
Log likelihood	-17831.6	-23285.2	-457656.6

Note: Each column presents results from a fixed-effects Poisson model with robust standard errors (in parentheses) clustered at the user level. The dependent variables in each column are a count of the number of fake news websites, credible news websites, and all other websites (excluding any news sites and social sites) that user i visited on day t.

Table 5: Idiosyncratic Exposure to Supply of Fake News as Exogenous Variation (and Additional Tests)

DV:	(1) fake site visits	(2) fake site visits	(3) news site visits	(4) other site visits
FN supply × social visits	0.001		0.000	0.000
	(0.000)		(0.001)	(0.001)
FN supply (fwd) × social visits		0.000		
		(0.000)		
social media visits	0.021	0.022	0.030	0.022
	(0.007)	(0.007)	(0.003)	(0.003)
Weekend Interactions (weekend ×)				
FN supply × social visits	-0.002		-0.007	0.001
	(0.004)		(0.003)	(0.002)
FN supply (fwd) × social visits		0.001		
		(0.001)		
social media visits	0.018	0.017	0.016	0.010
	(0.007)	(0.005)	(0.005)	(0.005)
Calendar week dummies	Yes	Yes	Yes	Yes
Day of week dummies	Yes	Yes	Yes	Yes
User FE	Yes	Yes	Yes	Yes
Obs	19325	18141	17172	19325
Users	765	726	659	765
Log likelihood	-18966.2	-17853.8	-23273.6	-457899.4

Note: Each column presents results from a fixed-effects Poisson model with robust standard errors (in parentheses) clustered at the user level. The dependent variable in columns (1) and (2) is a count of the number of fake news websites that user i visited on day t; in columns (3) and (4) the dependent variables are, respectively, counts of credible news websites visited and of all other websites (excluding any news site and social sites) visited by user i on day t.

Table 6: Predicted Supply of Fake News as Exogenous Variation (and Additional Tests)

DV:	(1) fake site visits	(2) fake site visits	(3) news site visits	(4) other site visits
predicted FN supply	0.096		0.043	0.002
	(0.049)		(0.046)	(0.020)
predicted FN supply (fwd)		-0.001		
		(0.062)		
social media visits	0.021	0.022	0.030	0.022
	(0.007)	(0.007)	(0.003)	(0.003)
DJT online mentions	-0.022	-0.030	0.015	0.003
	(0.017)	(0.015)	(0.008)	(0.004)
HRC online mentions	0.023	0.034	-0.014	-0.005
	(0.019)	(0.017)	(0.008)	(0.004)
Weekend Interactions (weekend $\times$)				
predicted FN supply	-0.162		-0.318	-0.074
	(0.139)		(0.204)	(0.062)
predicted FN supply (fwd)		-0.191		
		(0.308)		
social media visits	0.017	0.017	0.007	0.011
	(0.005)	(0.005)	(0.004)	(0.004)
DJT online mentions	0.001	-0.009	0.004	-0.011
	(0.013)	(0.012)	(0.013)	(0.007)
HRC online mentions	0.028	0.019	-0.005	0.006
	(0.021)	(0.022)	(0.016)	(0.010)
Calendar week dummies	Yes	Yes	Yes	Yes
Day of week dummies	Yes	Yes	Yes	Yes
User FE	Yes	Yes	Yes	Yes
Obs	19325	18141	17172	19325
Users	765	726	659	765
Log likelihood	-18951.7	-17831.2	-23286.4	-457673.1

Note: Each column presents results from a fixed-effects Poisson model with robust standard errors (in parentheses) clustered at the user level. The dependent variable in columns (1) and (2) is a count of the number of fake news websites that user i visited on day t; in columns (3) and (4) the dependent variables are, respectively, counts of credible news websites visited and of all other websites (excluding any news site and social sites) visited by user i on day t.

Table 7: Effect of Fake News Supply on Likelihood of Visiting a Fake News Site (and Additional Tests)

	(1)	(2)	(3)	(4)
DV:	fake site (0/1)	fake site (0/1)	fake site (0/1)	news site (0/1)
FN supply	0.025	0.036		0.008
	(0.012)	(0.021)		(0.015)
FN supply (fwd)			0.001	
			(0.016)	
social media visits	0.003	0.013	0.012	0.069
	(0.003)	(0.021)	(0.026)	(0.009)
DJT online mentions	-0.006	-0.015	-0.008	0.001
	(0.009)	(0.014)	(0.016)	(0.015)
HRC online mentions	0.008	0.017	0.005	-0.000
	(0.010)	(0.015)	(0.018)	(0.016)
Weekend Interactions (weekend $\times$)				
FN supply	-0.017	-0.048		-0.102
	(0.032)	(0.046)		(0.046)
FN supply (fwd)			-0.076	
			(0.055)	
social media visits	0.008	0.012	0.012	0.036
	(0.022)	(0.010)	(0.015)	(0.014)
DJT online mentions	-0.003	0.003	-0.004	0.028
	(0.011)	(0.019)	(0.017)	(0.016)
HRC online mentions	0.023	0.025	0.027	-0.023
	(0.017)	(0.030)	(0.023)	(0.021)
Constant	-2.854			
	(0.715)			
Calendar week dummies	Yes	Yes	Yes	Yes
Day of week dummies	Yes	Yes	Yes	Yes
User FE	No	Yes	Yes	Yes
Obs	32020	18929	17773	16721
Users	1418	740	704	631
Log likelihood	-12062.6	-5479.5	-5163.6	-6341.1
Pseudo R-sqaured	0.0090	0.021	0.020	0.052

Note: Each column presents results from a logistic model. Clustered SEs (at the user level in column (1)) and bootstrapped SEs (in columns (2)–(4)) are reported in parentheses. The dependent variable in columns (1)–(3) is a binary variable set to 1 if user i visited at least one fake news website on day t (and set to 0 otherwise); the dependent variable in column (4) is a binary variable set to 1 only if at least one credible news website was visited on that day.

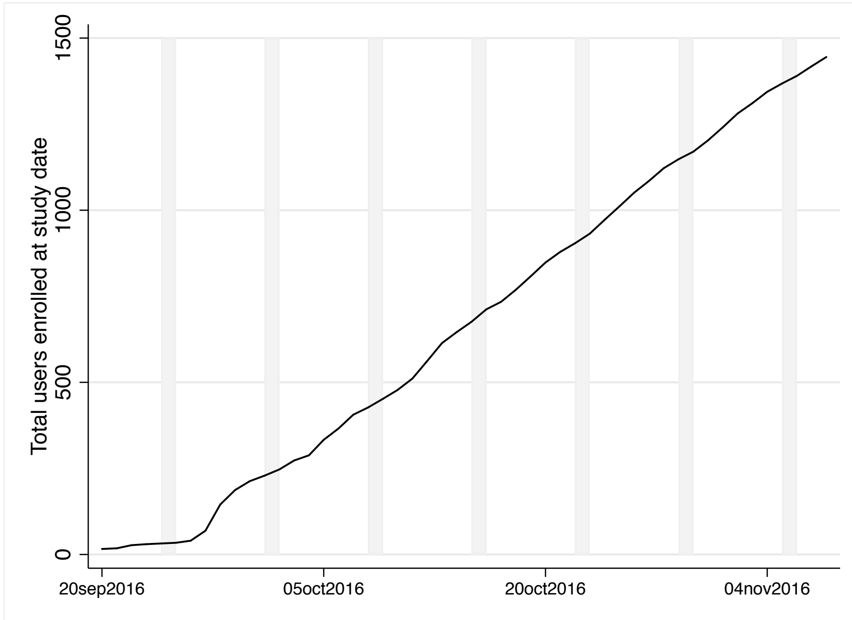

Figure 1: Cumulative Users by Day of Study

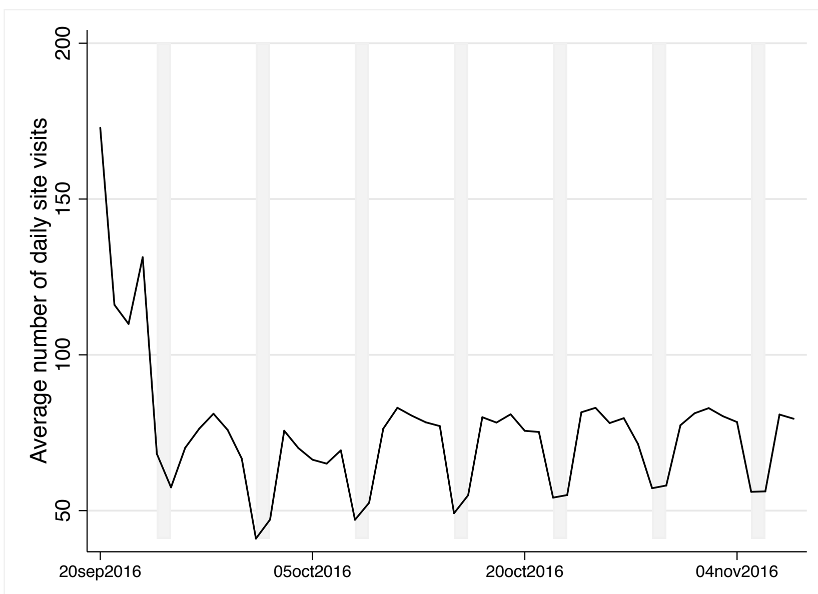

Figure 2: Average Unique Site Visits

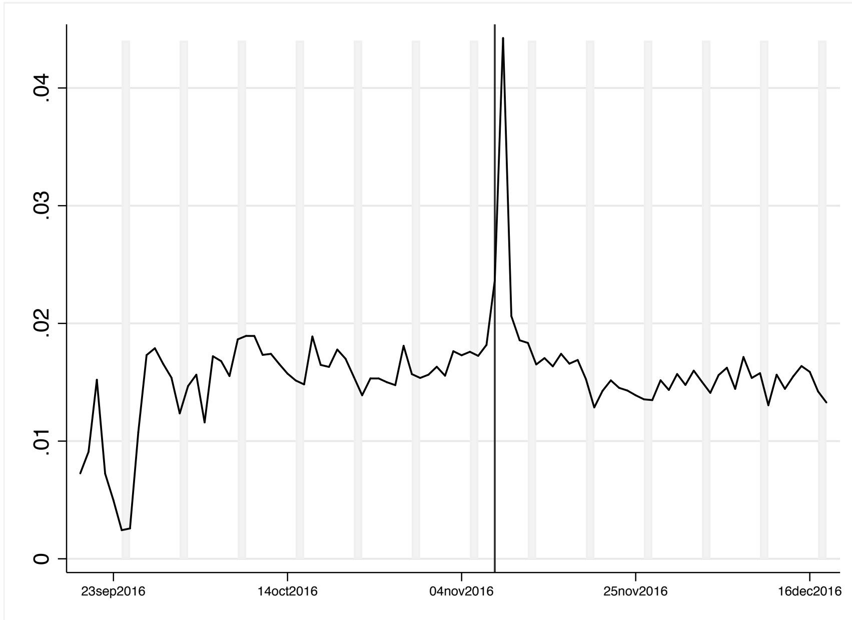

Figure 3: Trend in Consumption of Credible News

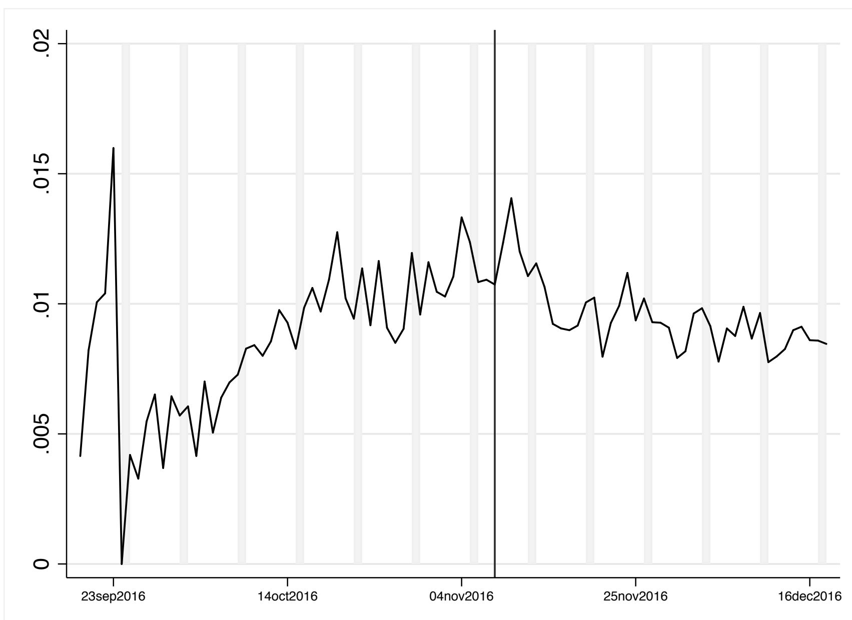

Figure 4: Trend in Consumption of Fake News

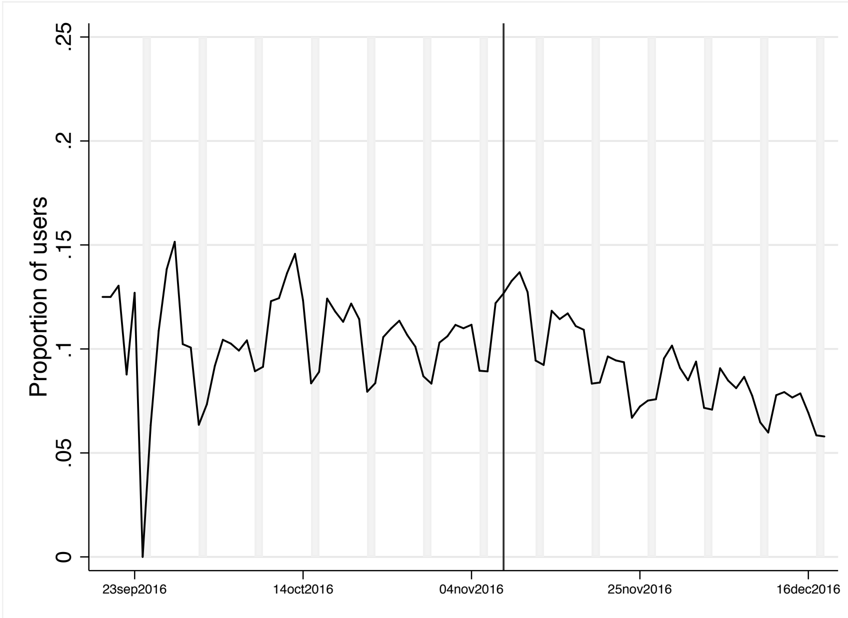

Figure 5: Daily Proportion of Users Who Visited at Least One Fake News Site

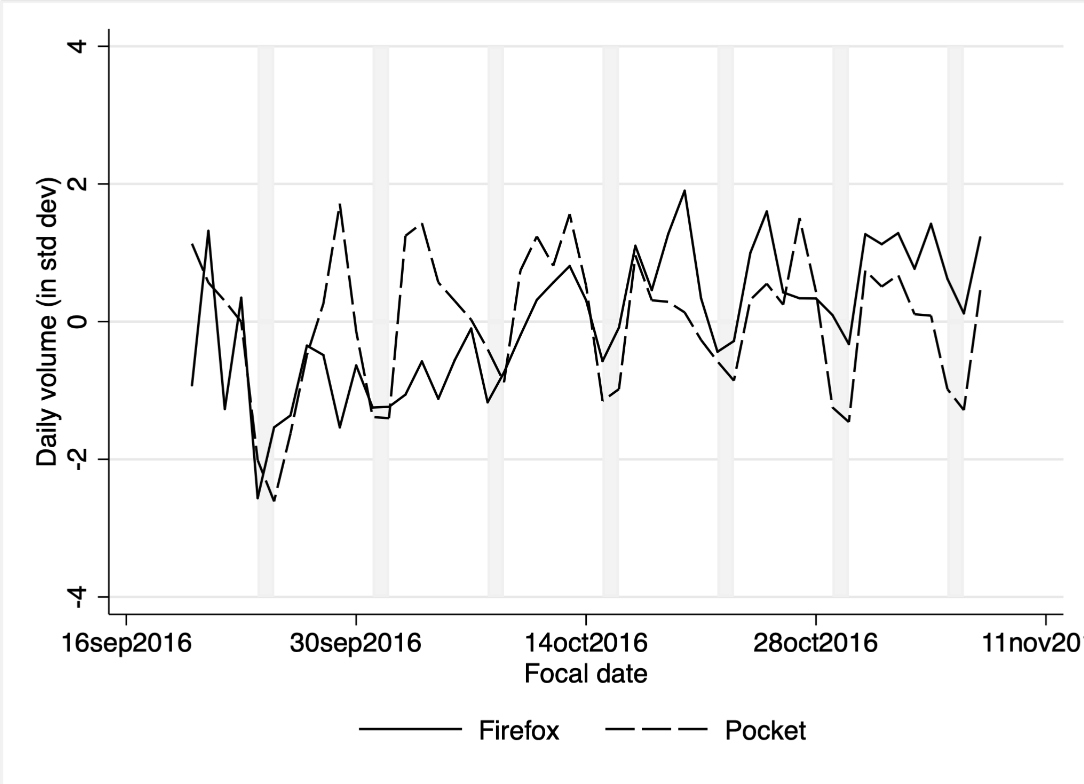

Figure 6: Firefox-Based versus Pocket-Based Visits to Fake News Sites

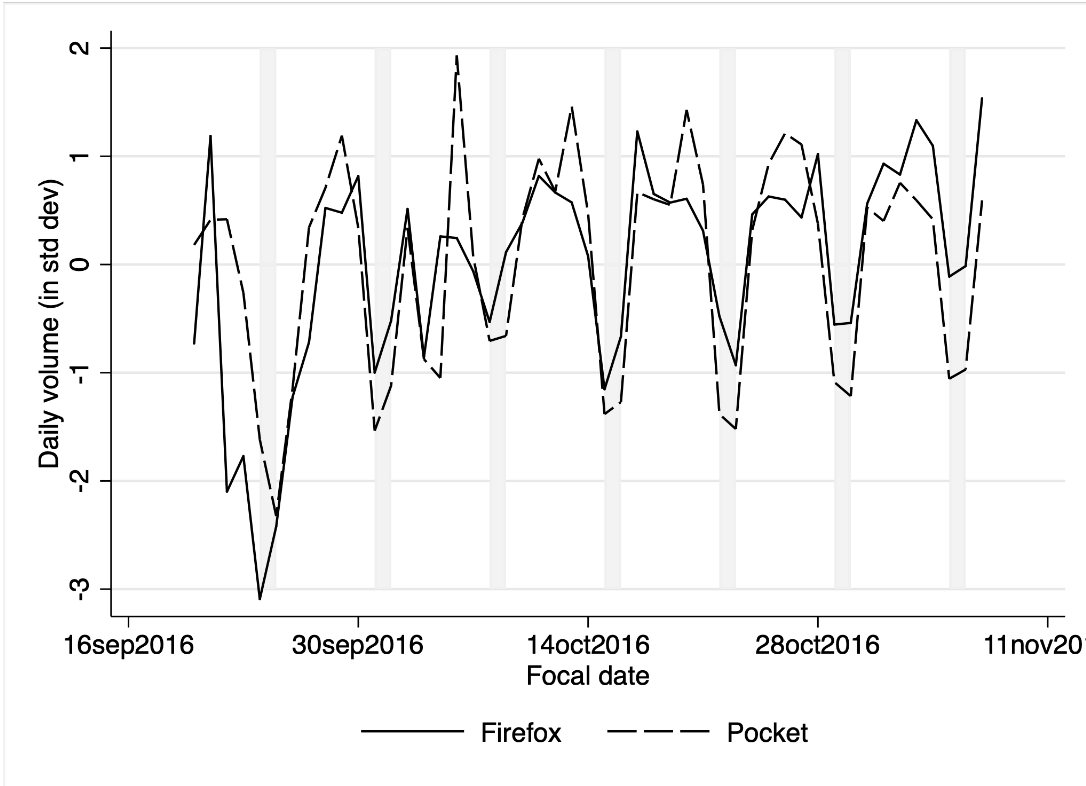

Figure 7: Firefox-Based versus Pocket-Based Visits to Credible News Sites

Data Appendix for "How the Supply of Fake News Affected Consumer Behavior during the 2016 US Election"

Anil R. Doshi
UCL School of Management

Sharat Raghavan
Haas School of Business, University of California at Berkeley

Rebecca Weiss
Mozilla Corporation and Stanford University

Eric Petitt
Mozilla Corporation

October 2018

A Introduction

This data appendix provides additional details on the collection of data and on the creation of variables for the supply of fake news and visits to fake news sites.

A.1 Site Visit Measures

Our site visit measures are obtained from Mozilla's Context Graph project. Our general appraoch was to provide Mozilla with a list of domains for each of our site visit measures—*fake site visits*, *news site visits*, and *social media visits*—and, in return, we received an aggregated daily count of visits to those domains by each user in the Context Graph. We also received a measure of the total daily sites visited by each user, which we used to compute our *other site visits* measure.

A.1.1 Fake Site Visits

Our fake site visits measure is derived by totaling the number of visits each user made to domains that are in the OpenSources[1] and Allcott and Gentzkow (2017; hereafter, AG) databases. From OpenSources, we acquired a list of 610 domains; those domains are provided in Table A1.

We then requested from Mozilla site visit counts to the 375 domains from the AG database. Of those domains, 130 were also in the initial OpenSources domains.

[1] The database can be found at http://www.opensources.co. The collection date of the URLs was January 26, 2017.

An additional 15 domains that served a fake news article based on AG are generally regarded as being legitimate news sources or general content platforms and were thus excluded. Those domains include: bloomberg.com, buzzfeed.com, dailymail.co.uk, getpocket.com, huffingtonpost.com, huffingtonpost.co.uk, independent.co.uk, nydailynews.com, nymag.com, nypost.com, people.com, slate.com, talkingpointsmemo.com, washingtontimes.com, and youtube.com. Finally, the data returned by Mozilla included 16 domains that were not matched with our request (for example, we requested 1776coalition.com and we received the site visit counts for coalition.com). Those 16 domains were: 1776coalitioncom, 24usainfocom, 710woriheartcom, chuckcallestoblogspotcom, electionfraud2016wordpresscom, eninstitutomanquehueorg, friendsofsyriawordpresscom, nuevoordenmundialreptilianoblogspotcom, u281p372newsninjaacom, u4281p6798newzfeednet, u4638p2660newsninjaacom, u7176p8678newzfeednet, u7434p6564mrsjekyllsaysnet, u7690p7141liamtheleprechaunco, and u8177p8978newzfeednet. After allowing for duplicative, legitimate and erroneously returned domains, we added the site visit counts to the remaining 214 domains, which are listed in Table A2.

Using the combined lists described above, we created our *fake site visits* measure based on a total of 824 domains. The data we acquire is a daily total number of visits to any of those 824 domains by user.

A.1.2 Credible News Site Visits

To obtain a list of domains for our measure of visits to credible news sites, we again started with the OpenSources database. When these data were collected, this database maintained a list of fake news sources and also a list of credible news sources.[2] We supplemented the Opensources sites with a hand-collected list of major online regional newspapers. Altogether, there are 69 domains on our list of credible news sites and they are listed in Table A3.

A.1.3 Social Media Site Visits

Our measure of social media visits consists of a count of daily visits made to the following six sites: twitter.com, facebook.com, instagram.com, snapchat.com, linkedin.com, and pinterest.com.

A.2 Fake News Supply

Our fake news supply measure is the article data set from Allcott and Gentzkow (2017). Their data consist of 948 articles (distinct URLs) published on 156 news topics and appearing on 375 websites (distinct domains) from May 15, 2016, to November 11 , 2016. The articles are about the presidential campaign or candidates, and they are cited as being false by the fact-checking services Snopes, Politifact, and/or Buzzfeed.

[2]The list of credible news sources is no longer maintained by OpenSources. However, it can be viewed by inspecting changes made to the database on April 2, 2017 at GitHub .

For our primary set of analyses, we count the daily number of articles in the database as our measure of $FNSupply_t$ over the seven weeks prior to the election. Table A4 lists all the sample domains that hosted at least one fake news article during the study period.

A.2.1 Machine Learning Prediction of Fake News Supply Used in Section 4.5.2

In the main text section entitled "Identifying Exogenous Producers of Fake News", we limit the count of articles in $FNSupply_t$ to include only those appearing on sites that were predicted to be taken down by June 2017. To construct this measure, we first visited each of the domains and assessed whether the site was completely shut down (the "strict" definition) or whether it was technically still being hosted but essentially dormant (the "loose" definition). For each definition, we document these observations in the respective "Site Down" columns of Table A4.

To create our ex ante prediction of whether the site would eventually be taken down, we first assembled a set of features from the domains' ICANN registration data (obtained with the assistance of DomainTools, an online security company). A list of those features is given in Table A5. We employ the LASSO algorithm, developed in the machine learning field (Friedman et al., 2010), to select variables that best predict the outcome of interest (here, whether the site was shut down). We employ a 10-fold cross-validation technique to estimate the coefficients; this involves using 90% of the sample as "training" data and witholding 10% of the data for out-of-sample testing. That process is repeated over 10 iterations, and the coefficients are averaged across the iterations. The resulting coefficient estimates (and thus the features that were included) for predicting the "site down" outcome are also presented—for both the strict and loose definition—in Table A5.

References

Allcott, H. and Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–236.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1.

Table A1: List of Domains from OpenSources Database Included in *fake site visits*

100percentfedup.com, 21stcenturywire.com, 24newsflash.com, 365usanews.com, 4threvolutionarywar.wordpress.com, 70news.wordpress.com, 82.221.129.208, aanirfan.blogspot.co.uk, abcnews.com.co, abcnewsgo.co, abeldanger.net, abovetopsecret.com, abriluno.com, aceflashman.wordpress.com, acting-man.com, activistpost.com, addictinginfo.org, adobochronicles.com, advocate.com, ahtribune.com, allnewspipeline.com, americanfreepress.net, americankabuki.blogspot.com, americanlookout.com, americannews.com, americanoverlook.com, americanpatriotdaily.com, americanreviewer.com, americantoday.news, americasfreedomfighters.com, ammoland.com, amplifyingglass.com, amren.com, amtvmedia.com, amusmentic.com, ancient-code.com, angrypatriotmovement.com, anonews.co, anonhq.com, anonnews.co, anotherdayintheempire.com, antiwar.com, antoniusaquinas.wp.com, asia-pacificresearch.com, assassinationscience.com, associatedmediacoverage.com, attn.com, automaticearth.com, automotostar.com, awarenessact.com, awdnews.com, awm.com, barenakedislam.com, bb4sp.com, beehivebugle.com, beforeitsnews.com, betootaadvocate.com, bients.com, bigamericannews.com, bigbluedimension.com, bigbluevision.com, bigbluevision.org, bighairynews.com, bignuggetnews.com, bigpzone.com, bipartisanreport.com, blackagendareport.com, blacklistednews.com, bluenationreview.com, boilingfrogspost.com, borowitzreport.com, bostonleader.com, breaking911.com, breitbart.com, brotherjohnf.com, bullionbullscanada.com, burrardstreetjournal.com, buzzfeedusa.com, bvanews.com, callthecops.net, canadafreepress.com, cap-news.com, cbsnews.com.co, celebtricity.com, channel-7-news.com/, chaser.com.au, checkoutthehealthyworld.com, chicksontheright.com, christianfightback.com, christiantimesnewspaper.com, christwire.org, chronicle.su, cityworldnews.com, civictribune.com, clashdaily.com, clickhole.com, cnnnext.com, cnsnews.com, coasttocoastam.com, collective-evolution.com, collectivelyconscious.net, commondreams.org, concisepolitics.com, consciouslifenews.com, conservativebyte.com, conservativedailypost.com, conservativefiringline.com, conservativefrontline.com, conservativeinfidel.com, conservativeoutfitters.com, conservativerefocus.com, conservativespirit.com, conservativestate.com, conservativetribune.com, consortiumnews.com, conspiracywire.com, corbettreport.com, countdowntozerotime.com, countercurrents.org, counterinformation.wordpress.com, counterpsyops.com, counterpunch.com, counterpunch.org, creambmp.com, crystalair.com, dailybuzzlive.com, dailycurrant.com, dailydiscord.com, dailyheadlines.com, dailyheadlines.net, dailykos.com, dailyleak.org, dailynewsbin.com, dailynewspolitic.com, dailyoccupation.com, dailypolitics.info, dailypoliticsusa.com, dailysignal.com, dailysquib.co.uk, dailystormer.com, dailywire.com, dandygoat.com, darkmoon.me, darkpolitricks.com, davejanda.com, davidduke.com, davidstockmanscontracorner.com, davidwolfe.com, dcclothesline.com, dcgazette.com, dcleaks.com, deadlyclear.wordpress.com, defenddemocracy.press, delectabledietofpics.net, dennismichaellynch.com, denverguardian.com, departed.co, derfmagazine.com, dineal.com, disclose.tv, disclosuremedia.net, dissentmagazine.org, diversitychronicle.wordpress.com, dollarvigilante.com, donaldtrumpnews.co, dont-tread-on.me, downtrend.com, drudgereport.com, drudgereport.com.co, duffleblog.com, duhprogressive.com, dutchsinse.com, eaglerising.com, ebolahoax.com, educate-yourself.org, educateinspirechange.org/health, electionnightgatekeepers.com, elelephantintheroom.blogspot.com, elitereaders.com, elkoshary.com, elmundotoday.com, embols.com, empireherald.com, empirenews.net, empiresports.co, emptywheel.net, enabon.com, endingthefed.com, endoftheamericandream.com, endtime.com, enduringvision.com, english.ruvr.ru, eutimes.net, eutopia.buzz, everydayworldnews.com, everythingnewdaily.com, ewao.com, expose1933.com, extraclubmagazine.com, eyeopening.info, fakingnews.com, familysecuritymatters.org, fantasticword.com, federalistpress.com, fellowshipoftheminds.com, filmsforaction.org, financialsurvivalnetwork.com, floridasunpost.com, flyheight.com, fmobserver.com, fognews.ru, foodbabe.com, foreignpolicyjournal.com, fort-russ.com, fourwinds10.net, fprnradio.com, freakoutnation.com, freebeacon.com, freedomdaily.com, freedomforceinternational.com, freedomoutpost.com, freedomsphoenix.com, freepatriot.org, freewoodpost.com, fridaymash.com, fromthetrenchesworldreport.com, frontpagemag.com, fusion.net, gaia.com, galacticconnection.com, gangstergovernment.com, gatesofvienna.net, geoengineeringwatch.org, geopolmonitor.com, globalresearch.ca, glossynews.com, godlikeproductions.com, gomerblog.com, goneleft.com, gonzalolira.blogspot.com, gopthedailydose.com, govtslaves.info, greanvillepost.com, guardianlv.com, guccifer2.wordpress.com, gulagbound.com, hangthebankers.com, healthimpactnews.com, healthnutnews.com, heatst.com, henrymakow.com, heresyblog.net, holyobserver.com, humansarefree.com, humortimes.com, huzlers.com, ifyouonlynews.com, ihavethetruth.com, ijr.com, ilovemyfreedom.org, in5d.com, indiaarising.com, informationclearinghouse.info, informetoday.com, infostormer.com, infowars.com, instaworldnews.com, intellihub.com, intrendtoday.com, intrepidreport.com, investmentresearchdynamics.com, investmentwatchblog.com, ironictimes.com, islamicanews.com, itaglive.com, itmakessenseblog.com, iwanttoexplore.com, jackpineradicals.com, jacobinmag.com, jamesrgrangerjr.com, jesus-is-savior.com, jewsnews.co.il, johnnyrobish.com, jonesreport.com, journal-neo.org, katehon.com, katehon.org, landoverbaptist.org, legorafi.fr, lewrockwell.com, liberalamerica.org, liberalbias.com, liberaldarkness.com, libertyblitzkrieg.com, libertyfederation.com, libertymovementradio.com, libertynews.com, libertytalk.fm, libertyunyielding.com, libertyvideos.org, libertywritersnews.com, lifeandabout.com, lifenews.com, lifeprevention.com, lifesitenews.com, lifezette.com, liveactionnews.org, livefreelivenatural.com, livevote.com, lushforlife.com, madpatriots.com, madworldnews.com, magafeed.com, makeamericagreattoday.com, mediamass.net, mediazone.news, megafreshnews.com, megynkelly.us, militianews.com, mintpressnews.com, moonofalabama.org, morningnewsusa.com, mpidailymagazine.com, mrconservative.com, msnbc.website, mydailyrelaxation.com, myfreshnews.com, myzonetoday.com, nahadaily.com, nakedcapitalism.com, nationalreport.net, nationindistress.weebly.com, nationonenews.com, naturalblaze.com, naturalnews.com, nbc.com.co, nbcpolitics.org, nbcpoll.com, ncscooper.com, nevo.news, newcenturytimes.com, newcoldwar.org, news4ktla.com, newsbbc.net, newsbiscuit.com, newsbreakers.org, newsbuzzdaily.com, newscenterusa.com, newscorpse.com, newsexaminer.net, newsfrompolitics.com, newslo.com, newsmax.com, newsmutiny.com, newsninja2012.com, newsopening.com, newstarget.com, newsthump.com, newstoad.net, newswatch28.com, newswatch33.com, newswire-24.com, newswithviews.com, newyorker.com/humor, nodisinfo.com, nomorefakenews.com, northcrane.com, notallowedto.com, now8news.com, nowtheendbegins.com, nutritionfacts.org, nymeta.co, nyuzer.com, objectiveministries.org, occupydemocrats.com, occupyliberals.com, odgossip.com, off-guardian.org, oftwominds.com, oilgeopolitics.net, onlineconservativepress.com, opednews.com, openmindmagazine.com, orientalreview.org, other98.com, pakalertpress.com, pamelageller.com, patdollar.com, patriotchronicle.com, patriotnewsdaily.com, patriotrising.com, patriotupdate.com, paulcraigroberts.org, platosguns.com, politicalblindspot.com, politicalcult.com, politicalears.com, politicalo.com, politicalsitenews.com, politicaltimes.org, politicalupdator.com, politicops.com, politicsbreaking.com, politicsinformation.com, politicsinfotoday.com, politicsintheusa.com, politicsinusa.com, politicususa.com, powerpoliticians.com, pravda.ru, pravdareport.com, prepperwebsite.com, presidentialvoting2016.com, press24.us, presstv.com, presstv.ir, prisonplanet.com, prisonplanet.tv, prntly.com, projectveritas.com, proudcons.com, proudemocrat.com, qpolitical.com, randpaulreview.com, rawforbeauty.com, rawstory.com, rawws.com, rbth.com, react365.com, readconservatives.news, readynutrition.com, reagancoalition.com, realfarmacy.com, realnewsrightnow.com, realplanetnews.com, realprogress.online, realtimepolitics.com, redflagnews.com, redstate.com, redstatewatcher.com, reductress.com, regated.com, remedydaily.com, rense.com, responsibletechnology.org, returnofkings.com, revolutions2040.com, rhotv.com, rickwells.us, rightalert.com, righton.com, rightwingnews.com, rilenews.com, rinf.com, rockcitytimes.com, ronpaulinstitute.org, rumormillnews.com, ruptly.tv, russia-direct.org, russia-insider.com, satiratribune.com, satirewire.com, scrappleface.com, secretsofthefed.com, sensationalisttimes.com, sentinelblog.com, sheepkillers.com, shoebat.com, silver-coin-investor. com, silverbearcafe.com, silverdoctors.com, silverstealers.net, silverstrategies.com, sjlendman.blogspot.com, skeptiko.com, sonsoflibertyradio.com, sportspickle.com, stneotscitizen.com, stormcloudsgathering.com, stuppid.com, subjectpolitics.com, supremepatriot.com, surrealscoop.com, theamericanindependent.wordpress.com, thebeaverton.com, theblaze.com, thebostontribune.com, thecommonsenseshow.com, thecontroversialfiles.net, thedailybeast.com, thedailymash.co.uk, thedailysheeple.com, thedailywtf.com, theduran.com, theeconomiccollapseblog.com, theeventchronicle.com, theextinctionprotocol.com, thefederalistpapers.org, theforbiddenknowledge.com, thefreepatriot.org, thefreethoughtproject.com, thegatewaypundit.com, thegoldandoilguy.com/articles, thehardtimes.net, theineptowl.com, theinformedamerican.net, thelastgreatstand.com, thelibertybeacon.com, thelibertymill.com, themadisonmisnomer.com, themindunleashed.com, themindunleashed.org, themuslimissue.wordpress.com, thenewinquiry.com, thenewsnerd.com, theonion.com, thephaser.com, thepoke.co.uk, thepoliticalinsider.com, theracketreport.com, therealstrategy.com, thereporterz.com, therightists.com, therightstuff.biz, theshovel.com.au, theskunk.org, thespoof.com, thestatelyharold.com, thetimesoftheworld.com, thetruthdivision.com, thetruthseeker.co.uk, theunrealtimes.com, theuspatriot.com, thevalleyreport.com, thewatchtowers.com, threepercenternation.com, topekasnews.com, topinfopost.com, trueactivist.com, truepundit.com, trumpvision365.com, truthandaction.org, truthbroadcastnetwork.com, truthfeed.com, truthfrequencyradio.com, truthkings.com, truthrevolt.org, twitchy.com, ufoholic.com, unclesamsmisguidedchildren.com, unconfirmedsources.com, undergroundworldnews.com, unitedmediapublishing.com, us.blastingnews.com, usadailypolitics.com, usahitman.com, usanewsflash.com, usanewsinsider.com, usapoliticstoday.com, usasupreme.com, uschronicle.com, usconservativetoday.com, usdefensewatch.com, ushealthyadvisor.com, usherald.com, uspoliticslive.com, usuncut.com, veteranstoday.com, vigilantcitizen.com, viralliberty.com, wakingupwisconsin.com, washingtonexaminer.com, washingtonsblog.com, waterfordwhispersnews.com, wearechange.org, webdaily.com, weeklyworldnews.com, westernjournalism.com, whatdoesitmean.com, whatreallyhappened.com, whitepower.com, whowhatwhy.com, whydontyoutrythis.com, wikileaks.com, wikileaks.org, willyloman.wordpress.com, winkprogress.com, winningdemocrats.com, witscience.org, wnd.com, wonkie.com, world24monitor.com, worldcallyoutoday.com, worlddaily.info, worldnewsdailyreport.com, worldnewspolitics.com, worldpoliticsus.com, worldrumor.com, worldstoriestoday.com, worldtruth.tv, worldwidehealthy.com, wundergroundmusic.com, rt.com, yellowhammernews.com, yesimright.com, youngcons.com, yourfunpage.com, yournewswire.com, zerohedge.com, zootfeed.com

Table A2: List of Additional Domains Included in *fake site visits* from AG

about2day.com, activeopinion.com, aldipest.com, alexanderhiggins.com, alternativenewsnetwork.net, americafans.com, americanflare.com, americanjournalreview.com, americanmilitarynews.com, americanow.com, americanpoliticnews.com, americanupdater.com, americarightnow.com, americasnewest.com, amunweb.com, anews24.org, angrypatriots.com, baltimoregazette.com, bizpacreview.com, bolly.news, breitbartt.co, butthatsnoneofmybusiness.com, celebrityhealthfitness.com, centrictv.com, channel16news.com, choiceandtruth.com, cjpearson.org, consamerica.com, consamericans.com, consciouslyenlightened.com, conservativearmy88.com, conservativebase.com, conservativeeagles.com, conservativefighters.com, conservativeinsider.co, conservativeintel.com, conservativepost.com, conservativestudio.com, conservativesus.com, consfreedom.com, consnation.com, cooltobeconservative.com, currenttopnews.com, daily-sun.com, dailycaller.com, dailydot.com, dailynewsposts.info, dailyo.in, dailypresser.com, damnlikes.com, dangerandplay.com, dcwatchdog.org, deathandtaxesmag.com, defund.com, democracynow.org, deprogramyourself.org, digg.com, distractify.com, diyhilfe.com, drrichswier.com, duanelester.com, eheadlines.com, en-volve.com, everyday24.net, everynewshere.com, extensivenews.com, extremelynewsworthy.com, fanzinger.com, fedsalert.com, fox17online.com, fox4kc.com, freedomsfinalstand.com, freemarketcentral.com, funkydineva.com, fury.news, glennbeck.com, greenmond.com, greenvillegazette.com, guerilla.news, halturnershow.com, hannity.com, hiddenamericans.com, hillarydaily.com, hitpolitics.com, imjussayin.co, incredibleusanews.com, informationliberation.com, inquisitr.com, ipatriot.com, joeforamerica.com, kdvr.com, kfor.com, khou.com, latest.com, liberalsociety.com, linktv.org, mainerepublicemailalert.com, maxkeep.com, mediaite.com, minds.com, miniplanet.us, mirrorspectrum.com, mostextreme.us, mrcblog.com, msfanpage.link, murbles.com, myfox8.com, nationalinsiderpolitics.com, neonnettle.com, netlivemedia.com, newromantimes.com, newsbian.com, newsinworld365days.com, newsiosity.com, noscomunicamos.com, organicandhealthy.org, overpassesforamerica.com, patriothangout.com, patriotnewsagency.com, patriottribune.com, pjmedia.com, politicono.com, politicot.com, politicscorner.today, politicsforum.online, politistick.com, politleague.com, professionalmac.com, pundittoday.com, puppetstringnews.com, rebelcowgirlroundup.com, redalertpolitics.com, religiousmind.com, reportme24.com, samuel-warde.com, sbs.com.au, schatziesearthproject.com, simplecapacity.com, sonsoflibertymedia.com, sourcesnews.com, spinzon.com, spotlighttimes.com, statenation.co, stateofthenation2012.com, superstation95.com, tap-news.com, tdnewswire.com, tdtalliance.com, teaparty.org, text143.com, tfhnews.com, the-insider.co, theamericanmirror.com, theantimedia.org, thebiafraherald.co, thecarsmagazine.com, theconservativeclub.us, thehayride.com, theintellectualist.co, theinternationalreporter.org, thelastlineofdefense.org, thelibertarianrepublic.com, thenationalsun.com, thenewsclub.info, thenewyorkevening.com, thetrumptruck.com, thewashingtonstandard.com, tmn.today, tmzcomedy.com, todaychristian.net, trendingcult.com, trumpnews2016.org, truthinsideofyou.org, truthorfiction.com, uconservative.com, ufpnews.com, unilad.co.uk, unitedstates-politics.com, untoldnews.net, urbannewsletter.com, usa2016elections.com, usaaroundtheworldnews.com, usadailyinfo.com, usadailytime.com, usainfobox.com, usainfonews.com, usalibertynews.com, usanewshome.com, usapoliticsnow.com, usatodaypolitics.com, usatwentyfour.com, usbreakingnewsfeed.com, usdailypolitic.com, uspoln.com, vesselnews.io, veteransnewsnow.com, vidaguerrablog.com, viraldiesel.com, viralows.com, voiceofshadows.com, vote.us.org, weeklypopnews.com, westernsentinel.com, wgntv.com, whatsupic.com, wizardofviral.com, wordondastreet.com, worldaily.info, worldinformation24.info, worldpoliticus.com, wtoe5news.com, wuc-news.com, yepsee.com, zgarlic.com

Table A3: List of Domains Included in *news site visits*

abcnews.com, ajc.com, azcentral.com, baltimoresun.com, bbc.co.uk, bbc.com, bostonglobe.com, bostonherald.com, cbsnews.com, charlotteobserver.com, chicagotribune.com, chron.com, cincinnati.com, cleveland.com, cnn.com, courier-journal.com, dallasnews.com, denverpost.com, detroitnews.com, dispatch.com, economist.com, foxnews.com, freep.com, ft.com, independent.co.uk, indystar.com, jsonline.com, kansascity.com, latimes.com, mercurynews.com, miamiherald.com, msnbc.com, mysanantonio.com, nature.com, nbcnews.com, newsday.com, newyorktimes.com, nj.com, nydailynews.com, nypost.com, nytimes.com, ocregister.com, oklahoman.com, oregonlive.com, orlandosentinal.com, philly.com, phys.org, post-gazette.com, reuters.com, richmond.com, rockymountainnews.com, sacbee.com, sandiegouniontribune.com, scientificamerican.com, seattletimes.com, sfchronicle.com, sfgate.com, slate.com, sltrib.com, sptimes.com, star-telegram.com, startribune.com, stltoday.com, sun-sentinel.com, tampabay.com, twincities.com, usatoday.com, washingtonpost.com, and wsj.com.

Table A4: List of Domains of Fake News Articles Included in Fake News Supply Measure

	Article Count		Strict Definition		Loose Definition	
Domain	Total	In Sample	Site Down	Prediction	Site Down	Prediction
100percentfedup.com	7	3	No	0.008	No	0.015
1776coalition.com	1	1	No	0.008	No	0.013
24usainfo.com	2	2	Yes	0.999	Yes	0.979
365usanews.com	1	0	Yes	0.969	Yes	0.972
710wor.iheart.com	1	1	No	0.094	No	0.178
abcnews.com.co	2	1	No	0.036	No	0.229
about2day.com	1	0	Yes	0.670	Yes	0.758
activeopinion.com	1	1	No	0.043	Yes	0.919
aldipest.com	2	2	Yes	0.871	Yes	0.976
alexanderhiggins.com	1	1	No	0.010	No	0.038
allnewspipeline.com	1	0	No	0.036	No	0.081
alternativenewsnetwork.net	1	1	No	0.043	No	0.084
americafans.com	1	1	No	0.011	No	0.315
americanflare.com	2	0	No	0.143	No	0.578
americanjournalreview.com	2	0	No	0.060	No	0.158
americanlookout.com	1	1	No	0.002	No	0.161
americanmilitarynews.com	1	1	No	0.004	No	0.052
americannews.com	2	0	No	0.003	No	0.022
americanow.com	1	1	No	0.029	No	0.037
americanpoliticnews.com	2	0	No	0.012	No	0.203
americantoday.news	1	0	No	0.001	No	0.056
americanupdater.com	1	1	No	0.010	Yes	0.234
americarightnow.com	1	0	Yes	0.888	Yes	0.875
americasfreedomfighters.com	4	3	No	0.003	No	0.025

Continued on next page

Table A4 – continued from previous page

Domain	Article Count: Total	Article Count: In Sample	Strict Definition: Site Down	Strict Definition: Prediction	Loose Definition: Site Down	Loose Definition: Prediction
americasnewest.com	2	1	Yes	0.939	Yes	0.907
amunweb.com	1	1	No	0.495	Yes	0.987
anews24.org	1	0	Yes	0.742	Yes	0.489
angrypatriotmovement.com	3	3	No	0.033	No	0.367
angrypatriots.com	1	1	No	0.007	Yes	0.602
anonews.co	1	1	No	0.266	No	0.308
anonhq.com	1	0	Yes	0.107	Yes	0.375
awarenessact.com	2	2	No	0.006	No	0.017
baltimoregazette.com	1	1	No	0.354	Yes	0.568
beforeitsnews.com	1	0	No	0.020	No	0.155
bients.com	5	3	Yes	0.630	Yes	0.673
bigbluedimension.com	1	0	Yes	0.628	Yes	0.832
bigbluevision.org	4	2	Yes	0.740	Yes	0.698
bignuggetnews.com	2	1	Yes	0.285	Yes	0.566
bipartisanreport.com	4	2	No	0.023	No	0.057
bizpacreview.com	2	2	No	0.015	No	0.144
bloomberg.com	1	0	No	0.001	No	0.004
bolly.news	1	0	Yes	0.957	Yes	0.919
breitbart.com	3	1	No	0.160	No	0.142
breitbartt.co	1	0	Yes	0.883	Yes	0.860
burrardstreetjournal.com	2	0	No	0.295	No	0.580
butthatsnoneofmybusiness.com	1	0	No	0.112	No	0.387
buzzfeed.com	1	0	No	0.007	No	0.042
buzzfeedusa.com	4	3	Yes	0.895	Yes	0.805
celebrityhealthfitness.com	1	0	No	0.009	No	0.016
centrictv.com	1	0	No	0.019	No	0.030
channel16news.com	1	0	Yes	0.319	Yes	0.612
chicksontheright.com	2	2	No	0.019	No	0.034
choiceandtruth.com	2	0	No	0.018	Yes	0.813
christiantimesnewspaper.com	2	2	Yes	0.374	Yes	0.594
chuckcallesto.blogspot.com	2	1	No	0.167	Yes	0.265
cjpearson.org	1	1	No	0.117	No	0.354
clashdaily.com	3	3	No	0.009	No	0.016
collective-evolution.com	1	1	No	0.008	No	0.009
consamerica.com	3	1	No	0.006	No	0.222
consamericans.com	3	2	No	0.006	Yes	0.229
consciouslyenlightened.com	3	1	No	0.258	Yes	0.578
conservativearmy88.com	1	0	No	0.288	No	0.468
conservativebase.com	1	1	No	0.011	No	0.024
conservativedailypost.com	4	4	No	0.604	No	0.544
conservativeeagles.com	1	0	Yes	0.663	Yes	0.668
conservativefighters.com	2	1	No	0.628	No	0.574

Continued on next page

Table A4 – continued from previous page

Domain	Article Count: Total	Article Count: In Sample	Strict Definition: Site Down	Strict Definition: Prediction	Loose Definition: Site Down	Loose Definition: Prediction
conservativefiringline.com	3	3	No	0.005	No	0.009
conservativeinsider.co	2	1	Yes	0.614	Yes	0.696
conservativeintel.com	1	1	No	0.007	No	0.008
conservativeoutfitters.com	3	2	No	0.021	No	0.202
conservativepost.com	6	4	No	0.030	No	0.182
conservativestate.com	1	1	Yes	0.655	Yes	0.658
conservativestudio.com	2	2	No	0.290	No	0.471
conservativesus.com	1	0	Yes	0.498	Yes	0.553
conservativetribune.com	5	4	No	0.014	No	0.100
consfreedom.com	2	0	Yes	0.527	Yes	0.682
consnation.com	2	1	No	0.523	No	0.678
cooltobeconservative.com	2	1	No	0.046	No	0.028
creambmp.com	1	0	No	0.016	No	0.081
currenttopnews.com	1	0	No	0.690	Yes	0.778
daily-sun.com	2	2	No	0.004	No	0.013
dailycaller.com	4	3	No	0.008	No	0.014
dailydot.com	1	0	No	0.002	No	0.081
dailyheadlines.net	4	3	No	0.046	No	0.328
dailymail.co.uk	1	1	No	0.000	No	0.024
dailynewsposts.info	4	1	No	0.001	No	0.021
dailyo.in	1	0	No	0.012	No	0.025
dailyoccupation.com	2	2	No	0.043	No	0.081
dailypresser.com	2	1	No	0.285	No	0.566
dailywire.com	2	1	No	0.003	No	0.033
damnlikes.com	1	1	No	0.031	Yes	0.893
dangerandplay.com	1	1	No	0.297	Yes	0.462
dcclothesline.com	2	1	No	0.010	No	0.109
dcwatchdog.org	1	1	Yes	0.118	Yes	0.348
deathandtaxesmag.com	1	0	No	0.016	No	0.646
defund.com	8	6	No	0.010	No	0.038
democracynow.org	1	0	No	0.029	No	0.035
departed.co	5	3	Yes	0.986	Yes	0.977
deprogramyourself.org	1	0	No	0.404	No	0.380
digg.com	1	0	No	0.065	No	0.070
dineal.com	3	1	No	0.681	No	0.769
distractify.com	1	0	No	0.010	No	0.140
diyhilfe.com	1	0	Yes	0.637	Yes	0.680
donaldtrumpnews.co	16	7	No	0.516	No	0.363
downtrend.com	1	1	No	0.227	No	0.080
drrichswier.com	1	1	No	0.038	No	0.308
duanelester.com	1	1	No	0.003	No	0.005
eaglerising.com	1	1	No	0.064	No	0.029

Continued on next page

Table A4 – continued from previous page

Domain	Article Count: Total	Article Count: In Sample	Strict Definition: Site Down	Strict Definition: Prediction	Loose Definition: Site Down	Loose Definition: Prediction
educateinspirechange.org	1	0	No	0.111	Yes	0.328
eheadlines.com	1	0	No	0.007	No	0.075
electionfraud2016.wordpress.com	1	0	No	0.155	No	0.245
embols.com	5	2	Yes	0.303	Yes	0.591
empirenews.net	1	1	No	0.030	No	0.293
en-volve.com	4	2	No	0.038	No	0.567
en.institutomanquehue.org	1	1	No	0.014	Yes	0.919
endingthefed.com	41	30	No	0.082	Yes	0.738
eutimes.net	3	3	No	0.008	No	0.026
everyday24.net	1	1	Yes	0.957	Yes	0.919
everynewshere.com	14	10	Yes	0.960	Yes	0.786
extensivenews.com	1	0	Yes	0.308	Yes	0.598
extremelynewsworthy.com	1	0	Yes	0.957	Yes	0.919
fanzinger.com	2	0	No	0.018	No	0.031
fedsalert.com	1	1	No	0.311	Yes	0.602
fellowshipoftheminds.com	1	0	No	0.037	No	0.300
fox17online.com	1	0	No	0.007	No	0.032
fox4kc.com	1	0	No	0.011	No	0.051
freakoutnation.com	1	0	No	0.022	No	0.036
freedomdaily.com	3	2	No	0.025	Yes	0.464
freedomoutpost.com	1	1	No	0.048	No	0.030
freedomsfinalstand.com	2	2	No	0.284	No	0.564
freemarketcentral.com	1	1	No	0.087	No	0.410
friendsofsyria.wordpress.com	3	3	No	0.155	No	0.245
funkydineva.com	1	0	Yes	0.681	Yes	0.461
fury.news	3	3	No	0.081	No	0.281
getpocket.com	1	0	No	0.097	No	0.273
glennbeck.com	1	1	No	0.001	No	0.006
globalresearch.ca	2	1	No	0.002	No	0.081
gopthedailydose.com	3	3	No	0.001	No	0.001
greenmond.com	1	0	No	0.091	Yes	0.763
greenvillegazette.com	2	2	No	0.026	No	0.301
guerilla.news	2	2	No	0.463	No	0.271
halturnershow.com	2	0	Yes	0.453	Yes	0.572
hannity.com	1	0	No	0.004	No	0.014
hiddenamericans.com	1	1	No	0.055	No	0.142
hillarydaily.com	4	3	No	0.032	No	0.264
hitpolitics.com	1	1	Yes	0.657	Yes	0.559
huffingtonpost.co.uk	1	0	No	0.000	No	0.033
huffingtonpost.com	1	0	No	0.009	No	0.028
ihavethetruth.com	8	7	No	0.029	No	0.076
ijr.com	4	2	No	0.003	No	0.009

Continued on next page

Table A4 – continued from previous page

Domain	Article Count Total	Article Count In Sample	Strict Definition Site Down	Strict Definition Prediction	Loose Definition Site Down	Loose Definition Prediction
ilovemyfreedom.org	4	2	No	0.168	No	0.372
imjussayin.co	1	0	No	0.001	No	0.002
incredibleusanews.com	1	1	Yes	0.750	Yes	0.964
independent.co.uk	1	0	No	0.000	No	0.024
informationliberation.com	1	1	No	0.015	No	0.051
infowars.com	6	4	No	0.043	No	0.091
inquisitr.com	2	2	No	0.006	No	0.074
intrendtoday.com	2	1	No	0.365	Yes	0.583
ipatriot.com	2	1	No	0.004	No	0.017
jewsnews.co.il	8	7	No	0.001	No	0.022
joeforamerica.com	4	3	No	0.001	No	0.002
kdvr.com	1	0	No	0.003	No	0.015
kfor.com	1	0	No	0.001	No	0.007
khou.com	1	0	No	0.001	No	0.006
latest.com	1	0	No	0.002	No	0.008
liberalamerica.org	3	1	No	0.001	No	0.022
liberalsociety.com	1	1	No	0.249	Yes	0.760
libertynews.com	1	1	No	0.034	No	0.054
libertywritersnews.com	2	1	No	0.059	No	0.593
lifenews.com	1	1	No	0.002	No	0.040
linktv.org	1	0	No	0.026	No	0.035
mainerepublicemailalert.com	2	1	No	0.038	No	0.304
maxkeep.com	1	1	Yes	0.957	Yes	0.919
mediaite.com	1	1	No	0.003	No	0.013
mediazone.news	6	6	Yes	0.467	Yes	0.275
minds.com	1	1	No	0.002	No	0.015
miniplanet.us	1	1	No	0.148	No	0.410
mirrorspectrum.com	1	0	Yes	0.344	Yes	0.556
morningnewsusa.com	2	2	No	0.019	No	0.334
mostextreme.us	1	1	Yes	0.957	Yes	0.919
mrcblog.com	1	0	No	0.301	No	0.588
msfanpage.link	3	2	No	0.043	No	0.081
murbles.com	1	0	Yes	0.994	Yes	0.922
myfox8.com	1	0	No	0.007	No	0.041
myfreshnews.com	5	1	No	0.036	Yes	0.551
nationalinsiderpolitics.com	6	4	Yes	0.535	Yes	0.544
naturalnews.com	6	3	No	0.010	No	0.063
ncscooper.com	1	1	No	0.036	Yes	0.919
neonnettle.com	1	0	No	0.010	No	0.104
netlivemedia.com	1	1	Yes	0.301	Yes	0.589
nevo.news	6	3	No	0.086	No	0.297
newromantimes.com	1	1	No	0.003	No	0.018

Continued on next page

Table A4 – continued from previous page

Domain	Article Count		Strict Definition		Loose Definition	
	Total	In Sample	Site Down	Prediction	Site Down	Prediction
newsbian.com	1	0	Yes	0.970	Yes	0.914
newsfrompolitics.com	1	0	No	0.663	No	0.619
newsinworld365days.com	1	0	Yes	0.685	Yes	0.773
newsiosity.com	1	1	No	0.103	No	0.361
newsninja2012.com	1	1	Yes	0.297	Yes	0.462
newyorker.com	4	2	No	0.001	No	0.004
noscomunicamos.com	3	2	Yes	0.957	Yes	0.919
nowtheendbegins.com	4	0	No	0.010	No	0.031
nuevoordenmundialreptiliano.blogspot.com	1	1	No	0.167	No	0.265
nydailynews.com	1	0	No	0.001	No	0.003
nymag.com	1	0	No	0.001	No	0.044
nypost.com	1	0	No	0.001	No	0.009
occupydemocrats.com	1	0	No	0.007	No	0.021
openmindmagazine.com	1	1	Yes	0.673	Yes	0.761
organicandhealthy.org	1	0	Yes	0.956	Yes	0.919
overpassesforamerica.com	1	0	No	0.001	No	0.025
patriothangout.com	1	0	No	0.044	Yes	0.969
patriotnewsagency.com	1	1	Yes	0.838	Yes	0.810
patriottribune.com	1	1	No	0.146	No	0.364
people.com	1	0	No	0.000	No	0.024
pjmedia.com	2	1	No	0.075	No	0.040
politicalsitenews.com	1	0	No	0.247	No	0.627
politicono.com	1	1	No	0.277	No	0.264
politicops.com	3	1	Yes	0.270	Yes	0.256
politicot.com	1	1	No	0.274	No	0.260
politicscorner.today	4	2	Yes	0.923	Yes	0.800
politicsforum.online	1	0	Yes	0.986	Yes	0.933
politicususa.com	3	0	No	0.038	No	0.050
politistick.com	1	1	Yes	0.280	Yes	0.740
politleague.com	1	1	No	0.043	Yes	0.919
presidentialvoting2016.com	1	0	No	0.603	Yes	0.991
press24.us	1	1	No	0.541	No	0.717
prntly.com	4	1	No	0.268	No	0.732
professionalmac.com	1	1	No	0.361	Yes	0.577
punditto day.com	1	0	No	0.310	Yes	0.601
puppetstringnews.com	1	1	No	0.043	No	0.081
rawstory.com	1	0	No	0.013	No	0.042
reagancoalition.com	1	1	No	0.009	No	0.025
rebelcowgirlroundup.com	1	0	Yes	0.881	Yes	0.979
redalertpolitics.com	1	0	No	0.040	No	0.047
redflagnews.com	8	7	No	0.003	No	0.003
redstatewatcher.com	16	11	No	0.140	No	0.412

Continued on next page

Table A4 – continued from previous page

	Article Count		Strict Definition		Loose Definition	
Domain	Total	In Sample	Site Down	Prediction	Site Down	Prediction
religiousmind.com	1	1	No	0.340	No	0.551
reportme24.com	1	1	Yes	0.679	Yes	0.768
rickwells.us	5	3	Yes	0.679	Yes	0.567
rightwingnews.com	2	2	No	0.084	No	0.152
rt.com	1	0	No	0.000	No	0.001
samuel-warde.com	1	1	No	0.288	No	0.450
sbs.com.au	1	0	No	0.043	No	0.081
schatziesearthproject.com	1	0	No	0.037	Yes	0.919
secretsofthefed.com	1	1	No	0.003	No	0.008
simplecapacity.com	1	1	No	0.239	Yes	0.695
slate.com	1	0	No	0.001	No	0.006
sonsoflibertymedia.com	3	3	Yes	0.818	Yes	0.671
sourcesnews.com	1	1	No	0.088	No	0.757
spinzon.com	3	2	No	0.638	No	0.681
spotlighttimes.com	1	1	Yes	0.689	Yes	0.777
statenation.co	4	4	Yes	0.902	Yes	0.941
stateofthenation2012.com	2	1	No	0.009	No	0.016
subjectpolitics.com	1	1	No	0.085	No	0.048
superstation95.com	2	1	Yes	0.590	Yes	0.834
talkingpointsmemo.com	1	1	No	0.005	No	0.071
tap-news.com	1	1	Yes	0.642	Yes	0.843
tdnewswire.com	2	2	Yes	0.992	Yes	0.988
tdtalliance.com	1	1	No	0.361	No	0.578
teaparty.org	7	5	No	0.002	No	0.012
text143.com	1	0	Yes	0.549	Yes	0.980
tfhnews.com	1	1	No	0.370	Yes	0.590
the-insider.co	1	0	Yes	0.917	Yes	0.747
theamericanmirror.com	1	1	No	0.015	No	0.089
theantimedia.org	1	1	No	0.003	No	0.007
thebiafraherald.co	1	1	No	0.000	No	0.059
theblaze.com	1	1	No	0.171	No	0.273
thecarsmagazine.com	1	0	Yes	0.583	Yes	0.398
theconservativeclub.us	1	1	Yes	0.577	Yes	0.504
thedailybeast.com	1	0	No	0.010	No	0.032
theduran.com	1	1	No	0.631	No	0.673
thefederalistpapers.org	4	2	No	0.119	No	0.246
thefreepatriot.org	2	2	No	0.044	No	0.081
thefreethoughtproject.com	1	1	No	0.038	No	0.304
thegatewaypundit.com	11	9	No	0.267	No	0.418
thehayride.com	1	1	No	0.008	No	0.014
theintellectualist.co	1	1	No	0.035	No	0.129
theinternationalreporter.org	2	2	No	0.164	No	0.364

Continued on next page

Table A4 – continued from previous page

Domain	Article Count		Strict Definition		Loose Definition	
	Total	In Sample	Site Down	Prediction	Site Down	Prediction
thelastlineofdefense.org	1	1	No	0.393	No	0.367
thelibertarianrepublic.com	1	1	No	0.003	No	0.002
themuslimissue.wordpress.com	1	1	Yes	0.155	Yes	0.245
thenationalsun.com	1	0	Yes	0.672	Yes	0.840
thenewsclub.info	2	1	Yes	0.984	Yes	0.929
thenewyorkevening.com	2	1	No	0.667	No	0.756
thepoliticalinsider.com	3	3	No	0.011	No	0.045
therealstrategy.com	6	4	No	0.129	No	0.266
therightists.com	2	2	No	0.265	No	0.382
thetrumptruck.com	1	1	Yes	0.951	Yes	0.897
thewashingtonstandard.com	1	1	No	0.264	No	0.537
tmn.today	2	2	No	0.016	No	0.086
tmzcomedy.com	1	0	No	0.296	Yes	0.582
todaychristian.net	1	1	No	0.042	Yes	0.301
topinfopost.com	1	1	No	0.034	Yes	0.276
trendingcult.com	1	0	Yes	0.969	Yes	0.910
trueactivist.com	1	1	No	0.057	No	0.109
trumpnews2016.org	1	0	Yes	0.607	Yes	0.880
truthandaction.org	5	2	No	0.182	No	0.276
truthfeed.com	16	9	No	0.327	No	0.738
truthinsideofyou.org	1	1	No	0.009	No	0.146
truthkings.com	3	2	No	0.050	Yes	0.632
truthorfiction.com	1	1	Yes	0.224	Yes	0.173
twitchy.com	1	1	No	0.003	No	0.024
u281p372.newsninjaa.com	1	0	Yes	0.971	Yes	0.914
u4281p6798.newzfeed.net	1	0	Yes	0.987	Yes	0.967
u4638p2660.newsninjaa.com	1	0	Yes	0.971	Yes	0.914
u7176p8678.newzfeed.net	1	0	Yes	0.987	Yes	0.967
u7434p6564.mrsjekyllsays.net	1	0	Yes	0.987	Yes	0.966
u7690p7141.liamtheleprechaun.co	1	0	Yes	0.992	Yes	0.965
u8177p8978.newzfeed.net	1	0	Yes	0.987	Yes	0.967
uconservative.com	8	4	No	0.222	Yes	0.637
ufpnews.com	1	1	No	0.037	Yes	0.303
unclesamsmisguidedchildren.com	1	1	Yes	0.287	Yes	0.448
unilad.co.uk	1	0	No	0.001	Yes	0.761
unitedstates-politics.com	2	1	Yes	0.953	Yes	0.909
untoldnews.net	1	0	Yes	0.939	Yes	0.917
uprootedpalestinians.wordpress.com	1	1	No	0.155	No	0.245
urbannewsletter.com	1	1	Yes	0.870	Yes	0.976
us.blastingnews.com	1	0	No	0.003	No	0.071
usa2016elections.com	1	1	Yes	0.714	Yes	0.989
usaaroundtheworldnews.com	1	0	Yes	0.442	Yes	0.588

Continued on next page

Table A4 – continued from previous page

Domain	Article Count		Strict Definition		Loose Definition	
	Total	In Sample	Site Down	Prediction	Site Down	Prediction
usadailyinfo.com	6	3	Yes	0.439	Yes	0.584
usadailytime.com	3	1	No	0.422	No	0.564
usainfobox.com	1	0	Yes	0.424	Yes	0.567
usainfonews.com	1	0	No	0.497	No	0.698
usalibertynews.com	1	1	No	0.003	Yes	0.521
usanewsflash.com	9	6	No	0.302	No	0.544
usanewshome.com	2	0	No	0.156	Yes	0.475
usanewsinsider.com	3	0	No	0.750	Yes	0.964
usapoliticsnow.com	3	2	No	0.125	No	0.518
usapoliticstoday.com	7	4	No	0.162	Yes	0.543
usasupreme.com	9	6	No	0.159	No	0.481
usatodaypolitics.com	2	0	Yes	0.773	Yes	0.944
usatwentyfour.com	3	3	No	0.155	No	0.276
usbreakingnewsfeed.com	1	0	Yes	0.622	Yes	0.663
usdailypolitic.com	1	0	No	0.689	Yes	0.651
usdefensewatch.com	1	1	Yes	0.336	Yes	0.544
ushealthyadvisor.com	2	2	No	0.638	No	0.681
usherald.com	3	2	No	0.001	Yes	0.552
uspoliticslive.com	1	0	Yes	0.253	Yes	0.636
uspoln.com	2	2	No	0.307	Yes	0.597
vesselnews.io	2	2	No	0.043	No	0.081
veteransnewsnow.com	1	1	No	0.284	No	0.443
vidaguerrablog.com	1	0	Yes	0.968	Yes	0.905
vigilantcitizen.com	1	1	No	0.007	No	0.013
viraldiesel.com	1	0	Yes	0.969	Yes	0.908
viralows.com	1	0	Yes	0.970	Yes	0.914
voiceofshadows.com	1	0	Yes	0.302	Yes	0.590
vote.us.org	1	1	No	0.030	No	0.081
washingtontimes.com	1	0	No	0.001	No	0.007
wearechange.org	2	2	No	0.064	No	0.017
weeklypopnews.com	1	0	No	0.091	Yes	0.763
westernjournalism.com	1	1	No	0.098	No	0.109
westernsentinel.com	2	1	No	0.286	No	0.567
wgntv.com	1	0	No	0.001	No	0.006
whatdoesitmean.com	3	3	No	0.003	No	0.017
whatsupic.com	3	3	No	0.005	No	0.005
winningdemocrats.com	1	0	Yes	0.360	Yes	0.576
wizardofviral.com	1	0	Yes	0.970	Yes	0.912
wnd.com	1	0	No	0.001	No	0.007
wordondastreet.com	1	0	No	0.009	No	0.017
worldaily.info	1	1	Yes	0.956	Yes	0.919
worldcallyoutoday.com	1	1	Yes	0.679	Yes	0.768

Continued on next page

Table A4 – continued from previous page

Domain	Article Count Total	Article Count In Sample	Strict Definition Site Down	Strict Definition Prediction	Loose Definition Site Down	Loose Definition Prediction
worldinformation24.info	2	1	No	0.118	No	0.200
worldnewsdailyreport.com	2	2	No	0.099	Yes	0.645
worldnewspolitics.com	8	3	Yes	0.314	Yes	0.657
worldpoliticus.com	6	4	No	0.515	No	0.520
worldtruth.tv	7	5	No	0.043	Yes	0.919
wtoe5news.com	1	0	Yes	0.305	Yes	0.594
wuc-news.com	1	1	Yes	0.659	Yes	0.748
yepsee.com	1	0	No	0.014	Yes	0.136
yesimright.com	6	5	No	0.091	Yes	0.608
youngcons.com	3	3	No	0.253	No	0.681
yournewswire.com	11	6	No	0.036	No	0.253
youtube.com	65	40	No	0.009	No	0.032
zerohedge.com	1	1	No	0.007	No	0.008
zgarlic.com	1	0	No	0.043	Yes	0.919
zootfeed.com	2	1	No	0.301	No	0.588

Table A5: Weights Obtained from Machine Learning Model Predicting Whether Site is Down

Variable/Feature	Weight Strict Definition	Weight Loose Definition
(Intercept)	-16.0655	-14.7938
create_date	0.0006	0.0007
create_year	0.0000	0.0000
state_al	-4.9945	-3.5136
state_az	0.3562	0.0000
state_ca	0.0000	0.0000
state_co	0.0000	0.0000
state_ct	0.0000	0.0000
state_dc	0.0000	0.0000
state_de	5.8015	1.0243
state_fl	0.0000	0.0000
state_ga	1.8367	0.7944
state_ia	0.0000	0.0000
state_il	0.0000	0.0000
state_intl	0.0000	0.0000
state_ky	0.0000	3.2838
state_la	0.0000	0.0000
state_ma	0.0000	0.0000
state_mi	-0.0017	0.0000

Continued on next page

Table A5 – continued from previous page

Variable/Feature	Weight	
	Strict Definition	Loose Definition
state_mn	-1.2472	2.0745
state_mo	0.0000	-0.2959
state_mt	0.0000	-1.1687
state_nc	0.0000	-1.7485
state_nh	0.0000	1.3079
state_nj	0.0000	0.0000
state_nv	0.0000	0.0000
state_ny	0.0000	0.0000
state_ok	0.0000	0.0000
state_other	0.0000	0.0000
state_pa	0.0000	-0.0737
state_sc	0.0000	-2.2682
state_tx ×	-0.4878	0.3954
state_ut	0.0000	-0.6723
state_va	-0.2904	-1.6888
state_wa	2.5032	0.2208
state_wi	1.1280	1.6137
state_wy	0.0000	0.0000
country_al	-1.0997	2.7340
country_ar	0.6007	1.2626
country_au	1.3753	1.4032
country_bd	0.0000	0.0000
country_ca	-0.9533	-1.1136
country_ch	0.0000	0.0000
country_cl	0.0000	1.8229
country_de	-0.7137	-2.9715
country_ee	0.0000	0.0000
country_gb	0.0000	-0.0318
country_ge	1.1134	1.2062
country_il	-2.7294	-1.0120
country_in	0.0000	-1.7333
country_kv	0.0053	0.2968
country_lk	0.6311	1.2727
country_mk	1.9782	0.0000
country_other	0.0000	0.0000
country_pa	0.0000	0.0000
country_pk	-1.6178	2.3292
country_ru	0.0000	0.0000
country_ua	-4.8892	1.1569
country_us	0.0000	-0.0027
ext_au	-1.0729	-2.6085
ext_ca	0.0000	-0.1789

Continued on next page

Table A5 – continued from previous page

Variable/Feature	Weight	
	Strict Definition	Loose Definition
ext_co	-1.5308	-1.3805
ext_com	0.0000	0.0000
ext_il	0.0000	0.0000
ext_in	0.0000	0.0000
ext_info	0.0000	0.0000
ext_io	-2.4058	-3.0160
ext_link	-1.1155	-2.0521
ext_net	0.0000	0.2061
ext_news	0.0000	0.0000
ext_online	0.0000	0.0000
ext_org	0.0000	-0.1869
ext_today	2.6847	0.9177
ext_tv	-0.5589	1.3135
ext_uk	0.0000	0.0000
ext_us	0.1718	0.0000
reg_type_individual	-0.0071	0.0000
reg_type_institution	0.0000	0.0000
reg_type_private	0.8630	0.0000
tech_type_individual	0.0000	-0.8090
tech_type_institution	0.0000	0.0000
tech_type_private	0.0000	0.0000
admin_type_individual	0.0000	0.0000
admin_type_institution	0.0000	0.8137
admin_type_private	0.0000	0.0000
year_1991	0.0000	0.0000
year_1993	0.0000	0.0000
year_1994	0.0000	0.0000
year_1995	0.0000	0.0000
year_1996	0.0000	0.0000
year_1997	0.0000	0.0000
year_1998	0.0000	0.0000
year_1999	4.0001	1.3652
year_2000	3.8317	2.4722
year_2001	0.0000	0.9613
year_2002	0.0000	0.0000
year_2003	0.0000	0.0000
year_2004	0.0000	-0.4042
year_2005	-0.2399	-1.1020
year_2006	-0.7339	-1.0621
year_2007	-0.3670	0.0826
year_2008	-1.2848	-1.9608
year_2009	-0.3691	-2.1053

Continued on next page

Table A5 – continued from previous page

	Weight	
Variable/Feature	Strict Definition	Loose Definition
year_2010	0.0000	0.0799
year_2011	1.7999	0.5631
year_2012	-2.2003	-3.5825
year_2013	-1.0120	-0.6371
year_2014	0.0000	-0.4561
year_2015	1.2395	0.0000
year_2016	0.5910	0.0014
year_2017	6.9361	1.7208
dword_america	-2.1821	-1.8442
dword_patriot	-0.0014	-0.4140
dword_politic	0.0304	0.0006
dword_trump	-1.7079	-0.3937
dword_2016	0.0000	0.0000
dword_truth	0.0000	0.9032
dword_conservative	-0.1138	-0.5203
dword_liberty	0.2225	0.8749
dword_usa	-0.6864	-0.2865
bill_name_known	0.8716	1.4536
year_sq	0.0000	0.0000
year × state_al	0.0000	0.0000
year × state_az	0.0000	0.0000
year × state_ca	0.0000	0.0000
year × state_co	0.0000	0.0000
year × state_ct	0.0000	0.0000
year × state_dc	0.0000	0.0000
year × state_de	0.0000	0.0000
year × state_fl	0.0000	0.0000
year × state_ga	0.0003	0.0000
year × state_ia	0.0000	0.0000
year × state_il	0.0000	0.0000
year × state_intl	0.0000	0.0000
year × state_ky	0.0000	0.0002
year × state_la	0.0000	0.0000
year × state_ma	0.0000	0.0000
year × state_mi	0.0000	0.0000
year × state_mn	-0.0000	0.0000
year × state_mo	0.0000	-0.0000
year × state_mt	0.0000	-0.0000
year × state_nc	0.0000	-0.0000
year × state_nh	0.0000	0.0000
year × state_nj	0.0000	0.0000
year × state_nv	0.0010	0.0001

Continued on next page

Table A5 – continued from previous page

Variable/Feature	Weight	
	Strict Definition	Loose Definition
year × state_ny	0.0000	0.0000
year × state_ok	0.0000	0.0000
year × state_other	0.0000	0.0000
year × state_pa	0.0000	0.0000
year × state_sc	0.0000	-0.0004
year × state_tx	0.0000	0.0001
year × state_ut	0.0000	-0.0001
year × state_va	-0.0000	-0.0001
year × state_wa	0.0000	0.0004
year × state_wi	0.0000	0.0000
year × state_wy	0.0000	0.0000
year_sq × state_al	0.0000	0.0000
year_sq × state_az	0.0000	0.0000
year_sq × state_ca	0.0000	0.0000
year_sq × state_co	0.0000	0.0000
year_sq × state_ct	0.0000	0.0000
year_sq × state_dc	0.0000	0.0000
year_sq × state_de	0.0000	0.0000
year_sq × state_fl	0.0000	0.0000
year_sq × state_ga	0.0000	0.0000
year_sq × state_ia	0.0000	0.0000
year_sq × state_il	0.0000	0.0000
year_sq × state_intl	0.0000	0.0000
year_sq × state_ky	0.0000	0.0000
year_sq × state_la	0.0000	0.0000
year_sq × state_ma	0.0000	0.0000
year_sq × state_mi	-0.0000	0.0000
year_sq × state_mn	-0.0000	0.0000
year_sq × state_mo	0.0000	-0.0000
year_sq × state_mt	0.0000	-0.0000
year_sq × state_nc	0.0000	-0.0000
year_sq × state_nh	0.0000	0.0000
year_sq × state_nj	0.0000	0.0000
year_sq × state_nv	0.0000	0.0000
year_sq × state_ny	0.0000	0.0000
year_sq × state_ok	0.0000	0.0000
year_sq × state_other	0.0000	0.0000
year_sq × state_pa	0.0000	0.0000
year_sq × state_sc	0.0000	-0.0000
year_sq × state_tx	0.0000	0.0000
year_sq × state_ut	0.0000	0.0000
year_sq × state_va	-0.0000	-0.0000

Continued on next page

Table A5 – continued from previous page

Variable/Feature	Weight	
	Strict Definition	Loose Definition
year_sq × state_wa	0.0000	0.0000
year_sq × state_wi	0.0000	0.0000
year_sq × state_wy	0.0000	0.0000
year_2016 × state_al	0.0000	0.0000
year_2016 × state_az	0.0000	-0.1432
year_2016 × state_ca	1.2624	0.0505
year_2016 × state_co	4.6433	1.8051
year_2016 × state_fl	0.0000	0.3093
year_2016 × state_ga	0.0000	2.4465
year_2016 × state_intl	0.0000	0.0000
year_2016 × state_ma	9.1458	3.7324
year_2016 × state_mi	-7.1987	-3.8265
year_2016 × state_nj	-5.0544	-3.5861
year_2016 × state_other	4.0230	0.4370
year_2016 × state_pa	-2.0051	-0.1251
year_2016 × state_tx	1.6299	0.1222
year_2016 × state_ut	3.1243	4.1428
year_2016 × state_va	-3.6182	-1.8781
year_2016 × state_wa	0.0000	-1.2543
year_2016 × state_wi	0.2415	0.1911
year × country_al	-0.0000	0.0002
year × country_ar	0.0000	0.0000
year × country_au	0.0000	0.0000
year × country_bd	0.0000	0.0000
year × country_ca	0.0000	-0.0001
year × country_ch	0.0000	0.0000
year × country_cl	0.0000	0.0001
year × country_de	-0.0000	-0.0003
year × country_ee	0.0000	0.0000
year × country_gb	-0.0011	-0.0006
year × country_ge	0.0000	0.0000
year × country_il	0.0000	-0.0000
year × country_in	0.0000	-0.0000
year × country_kv	0.0000	0.0000
year × country_lk	0.0000	0.0000
year × country_mk	0.0000	0.0000
year × country_other	0.0000	0.0000
year × country_pa	0.0000	0.0000
year × country_pk	0.0000	0.0002
year × country_ru	0.0000	0.0000
year × country_ua	0.0000	0.0000
year × country_us	0.0000	0.0000

Continued on next page

Table A5 – continued from previous page

Variable/Feature	Weight	
	Strict Definition	Loose Definition
year_sq × country_al	-0.0000	0.0000
year_sq × country_ar	0.0000	0.0000
year_sq × country_au	0.0000	0.0000
year_sq × country_bd	0.0000	0.0000
year_sq × country_ca	0.0000	0.0000
year_sq × country_ch	0.0000	0.0000
year_sq × country_cl	0.0000	0.0000
year_sq × country_de	0.0000	0.0000
year_sq × country_ee	0.0000	0.0000
year_sq × country_gb	-0.0000	-0.0000
year_sq × country_ge	0.0000	0.0000
year_sq × country_il	0.0000	-0.0000
year_sq × country_in	0.0000	-0.0000
year_sq × country_kv	0.0000	0.0000
year_sq × country_lk	0.0000	0.0000
year_sq × country_mk	0.0000	0.0000
year_sq × country_other	0.0000	0.0000
year_sq × country_pa	0.0000	0.0000
year_sq × country_pk	0.0000	0.0000
year_sq × country_ru	0.0000	0.0000
year_sq × country_ua	-0.0000	0.0000
year_sq × country_us	0.0000	0.0000
year_2016 × country_ar	0.3415	0.0678
year_2016 × country_au	-1.0915	-0.7944
year_2016 × country_ca	0.0000	1.8788
year_2016 × country_gb	5.9610	3.4761
year_2016 × country_ge	0.2479	0.0018
year_2016 × country_il	0.0000	-0.0053
year_2016 × country_in	-5.7412	-2.9084
year_2016 × country_kv	0.0015	0.0032
year_2016 × country_lk	0.3634	0.0652
year_2016 × country_mk	-3.0917	0.0000
year_2016 × country_other	0.2354	0.0002
year_2016 × country_pa	1.9751	0.2391
year_2016 × country_pk	0.0000	0.0000
year_2016 × country_ru	8.1487	5.9630
year_2016 × country_ua	0.0000	0.0241
year_2016 × country_us	0.2237	0.0000
year × ext_au	-0.0000	-0.0002
year × ext_ca	0.0000	-0.0000
year × ext_co	-0.0001	-0.0001
year × ext_com	0.0002	0.0001

Continued on next page

Table A5 – continued from previous page

	Weight	
Variable/Feature	Strict Definition	Loose Definition
year × ext_il	0.0000	0.0000
year × ext_in	0.0000	0.0000
year × ext_info	0.0000	0.0000
year × ext_io	-0.0000	-0.0000
year × ext_link	-0.0000	-0.0001
year × ext_net	0.0000	0.0000
year × ext_news	0.0000	0.0000
year × ext_online	0.0000	0.0000
year × ext_org	0.0000	-0.0000
year × ext_today	0.0000	0.0000
year × ext_tv	-0.0002	0.0000
year × ext_uk	0.0000	0.0003
year × ext_us	0.0000	0.0000
year_sq × ext_au	-0.0000	-0.0000
year_sq × ext_ca	0.0000	-0.0000
year_sq × ext_co	0.0000	0.0000
year_sq × ext_com	0.0000	0.0000
year_sq × ext_il	0.0000	0.0000
year_sq × ext_in	0.0000	0.0000
year_sq × ext_info	0.0000	0.0000
year_sq × ext_io	0.0000	-0.0000
year_sq × ext_link	-0.0000	-0.0000
year_sq × ext_net	0.0000	0.0000
year_sq × ext_news	0.0000	0.0000
year_sq × ext_online	0.0000	0.0000
year_sq × ext_org	0.0000	-0.0000
year_sq × ext_today	0.0000	0.0000
year_sq × ext_tv	-0.0000	0.0000
year_sq × ext_uk	0.0000	0.0000
year_sq × ext_us	0.0000	0.0000
year_2016 × ext_co	4.1390	3.1931
year_2016 × ext_com	0.0000	0.0000
year_2016 × ext_info	-2.6328	-3.5030
year_2016 × ext_io	-1.7092	-0.4182
year_2016 × ext_net	1.8452	1.3321
year_2016 × ext_news	-0.5215	-0.6853
year_2016 × ext_org	1.4905	0.0000
year_2016 × ext_today	0.2320	0.8617
year_2016 × ext_us	-0.3623	-1.5459
year × reg_type_individual	-0.0000	0.0000
year × reg_type_institution	0.0000	0.0000
year × reg_type_private	0.0000	0.0000

Continued on next page

Table A5 – continued from previous page

Variable/Feature	Weight	
	Strict Definition	Loose Definition
year_sq × reg_type_individual	0.0000	0.0000
year_sq × reg_type_institution	0.0000	0.0000
year_sq × reg_type_private	0.0000	0.0000
year_2016 × reg_type_individual	0.0000	0.0000
year_2016 × reg_type_institution	-3.4814	0.0000
year_2016 × reg_type_private	0.0000	0.0000
year × tech_type_individual	0.0000	0.0000
year × tech_type_institution	0.0000	0.0000
year × tech_type_private	0.0000	0.0000
year_sq × tech_type_individual	0.0000	0.0000
year_sq × tech_type_institution	0.0000	0.0000
year_sq × tech_type_private	0.0000	0.0000
year_2016 × tech_type_individual	0.1870	0.3543
year_2016 × tech_type_institution	0.0000	0.0000
year_2016 × tech_type_private	0.0000	0.0000
year × admin_type_individual	0.0000	0.0000
year × admin_type_institution	0.0000	0.0002
year × admin_type_private	0.0000	0.0000
year_sq × admin_type_individual	0.0000	0.0000
year_sq × admin_type_institution	0.0000	0.0000
year_sq × admin_type_private	0.0000	0.0000
year_2016 × admin_type_individual	3.8418	1.1041
year_2016 × admin_type_institution	0.0000	0.0000
year_2016 × admin_type_private	0.0000	0.0000
year × dword_2016	0.0000	0.0000
year × dword_america	0.0000	-0.0001
year × dword_conservative	-0.0000	-0.0000
year × dword_liberty	0.0011	0.0001
year × dword_patriot	-0.0013	-0.0002
year × dword_politic	0.0000	0.0001
year × dword_trump	-0.0000	0.0000
year × dword_truth	0.0000	0.0000
year × dword_usa	-0.0001	-0.0001
year_sq × dword_2016	0.0000	0.0000
year_sq × dword_america	0.0000	0.0000
year_sq × dword_conservative	0.0000	0.0000
year_sq × dword_liberty	0.0000	0.0000
year_sq × dword_patriot	-0.0000	-0.0000
year_sq × dword_politic	0.0000	0.0000
year_sq × dword_trump	-0.0000	-0.0000
year_sq × dword_truth	0.0000	0.0000
year_sq × dword_usa	0.0000	0.0000

Continued on next page

Table A5 – continued from previous page

Variable/Feature	Weight	
	Strict Definition	Loose Definition
year_2016 × dword_2016	1.2270	4.2812
year_2016 × dword_america	-3.0582	0.0000
year_2016 × dword_conservative	0.0000	0.0000
year_2016 × dword_liberty	-6.6722	-0.8944
year_2016 × dword_patriot	5.3433	2.0138
year_2016 × dword_politic	-0.0496	-0.6860
year_2016 × dword_trump	-0.1469	-1.8784
year_2016 × dword_usa	0.0000	0.0000