Political Knowledge and Misinformation in the Era of Social Media: Evidence from the 2015 U.K. Election*

Kevin Munger[†], Patrick Egan[‡], Jonathan Nagler[§], Jonathan Ronen[¶] and Joshua A. Tucker[‖]

April 4, 2017

*This project is a result of research collaboration between the NYU Social Media and Political Participation (SMaPP) lab and the survey research firm YouGov, and the authors are extremely grateful to YouGov for its support of this project. The research of SMaPP lab, which is Co-Directed by Nagler and Tucker along with Richard Bonneau and John T. Jost, is generously supported by INSPIRE program of the National Science Foundation (Award SES-1248077), the New York University Global Institute for Advanced Study, and Dean Thomas Carews Research Investment Fund at New York University. All of the authors contributed to theorizing, research design, and writing of the manuscript. Munger and Ronen wrote the computer programs, and Munger analyzed the data and prepared the tables, figures, and supplementary materials. We thank Duncan Penfold-Brown and Yvan Scher for additional programming and data science support, and we thank Ken Benoit, Jaime Settle and participants at the NYU Political Economy Lunch for helpful comments.

[†]NYU, Department of Politics

[‡]NYU, Department of Politics

[§]NYU, Department of Politics

[¶]Berlin Institute for Medical Systems Biology, Max Delbrck Center for Molecular Medicine

[‖]NYU, Department of Politics

Abstract

Social media is playing a prominent role in election campaigns. Does it educate voters, or mislead them? We explore this question by measuring change in political knowledge among a panel of voters surveyed during the U.K. 2015 general election campaign. Most panelists allowed us to monitor their exposure to political information—both its content and its source—via the Twitter social media platform. Our panel design permits identification of the effect of shifts in information exposure on changes in political knowledge. We show that information from news media increased knowledge of politically relevant facts, and that information from political parties increased knowledge of party platforms. But in a troubling demonstration of campaigns' ability to manipulate knowledge, we also find that exposure to partisan information shifted voters' assessments of the economy and immigration in directions favorable to the parties' platforms, and that much of this movement was in an inaccurate direction.

Do election campaigns provide voters with the information they need to make choices aligned with their interests and values? This question has long captured the attention of political scientists, who have reached mixed conclusions about the extent to which campaigns improve objective measures of political knowledge (Bartels, 2000; Gilens, Vavreck and Cohen, 2007; Huber and Arceneaux, 2007; Johnston, Hagen and Jamieson, 2004; Kelley, 1960; Koch, 2008; Lau, Sigelman and Rovner, 2007; Milazzo, 2015).

In a related vein, another group of scholars has examined how ever-evolving media technology—including newspapers, radio, television, cable, and now the Internet—have shaped and reshaped how voters are exposed to information about public affairs and are targeted by those seeking elected office (Gentzkow, 2006; Gentzkow, Shapiro and Sinkinson, 2014; Iyengar and Hahn, 2009; Prior, 2007; Strömberg, 2004).

Here we contribute to both of these literatures by documenting how an important recent development in media technology—the widespread use of social media—is affecting the public's level of objective political knowledge in election campaigns. Social media is broadly defined as those Internet applications that allow users to create and share content over network ties. Roughly half to three-quarters of all adult Internet users across nations in the developed world are now users of these services, the most prominent of which at present include platforms such as Facebook, Twitter, LinkedIn and Instagram (Greenwood, Perrin and Duggan, 2016). Social media has become an important source of information: about one in five Americans now say they "often" get news from a social networking site (Gottfried and Shearer, 2016). Reliance on social media for political news is particularly pronounced among young people, suggesting that the aggregate importance of social media as a news source will rise over time (Gottfried et al., 2016).

A potentially critical innovation of social media technology for political knowledge is that it gives politicians and political parties virtually unmediated access to those who choose to follow them on social media sites. Content from these sources can then be shared by users, amplifying its potential persuasive effects.

These messages—such as the "tweets" released by politicians to their Twitter followers—can become news stories in themselves, and are now commonly incorporated in mainstream news coverage. We are only in the earliest days of beginning to understand the political implications of these developments. In particular, we still know very little about whether social media use causes people to become more, or less, informed about politics—and whether its availability aggravates or ameliorates the ideologically homogeneous environments to which people are already exposed by their offline social networks and their selective consumption of traditional media (Bakshy, Messing and Adamic, 2015; Barberá et al., 2015; Guess, 2016).

In this paper, we examine the effect of social media on voter knowledge during the campaign for the U.K. general election held in May 2015. The U.K.'s party system was in substantial flux at the time, making the election an excellent case with which to study political learning. The Labour Party had recently shifted left on economic and social welfare policies, veering away from the market-based "New Labour" approach championed by former Prime Minister Tony Blair (Whiteley et al., 2013). The Liberal Democrats had also altered their stances, making substantial concessions as the junior member of their governing coalition with the Conservatives. This included the party's endorsement (contrary to its campaign promises) of an increase in university tuition fees (Weaver, 2015). Yet another change was the steep ascent of the nativist U.K. Independence Party (UKIP), riding a wave of disenchantment with establishment policies on immigration and the European Union (Evans and Mellon, 2015). Following a six-month campaign, the Conservatives won an outright majority of seats in Parliament as the Liberal Democrats suffered a stunning collapse in voter support and Labour lost all but one seat in its traditional stronghold, Scotland, to the Scottish National Party (SNP).

Our data come from a panel survey conducted in four waves beginning nearly a year before the election and concluding shortly afterward. Most panelists who were users of Twitter permitted us to obtain the tweets they were exposed to, including the issue-specific content and ideological leanings of the sources that made up their Twitter feeds. To address potential threats to inference due to users' selective consumption of political news, our panel design permits a differences-in-differences identification of the causal effect of changes in media exposure on changes in objective political knowledge of the same respondent over time across different issues. We show that our results are robust to concerns about selection bias and other challenges to inference.

Our findings provide some cause for optimism about the effect of social media on political knowledge, which we measure with factual questions about issues that were salient during the campaign and with items in which respondents placed the parties' platforms on a left-right scale on these issues. Tweets from news media accounts were generally associated with increased factual knowledge, but not with the ability to correctly place the parties' platforms. By contrast, tweets about particular issues from accounts associated with the parties led to an increase in panelists' ability to correctly rank the parties' platforms on these issues.

But our results also raise some important concerns. Exposure to partisan messages was related to a net increase in general factual knowledge. However, partisan information on specific issues increased some and reduced other voters' knowledge on these issues; these changes were in directions consistent with the parties' strategic interests. Tweets from anti-immigration UKIP on the topic of immigration tended to increase voters' assessments of the rate of immigration, leading many to over-estimate this rate. Tweets from the incumbent Conservatives tended to decrease estimates of the rate of unemployment while tweets from opposition parties tended to increase these estimates, both results which are consistent with the strategic interests of the parties in shifting voters beliefs. While the net effect of partisan messages on voter knowledge was positive, this masks significant heterogeneity and the fact that a substantial share of voters became more likely to hold inaccurate beliefs over the course of the campaign due to their exposure to social media.

Our findings suggest that as social media plays an ever more prominent role in political life, its effects on political knowledge will in many ways reinforce those of traditional media. This is particularly the case with exposure to non-partisan news, which appears to perform the same function of raising information levels via social media as it does through other channels. But social media presents myriad opportunities for parties and politicians to transmit information that is unmoored from the gatekeeping and context provided by traditional news media. This appears to be aggravating information polarization, yielding a pattern of responses to questions about politically relevant facts that is observationally equivalent to partisan motivated reasoning (Bartels, 2002) or partisan cheerleading (Bullock et al., 2015). In an era of widespread media disruption and the concurrent decline of traditional news media, these developments are troubling for those who see an informed electorate as critical to the functioning of mass democracy.

Campaigns, Media Environments, and Political Knowledge

Defined by (Carpini and Keeter, 1997, 10) as "the range of factual information about politics that is stored in long-term memory," political knowledge is vital for the healthy functioning of democracy. People who are more informed about politics cast votes and engage in other political behavior that is better aligned with their preferences and interests. While political scientists have shown that knowledge deficits can in some circumstances be overcome using cognitive shortcuts, they nevertheless agree that increased political knowledge is desirable and ignorance of basic facts about public affairs is problematic (Bartels, 1996; Fowler and Margolis, 2014; Lau and Redlawsk, 2001; Lupia, 1994).

Recent work suggests that once individuals finish formal schooling, there is little hope for improvement of their "static" political knowledge Barabas et al. (2014) of long-standing institutions (e.g., "How many members are in the House of Commons?") and long-established policies and matters of law ("Are children born in the U.S. to unauthorized immigrants automatically U.S. citizens?"). By contrast, exposure to news coverage can be associated with increases in "surveillance knowledge" of more recent developments in public affairs ("Has crime risen or fallen in the past year?") and policy ("Does Cameron's recently proposed immigration bill aim to increase or decrease the number of immigrants to Britain?"). As a given policy topic attracts more media attention, people tend to become more knowledgable about that topic in particular (Barabas and Jerit, 2009). Political campaigns may play a similar role, raising knowledge of topical political issues and party platforms (Andersen, Tilley and Heath, 2005; Banducci, Giebler and Kritzinger, 2015; Tillman, 2012), although these effects can be weak (Bartels, 2000). In theory, then, news coverage and political campaigns should provide voters with the surveillance knowledge they need to understand candidates' evolving policy stances and accurately hold elected officials accountable for their successes and mistakes.

But most of what we know about what leads to shifts in political knowledge comes from studies of news coverage and campaigns conducted through traditional media, such as newspapers, radio and television. Previous work suggests caution in assuming that these findings would hold in the era of social media. New developments in media technologies often upend the way people are exposed to political information and thus change both aggregate levels of political knowledge and its variance across citizens (Prior, 2007). The introduction of broadcast television reduced voter turnout, likely due to viewers substituting out of older media such as newspapers and radio which provided more political coverage (Gentzkow, 2006). The increased number of media choices available in the post-broadcast television era has swelled the knowledge gap between those who prefer news (who consume more of it) and those who prefer entertainment (who consume less) Prior (2005). These findings are reflected in cross-national comparisons that provide evidence of a general equilibrium effect of media environments, with "public service" television broadcast systems in Denmark and Switzerland tending to produce a more informed citizenry relative to the market-based systems in the U.S. and U.K. that provide a wider array of choices to viewers (Curran et al., 2009; Iyengar et al., 2009). In addition to affording people a great degree of choice about whether to consume news at all, the tremendous number of options in today's media environment provide more opportunities to selectively consume news that they find ideologically agreeable (Iyengar and Hahn, 2009; Stroud, 2008) and become persuaded by it (DellaVigna and Kaplan, 2006; Martin and Yurukoglu, 2014).

At first blush, we might expect social media to make these developments more pronounced. Some work has shown that social media allows people to further personalize their information environments, creating an "echo chamber" is which individuals are now even less likely to encounter views with which they disagree (Conover et al., 2012; Flaxman, Goel and Rao, 2013). However, other research counters that social media can actually increase cross-partisan exchange. This is because social media exposes users to people in their networks with whom they have weak ties—individuals with whom they would have little contact were it not for social media. Because this wider web tends to be more politically heterogeneous than the smaller group of people with whom one has direct, face-to-face contact, social media can in fact lead people to encounter a more diverse set of viewpoints online than offline (Bakshy, Messing and Adamic, 2015; Barberá et al., 2015).

In addition to being of substantive interest, social media technology has unique aspects that permit the opportunity to construct particularly strong research designs to explore classic questions about media effects. Measures of actual individual exposure to traditional media like newspapers, radio and television are rare, and self-reports of exposure to all kinds of media can be quite unreliable (Prior, 2012; Scharkow, 2016). Therefore even the strongest observational studies have typically relied upon some measure of aggregate media exposure, such as residence in a television media market, as a proxy for individual exposure (Huber and Arceneaux, 2007). By contrast, on social media it is possible to objectively and unobtrusively measure individuals' exposure to information and to record both the content and its source. This permits more precise estimation of media effects at the individual level and makes it easier to assess the varying degrees of influence of different information sources, which we do here.

Hypotheses about the Effects of Social Media Exposure on Political Knowledge

We explore hypotheses about how respondents learn from different types of political tweets that are motivated by long-standing theories and questions about media effects. Following Barabas et al. (2014), we have little reason to think that exposure to social media should have any influence on static political knowledge. Rather, we expect that messages about politics transmitted on social media should affect levels of surveillance knowledge of topics like the shifting positions of the parties or politically relevant facts on issues like immigration and unemployment. Our first hypothesis is thus that—as has been shown with traditional media—exposure to messages about a particular issue on social media will increase surveillance knowledge about that issue. While we offer refinements on this hypothesis below taking into account the strategic interests of the sender – some tweets may of course not be truthful and thus *not* likely to increase knowledge – we start with a straightforward hypothesis about the likely effect of tweets: that they will cause a net increase in knowledge.

Hypothesis 1 *Exposure to information on Twitter about an issue will cause a net increase in knowledge about that issue.*

Our next hypothesis specifies how exposure to social media improves surveillance knowledge of where the major parties stand on the most relevant issues in the campaign. As discussed above, on balance previous work has suggested that voters become more informed about the relative placements of political parties on different issues over the course of election campaigns. Our research design permits the exploration of a more refined hypothesis: that this kind of knowledge is increased via exposure to messages from the parties themselves. We assume it is in a given party's interest to attempt to differentiate themselves from other parties on various issue dimensions.[1] This expectation concords with the findings in Ander-

[1]This means we are *not* assuming a Downsian world whereby there is some "correct" position on each issue for maximizing votes. If that were the case, then we would expect the exact

sen, Tilley and Heath (2005) that knowledge of party platforms increased during election campaigns in the United Kingdom. Furthermore, Banducci, Giebler and Kritzinger (2015) find that exposure to relevant media coverage tends to increase knowledge of party platforms, but that this effect is stronger in low-quality news outlets—tabloids and other purveyors of "soft news." These outlets tend to be more openly biased and inflammatory, we use this finding to motivate our hypothesis that accounts associated with political parties will have more of an impact than those associated with the media. Thus our second hypothesis is that the cumulative effect of exposure to Tweets from parties about specific issues should be a better understanding of where the parties stand *vis a vis* each other on these issues:

Hypothesis 2 *Exposure to information on Twitter about a political topic sent by political parties will increase knowledge of the parties' relative positions on that issue.*

Another kind of surveillance knowledge that we would hope voters acquire over the course of an election campaign is that of relevant political facts that allow them to judge the performance of the incumbent government. These facts—such as the state of the economy—are the types of questions for which Barabas et al. (2014) and others have found that accuracy rises with news media coverage.

Hypothesis 3 *Exposure to information on Twitter sent by a media organization on a specific issue will increase knowledge of the facts associated with that issue.*

But our expectations become less straightforward with regard to the messages transmitted about these facts by the political parties. Even in an environment where it is difficult to directly mislead voters about these facts, the parties still have the ability to strategically highlight certain facts and deemphasize others.

opposite effect: additional information from parties should make it harder for voters to correctly order the parties, as each party would locate on this ideal spot and scatter the remaining parties far away. It also means we are assuming that parties do not try to obfuscate their positions.

As unemployment had been steadily falling in the years leading up to the 2015 U.K. election, we assume that (as Vavreck (2009) as shown in the case of U.S. presidential elections) the incumbent parties (the Conservatives and the Liberal Democrats) would want this fact to be widely known while the opposition parties (Labour and UKIP) would want to obscure it. Another important issue in the 2015 election was legal immigration from the EU to the United Kingdom. Concerns about the rate of immigration were instrumental in the rise of UKIP, a party whose anti-immigration stance resonated with a large number of voters. We assume that UKIP wanted to draw as much attention to the issue as possible, and we expect that exposure to tweets about immigration sent by UKIP could do one of two things for respondents' knowledge of immigration. Tweets by UKIP could increase respondents' chances of knowing the true number of immigrants. Or, as UKIP had incentive to have respondents believe the number of immigrants was larger than the actual number, exposure to tweets by UKIP on immigration could *increase* respondents' assessments of the total number of immigrants, leading many to inaccurately over-estimate this number. These strategic incentives are reflected in our final set of hypotheses.

Hypothesis 4 *Exposure to messages sent by political parties about issues will change levels of knowledge of political facts relevant to the issue in the direction that is strategically advantageous to the party transmitting the message:*

Empirical Implications of Hypothesis 4:

- (H4A) Tweets from incumbent parties will increase the (accurate) belief that unemployment rates were declining in 2015
- (H4B) Tweets from opposition parties will increase the (inaccurate) belief that unemployment rates were increasing in 2015
- (H4C) Tweets from UKIP will increase knowledge about the correct rate of immigration to the U.K.

- (H4D) Tweets from UKIP will increase belief in the number of immigrants coming to the United Kingdom, leading many to over-estimate this number.

Data

Panel Survey

We designed a 4-wave panel survey administered by the polling firm YouGov to respondents drawn from a population of social media users, what YouGov calls their Social Media Analysis tool (SoMA).[2] The SoMA sample was created by YouGov by asking respondents who had previously claimed to use social media if they would like to participate in surveys about their social media use. A subset of these users who used Twitter also gave their Twitter account information to YouGov, who shared with us the Twitter timelines of each respondent. To preserve anonymity, YouGov did not share the actual Twitter accounts of the respondents. Thus we have no data on the tweets *sent* by our respondents, only tweets they have seen. We refer to our respondents drawn from the SoMA sample as our Social Media Users (SMU) sample. The SMU sample contains respondents from all four countries in the United Kingdom (England, Scotland, Wales and Northern Ireland).

These respondents received a financial benefit for their participation in the survey. The surveys were conducted online. Each wave lasted approximately 10 minutes, and contained between 50 and 70 questions. We supplemented these surveys responses with demographic information that YouGov asks of all of their respondents.

The retention rates for different waves of the survey can be seen in Table 1.

[2]The SoMA sample was maintained by YouGov to be able to link survey responses to observable happenings in on the social media world, and consists of 14,000 respondents, 7,000 each selected for their use of Twitter or Facebook. They recently changed the name of the sample to YouGov Social.

Overall, there were 1308 respondents retained for all 4 waves of the SMU sample, out of the 3,846 who appeared in at least one wave.[3] The retention was lowest between waves 1 and 2, but was otherwise similar to what is often seen in online panel surveys (Chang and Krosnick, 2009). Notice that the retention rate is highest between waves 3 and 4. YouGov made an intensive effort to enroll as many previous respondents for the final, post-election wave as possible. Also, wave 4 consists only of respondents who had participated in at least one of the previous three waves, to best take advantage of the panel design.

[**Table 1 Here**]

The four waves of the survey took place over the course of almost a year: wave 1 lasted 22 days and concluded on July 31, 2014; wave 2 lasted 8 days and concluded on December 11, 2014; wave 3 lasted 12 days and concluded on March 30, 2015; and wave 4 lasted 26 days and concluded on June 17, 2015. Wave 4 was in the field for an especially long time as part of the effort to increase the retention rate, and it began 2 weeks after the day of the general election on May 7, 2015.

The timing of the survey allowed us to measure attitudes and knowledge before, during and after the 2015 U.K. Parliamentary campaign and election. The “long campaign,” during which spending is regulated, officially began on December 19th, 2014, and the “short campaign,” in which parties are given time slots to broadcast their messages on TV, began March 30th (Hope, 2015). The Conservatives and Labour parties had the most seats in parliament, while the Liberal Democrats experienced a sharp decline in popular support after joining the previous coalition government with the Conservatives. The rise of UKIP was a manifestation of the dissatisfaction of the nativist right with the U.K.’s position on immigration and the EU (Evans and Mellon, 2015). The election results turned out to be a surprise, as pre-election polls badly underestimated Conservative support (Lauderdale, 2015). The Conservatives won enough seats to govern without a coalition and the Liberal Democrats were all but removed from Parliament. Despite winning 13% of the vote, UKIP won a only a single seat.

[3] In order to maintain the size of the waves, YouGov also replenished the sample, adding respondents in later waves who were not in the first wave.

In this paper, we focus on a subset of our SMU sample. Some of the respondents drawn from YouGov's pool of social media users agreed to share the contents of their Twitter feed with us in addition to taking the surveys. We call this sample the "SMU Plus" sample. Analyzing this group allows us to make an inference about the impact of exposure to political information on Twitter *among people with Twitter accounts*, this is far from a representative sample of the population, and an understanding of the differences among the populations is essential. The covariate information presented in Table 2, Panel A was asked in waves 1 and 4, and in the cases in which respondents selected different answers in different waves, the modal responses are reported.

[Table 2 Here]

Table 2, Panel A demonstrates that there are sizable difference between the SMU sample and the voting population as a whole—the SMU sample tends to be more male, better educated, higher socio-economic class, younger and more liberal, all of which is to be expected among social media users.[4] The SMU Plus sample, who shared their Twitter accounts with YouGov, are slightly more male and better educated, but in general are a reasonably representative sample of SMU users. The data in the third column are from the British Election Study's 30,000 person post-election survey (Fieldhouse et al., 2015), and serves as the best available estimate of the true values of these demographics in the British electorate. This electorate is non-representative of the population, as demonstrated in the fourth column featuring statistics from the 2011 British Census.

The SMU respondents are also more likely than the general electorate to have voted for Labour and especially the Green party in the 2015 election, as can be seen in Table 2(b). Our sample also systematically under-reports support for UKIP. Among both samples, the breakdown by country of resident is similar, but as shown in Table 2(c), our samples are light on respondents from Scotland and Northern Ireland and heavy on respondents from Wales.

[4]There might be a concern that these median values mask some over-representation of particular demographics, especially young or wealthy people. However, only 10% of our sample is under 30, and only 4% reported a household income over £100,000.

As a "control" group, we drew respondents from another YouGov sample—the "Nationally Representative" (NR) sample. These respondents were entirely distinct from the SMU group, but received an identical 4-wave panel survey. Because this sample was representative of the U.K. population, it included a large number of Twitter users, but because we did not have access to their Twitter accounts, we could not include them in our analysis.

However, we do include those respondents in the NR sample who did not use Twitter. Below, we perform analyses that use exposure to tweets as an explanatory variable. For these NR respondents, we assume that they were exposed to 0 tweets. Including these respondents thus allows us to track changes in political knowledge among non-Twitter users. Overall, there were 389 NR non-Twitter users who appeared in both waves 1 and 4 (for the party placement analysis below), and 632 who appeared in both waves 2 and 3 (for the factual question analysis).

Tweets

The "SMU Plus" subsection of respondents provided YouGov with their Twitter handles, and while we do not have access to their individual Twitter profiles or what they tweeted or retweeted, the novel aspect of our dataset is that we match the respondents' responses with the content of their Twitter timelines.[5] The timelines consist of all of the tweets to which they could potentially have been exposed during the time period from January 1st, 2014 until May 22nd, 2015,[6] divided into 4 periods: from January 1st, 2014 to the beginning of wave 1 of our survey; from the end of wave 1 to the beginning of wave 2; from the end of wave 2 to the beginning of wave 3; and from the end of wave 3 until the beginning of wave 4. We thus have access to everything tweeted by every account the respondents

[5]Our overall setup is similar to Barabas and Jerit (2009). They measure the aggregate number of times specific policy-relevant topics are covered by the media and use these general trends to explain changes in political knowledge. We are able to measure the exact distribution of topics mentioned by the media and by politicians in each respondent's Twitter timeline, giving us a more individualized measure.

[6]Excluding the days during which the surveys were actually in the field.

followed.[7]

Unlike Facebook, which uses an algorithm to tailor the order that information from friends is displayed on the user's news feed, the stream of tweets in a user's timeline is strictly chronological.[8] We cannot know which tweets among those on the timeline the user actually saw. But because the timeline is uncurated, it is reasonable to treat the tweets they saw as a random sample from all of those they received.[9] Self-reported measures of media use are fraught with measurement error (Prior, 2013). Although we ask respondents outright how often they use Twitter, the validity of this information is difficult to verify. We use this variable as a covariate in our analyses, but hesitate to use it to make assumptions about our independent tweet count variables.[10]

To determine the impact of information seen on Twitter on respondents' preferences we curate tweets in respondents' timeline on distinct topics that we measure their opinions on. And we aggregate those tweets based on the sources they come from. We chose to examine three key issues we felt to be relevant to the U.K. election: U.K. taxing/spending policy; the U.K.'s ties to EU; legal immigration to the U.K.; as well as the extent of ISIS' expansion To determine which tweets were politically relevant, we manually constructed short lists of terms related to our topics of interest. From these short lists of "anchor terms" we then identified which other terms most frequently co-occurred with the original terms. We then

[7]Note that for all users in our sample we have their self-report of what traditional media they follow.

[8]Twitter added a "while you were away" feature to highlight tweets that its algorithm predicts the user is likely to be interested in on January 21, 2015, but this represents a tiny fraction of the overall Twitter feed.

[9]This is actually a very tricky question unto itself, and undoubtedly there are data available that could help us do a better job of figuring out which tweets were more likely to be seen. For example, someone who only follows three people is certainly more likely to see all of their tweets than someone who follows 3,000. Similarly, holding constant the number of people being followed, someone who logs on hourly will see more tweets than someone who does monthly. Tweets during the day are probably more likely to be seen that in the middle of the night. While this remains an interesting question for future research, we think that at the individual level, taking the proportion of tweets in one's one feed on a given topic (or from a given ideological source) as a proxy for the proportion of tweets exposed to on that topic (from that ideological perspective) is reasonable as a first step.

[10]We re-did our main analysis restricted to the subset of respondents who claimed to use Twitter "Every few weeks" or more often; the results are not substantively changed.

use these expanded list of terms to determine to identify tweets related to each topics.

For example, our original search for "Ties to the EU" consisted of the terms "brexit" and "euro-skeptic"; this is not a comprehensive list of terms that could be related to the topic, but it is a list unlikely to produce many false positives. We calculated the absolute frequency of all words from all tweets, and separately, the frequency of all words in the subset of tweets s that contained either "brexit" or "euroskeptic." We then calculated a score for each word w in this subset:

$$Score_s^w = \frac{f_s^w N_s^w}{f^w}$$

Where f_s^w is the relative frequency of word w in subset s, f^w is the frequency of word w overall, and N_s^w is the count of word w in subset s. We then used the words with the top 25 highest scores to create the subset of tweets that we claimed to actually pertain to the topic "Ties to the EU." The list of the top 10 of these terms for "Ties to the EU", along with their scores can be seen in Table 3. "Brexit" seems to have been an excellent choice, whereas "euroskeptic" was fairly uncommon, and more appropriate terms expressing the same sentiment included "no2eu" and "betteroffout."[11]

[Table 3 Here]

We performed an additional categorization of relevant tweets based on the type of the account that created them: tweets from accounts associated with a politician or a political party (462 total accounts), and tweets from accounts associated with journalists or media outlets (987 total accounts). We further split the political accounts into those associated with each of the four major political parties under study. For media accounts, a research assistant identified the U.K.

[11]The advantage of this approach—as opposed to just coming up with our own longer list originally—is two-fold. First, it allows the data itself inform us about the correct terms to use in the list, which is especially valuable when using social media where language use is constantly evolving. In addition, the method is replicable: conditional on using the state start words, the algorithm always produces the same list of 25 most commonly co-occurring words. For a full list of the top 25 terms found for each issue topic, see Appendix A.

media organizations with the greatest number of Twitter accounts—including the accounts of journalists employed by those organizations—and we then divided them according to their ideological leanings. Major left-leaning media outlets are The Guardian and The Independent; right-leaning media outlets are The Times and The Sun; centrist media outlets are Scottish TV, the BBC, CNN and The Financial Times.

The number of political tweets from politicians and media sources in the timelines of our respondents ranged from 0 up to 370,000. To be included in this count, a tweet needed to be: (a) sent by one of the 462 political or 987 media accounts we identified, and (b) mention one of the topics or parties we study. Overall, 32 percent of respondents received 0 political tweets from either source, and 63 percent received 0 tweets from political accounts. The wide variation in this measure makes it useful as an explanatory variable. For those respondents who did receive at least one tweet from each source, Table 4 provides a summary of the distribution of the tweets across topics from that source.

[**Table 4 Here**]

The first column shows the number of respondents who received at least one tweet sent by a type of account about a topic of interest. Comparing the rows of this first column shows the relative "penetration" of each party/media type among our respondents: we see that Labour and the Conservatives, the two largest parties, have tweets that reach the most respondents, and that centrist media reaches the most respondents overall.

The other four columns summarize the distribution of tweets received by the respondents identified in the first column. The first row, for example, looks at all of the tweets by Labour and breaks them down by topic. Among those 532 people who received at least one tweet from Labour, the mean percentage of the Labour tweets about economic issues in their timeline was 49%.

Comparing the rows, there is a marked difference in the relative emphases placed on the four topics by each source. For example, nearly half of tweets sent by Labour or the Tories were about the economy, while UKIP tweeted about the economy much less than about immigration or the EU, providing face validity of our coding strategy. There is less variation within the media accounts, although the Left Media tended to avoid discussing immigration. On average, media accounts were more likely to tweet about ISIS than were the parties.

Results

Party Placements

The first outcome of interest is the change in the ability of the respondents to correctly rank the four major parties (Liberal Democrats, Labour, Conservatives, UKIP) on a 100-point left-right scale on three major issues in the 2015 election:

- U.K. taxing versus spending policy: far left = Social spending should be **increased** even if that means higher taxes; far right = Taxes should be **cut** even if that means lower social spending
- The degree of U.K.'s ties to the EU: far left = Britain should **develop stronger ties** with the European Union; far right = Britain should **leave** the European Union
- The level of legal immigration: far left = Legal immigration to Britain should **increase** a lot; far left = Legal immigration to Britain should **decrease** a lot

In each wave of the survey, we asked respondents to place themselves and each of the 4 parties on a 0 (leftmost) to 100 (rightmost) scale.[12]

[12] In wave 2 we asked these questions to half of the respondents, and in wave 3 we asked them of the other half, because of length constraints in the survey. This means that we cannot compare results from wave 2 to wave 3, and in practice, we find that there is too little power to use the results from waves 2 and 3 in our analysis.

One of the challenges in analysis of this sort is establishing a "ground truth" of where the parties actually stand (Tucker and Markowski, 2007). There are a wide variety potential measures of this ground truth, and we tested many of them, including: the mean of all the respondents' placements of the parties; the mean of the placements by respondents with a college degree; the mean of the party placements made by self-identified supporters of each party; and the mean of the self-placements of self-identified supporters of each party.

All of these placement estimates were highly correlated with each other at .93 or higher, and we use the simplest measure—the mean of the placement by all respondents—as our "ground truth."[13] As a further reality check, we compared these placements against the party placements in the 2014 edition of the Chapel Hill Expert Survey (Bakker et al., 2015). Every wave of our placements correlated with the CHES estimates at at least .95. The highest correlation was with wave 1, the soonest after the 2014 CHES was conducted, suggesting that differences in later waves could be due to actual movements of the parties; note, though, that none of the estimated party movements are statistically significant.

Figure 1 gives the mean placement of respondents, and of each of the 4 parties, on each of the three issues we looked at: the U.K.'s relationship to the EU; the tradeoff between taxes and spending; and levels of immigration. Placement is given both in wave 1 and wave 4 of the survey. We see tremendous stability for the mean placements. The Liberal Democrats were perceived to move right on the U.K.'s relationship to the EU, as was the Conservative Party. On spending UKIP was perceived to move substantially to the left, and the Tories a small amount to the right. On the issue of immigration we saw the most movement. The Liberal Democrats, Labour, and the Conservative Party were all perceived to move to the right over the course of the campaign.

[**Figure 1**]

[13] Among other advantages, this approach allows for tracking the movement of the parties during the campaign. Notably, the Liberal Democrats moved to the right on the issue of the EU, and all of the parties except UKIP moved to the right on immigration.

In order to determine if each of our individual respondents correctly placed the parties in each wave, we compared their placement to the the mean values of the parties as shown in Figure 1. However, for the instances in which two parties were close together (within 10 points on the 100 point scale), we allowed some leeway: in such instances we accepted either ordering of the two parties as correct. The percentage of respondents identifying the correct orderings among Twitter users and non-Twitter users can be seen in Tables 5 and 6. Note that the correct ordering for the parties on each issue was the same in both waves for the immigration and spending issues, but not for the topic of the EU: the Liberal Democrats moved to the right, making their position similar to that of Labour. As a result, we coded the respondent's ranking as "correct" if they placed Labour to the left of the Liberal Democrats or vice versa. This meant that the EU question got "easier," hence the high percentage who got the question wrong in wave 1 but right in wave 4.[14] Overall, ranking the parties on spending was the most difficult task, with only 55 percent of respondents in wave 4 doing so correctly among those who attempted to answer it in both waves; and the number of respondents who were able to give any answer to this question was considerably smaller than for the other questions.

[Tables 5 and 6 Here]

To test our first hypothesis, that exposure to information about a political topic on Twitter will increase knowledge about that topic we estimate a logit model of the respondent's ability to correctly place the parties in wave 4, including the log of the number of tweets on the topic as our key explanatory variable, and conditioning on their placement in wave 1, as well as a set of variables capturing respondents' characteristics that could make them likely to learn about the correct party placement between waves of the survey independent of information seen on Twitter.[15] We condition on standard demographic variables (gender, age, class,

[14]This convergence makes interpreting the "improvement" in ranking the parties on this issue difficult—if someone were to entirely ignore political news for eight months and rank Labour to the left of the Liberal Democrats in both wave 1 and wave 4, our coding strategy considers their political knowledge to have increased. There is no easy solution to this problem, but it should be kept in mind when considering the results.

[15]Throughout the analysis we use the log of (1.0001 plus) the number of tweets in a respondent's timeline because of the highly skewed nature of the distribution of tweets.

education, race, marital status, and religiosity), as well as measures of exposure to news. We include variables for self reported frequency of watching 'Newsnight' (a long-running news program) and of using the internet, as well as dummy variables capturing self-reported measures of which print media respondents read. [16] The dependent variable we use in each of three models is a binary dummy variable for whether or not the respondent correctly ranked the four parties on a specific issue (we thus estimate each model with logit). We report 95% confidence intervals around the relevant coefficients in Figure 2.[17] We see in Figure 2 that all three of the effects are positive, and that 2 (EU and spending) are significant at $p < .05$, while the effect on ranking the parties on immigration is just shy of significant at $p < .10$. These results support our hypothesis that respondents learn about the issue position of parties from receiving information on Twitter.

[Figure 2 here]

To get a sense of the magnitude of the effect sizes and the distribution of the independent variables, Figure 3 plots the distribution of relevant tweets on the x-axis against the predicted probability that the respondent correctly ranked the parties on that topic in wave 4. We sets all other independent variables to their mean values. The general effect is positive, although decreasing density of the tweet count variable on the upper end of the distribution means that at no point do the 95% confidence intervals fail to overlap. The slope of this effect is steepest for ranking the parties on spending: if the typical respondent had received 2 standard deviations more tweets (343 more tweets) about spending than the median number, their predicted probability of ranking the parties correctly increases from .56 to .61. This may seem like a small effect, but it is based on tweets from different sources. We show below that the impact of tweets varies by source, thus making this a conservative estimate for the potential impact of tweets.

[Figure 3 here]

[16] We use the British 5-category system for measuring class. We include specific dummies for whether or not respondents report reading "Blue Tops" , "Red Tops", or "Broadsheets" (The Guardian or The Telegraph).

[17] Full model estimates are reported in Appendix C.

To test H_2, that exposure to information on a topic sent from a party will increase knowledge of the parties relative positions on an issue we aggregate tweets over topic separately for: a) tweets by political parties (or related politicians); and b) tweets by media organizations (or affiliated journalists). We then estimate logit models of the probability of correctly placing the parties based on the logged number of tweets from each source on the topic, including identical sets of covariates as used in our previous model of party placement used to generate Figure 2. The results of the disaggregated model are shown in Figure 4, and demonstrate that the results in Figure 2 are driven by tweets from the parties, as these independently have a positive and significant effect on the topics of spending and the EU (but not immigration), while tweets from the media do not have a statistically significant effect on correctly ranking the parties. These findings suggest that tweets sent by the parties do help to inform the voters about the positions of the parties on issues.

[**Figure 4 here**]

Factual Knowledge

We also operationalize political knowledge through factual questions of politically relevant topics. In waves 2 and 3 of the survey, we asked three multiple choice questions (correct answers in **bold**):

- (ISIS) The Islamic militant group known as ISIS currently controls territory in which of these countries: **Syria**, Kuwait, Morocco, or Pakistan?[18]
- (Unemployment) Compared to a year ago, has unemployment in Great Britain increased, **decreased**, or stayed the same?
- (Immigration) Over the past 5 years, has the number of immigrants to the United Kingdom from other EU countries been: Less than 100,000 per year,

[18] In the Wave 2 version of this question, "Morocco" was "Egypt," and we made the switch because there some news reports of ISIS activity in Egypt after Wave 2.

Between 100,000 and 300,000 per year, Between 300,000 and 500,000 per year, More than 500,000 per year?

Table 7 reports the joint distribution of right and wrong responses for each question in waves 2 and 3. Panel A restricts the sample to those people who use Twitter at least "Every few weeks," while Panel B only includes respondents who use Twitter "Less often" or "Never." We can see that the question about ISIS revealed very little change in respondents' beliefs. However, on both Unemployment and Immigration substantial numbers of respondents were on the off-diagonals (changing their response from wave 2 to wave 3). For instance, among Twitter users, 30% of respondents gave an incorrect answer on Immigration in wave 2, but a correct answer in wave 3. And there was also 'unlearning' on this issue: 19% of respondents gave a correct answer in wave 2, but an incorrect answer in wave 3. Overall, the rate of *unlearning* the correct answer is similar to the rate of learning the correct answer—compare the bottom left and top right cell of each 2 by 2 box: about as many people got the question right in wave 2 but wrong in wave 3 as vice versa. This allows us to explore both the types of tweets that inform, and the types of tweets that confuse.[19]

[Table 7 Here]

Our analysis here uses the same specification we used for correct placement of the parties. We first test H_1, that exposure to political information on Twitter increases knowledge, by running 3 logistic regressions where the dependent variable is whether the respondent correctly answered that question in wave 3 and the key explanatory variable is the total number of tweets related to that topic that appeared in their feed between wave 2 and wave 3. Again, we control for demographics, media use and whether they correctly answered the question in wave 2. Figure 5 plots three horizontal lines that represent the logit coefficient (with

[19]Because these questions are multiple choice, it was possible to guess the right answer, and thus some of this difference is the result of random noise. However, respondents were able to select "Don't Know" instead of answering the question, so the rate of true guessing should be low.

standard errors) of the measures of topic-related tweets appearing in the feed.[20] [21]

[**Figure 5 Here**]

We see a positive and significant relationship between the number of tweets and respondents' knowledge of immigration, a positive but not statistically significant effect for number of tweets seen about ISIS, and an estimate almost indistinguishable from zero for number of tweets seen about unemployment. We note that the imprecision of the estimate on ISIS is to be expected given how few tweets there were on the topic (see Table 4), and how little variation there is on the dependent variable: 88% of respondents got the question correct in both wave 2 and wave 3, so there is little change in knowledge to be explained. More generally, we believe that the lack of effects here could be because we are aggregating tweets intended to persuade respondents of different facts: tweets sent from UKIP may be meant to convince respondents that immigration was high, tweets sent from Labor may be intended to convince respondents that immigration was low.

To test for distinct effects of tweets by different sources, we again disaggregate these tweets by their source (media or parties), and estimate logit models of correct responses on the factual questions with both explanatory variables – tweets by media and tweets by parties – included, as well as identical control variables and the previous wave's response to the knowledge question, to test the hypothesis (H_3) that only media tweets, not political party tweets, will increase factual knowledge. The results are plotted in Figure 6, and agree with our expectations.[22] For all three topics, the effect of media tweets (but not political party tweets) is positive and significant. In fact, on the question of unemployment, the effect of tweets from political parties is *negative* and significant, a finding we explore more below.

[**Figure 6 Here**]

We expect that some political parties have strategic incentives to emphasize different aspects of the same topic, or even to mislead voters on certain topics, and that these divergent emphases might have contrasting effects on change in

[20] For complete estimates from the models, see Appendix C.

[21] For space reasons, we do not include effect-size plots for each of these regressions in the body of the text, but see Appendix B.

[22] Complete estimates are presented in Appendix C.

knowledge (H_4). When discussing the economy, for example, the incumbent parties (the Conservatives and Liberal Democrats) should want to emphasize the fact that the U.K. unemployment rate was decreasing, while the opposition parties (especially Labour, as UKIP was more focused on non-economic issues) should criticize other aspects of the economy and make people less likely to believe the truth that unemployment had been going down. Also, fears over immigration were central to UKIP's platform, so they were likely to discuss immigration more often and in a more inflammatory fashion, to make it seem that the number of immigrants was high.

To test these ideas, we further disaggregate the political party tweets by party (Labour, Conservatives, Liberal Democrats and UKIP) and the media tweets by ideological leanings of the media sources (Left, Center and Right). We again estimate models of providing the correct answer in wave 3 as a function of the number of tweets seen about each topic from each source (thus there are 7 tweet count variables on the right hand side of each model), as well as the answer in the previous wave and our set of control variables.[23] We report estimates of these models for unemployment and immigration, but not for factual knowledge about ISIS. We have no theoretical expectation about which source should have the largest impact on knowledge of ISIS, and there are so few tweets about ISIS that we can not precisely estimate the effects of these tweets.

The results for the effects of tweets on knowledge of unemployment are shown in Figure 7, and are are largely in accord with our expectations. Tweets from Labour lead to a *negative* and significant change in knowledge about the unemployment rate, while tweets from the Conservatives have a positive effect, although it falls just shy of significance at $p < .1$. Tweets from Liberal Democrats may have a slightly negative effect, contrary to our expectations of a positive effect - but the estimate is not significant at traditional levels. And as the Liberal Democrats suffered greatly from their alliance with the Conservatives, and actually tried to distance themselves from the coalition government, they may have simply not emphasized a clear position on this.

[23] Complete estimates for the models are reported in Appendix C.

[Figure 7 Here]

Turning to immigration (Figure 8), we fail to find the expected positive effect of UKIP tweets on knowledge of immigration. The only party's tweets to lead to an increase in knowledge of immigration is Labour, and this association is in fact positive. This finding fails to support H_4. However, rather than informing respondents of the true value of the number of immigrants each year, UKIP may have been exaggerating the number.

[Figure 8 Here]

To examine these results further, we take advantage of the fact that the multiple choice questions used to measure factual political knowledge about immigration and unemployment had ordinal choices. Instead of merely analyzing changes in *correctness*, we can look at the *direction* of those changes. We fit an ordered probit model where the dependent variable is the difference in the respondent's answer across waves to the relevant multiple choice question: an indication of whether, and how much, they think the value in question increased across waves.[24]

Table 8 displays the results of these ordered probit estimates. We condition on the same suite of demographic and media use controls as in previous analyses. The results of column 1 give context to the evidence from Figure 7: tweets from Labour increase estimates of the change in the unemployment rate, and tweets from Conservatives and right-leaning media decrease those estimates. The fact of the matter is that unemployment had been decreasing. Because "Decreased" was the lowest possible response (lower than "Stayed the same" or "Increased"), this implies that Labour's tweets were associated with less accuracy and the Conservatives' with greater accuracy. This further supports H_4.

[Table 8 Here]

[24]For example, this dependent variable takes a value of 2 if the respondent's answer to the question about immigration went from "Between 100,000 and 300,000 [immigrants] per year" (the second-lowest category) to "More than 500,000 [immigrants] per year" (the highest category). If the respondent instead changed from "Between 100,000 and 300,000 [immigrants] per year" to "Less than 100,000 [immigrants] per year," the dependent variable takes a value of -1.

Column 2 of Table 8 explains the null results from the analysis of immigration knowledge in Figure 8. Labour tweets are significantly associated with a decreased estimate of the rate of immigration, while UKIP tweets are significantly associated with an increased estimate. This is precisely what we would expect, based on the strategic frames most useful to these parties and especially to UKIP. The reason that these changes did not necessarily reflect an increased chance of correctly answering the question is that the correct answer ("Between 100,000 and 300,000 [immigrants] per year") was the second lowest choice. UKIP's tweets caused some respondents to correctly raise their estimates from the lowest to the second lowest choice, but caused others to incorrectly raise their estimates beyond the second lowest choice.

We claimed that our fourth hypothesis: that exposure to tweets sent by political parties about issues will change beliefs about facts in directions strategically advantageous to the party transmitting the message had 4 possible empirical implications here. Our findings support three of those, and the only one that fails to find support is that tweets by UKIP would increase respondents' knowledge of Immigration, which was negated by support for our alternative hypothesis that tweets by UKIP would simply make respondents think immigration rates were higher. We thus find substantial support for the theory that parties strategically discuss issues in such a way to encourage their followers to hold factual beliefs that are advantageous for those parties.

We thus find that the effect of exposure to tweets is contingent on the source of the tweets and the topic of the tweets. Exposure to political information sent by parties tends to increase knowledge of party platforms but not of factual knowledge, while the inverse is true for information sent by media accounts. Political parties do, however, tend to affect factual knowledge of politicized issues in strategically coherent ways.

Causality

In the above discussion, we assume that the observed empirical relationships between exposure to tweets about a certain topic and changes in knowledge of that topic are causally related. This is intuitively plausible, we are measuring knowledge at two different times, and attributing the *change* in knowledge to exposure to tweets seen *between* the two measurements. But we lack a crucial element to conclusively treat our measurement as the causal effect of exposure to tweets on changes in knowledge: our subjects self-select into which Twitter accounts they follow, so the "treatment" is not randomly assigned. We acknowledge the potential problems this poses for causal inference, but we argue that the quantity of interest could not in fact be estimated through random assignment.

The primary objection to our claim that we are measuring the causal effect of exposure to tweets on changes in knowledge is omitted variable bias. For example, consider our analysis of what people know about unemployment. We find in Table 8 that exposure to tweets by Labour causes subjects to increase their estimate of the unemployment rate. If we are concerned about omitted variable bias, we need to maintain that interest in unemployment (or some other omitted variable) causes subjects to self-select into following Labour politicians *and* causes them to raise their estimate of the unemployment rate *between* waves of the survey. We can not rule this out with the data available. But we note that we are conditioning on the other media sources that respondents follow, as well as a host of other variables that should be related to information they are likely to seek out.

The alternative explanation, and the one we advance in this paper, is that tweets by Labour (the out-party) about unemployment are designed to downplay the successes of the incumbent party, and that exposure to these tweets *causes* treated subjects to increase their estimate of the unemployment rate. And, we note the meaning of the estimates here. We are conditioning on a large set of characteristics that would make a person likely to follow Labour. One would expect the set of people contained here to follow Labour based on some characteristic related to interest in the unemployment rate to be quite low, and unlikely to account for the observed estimate.

We do not have a randomized experiment or any other source of fully exogenous variation in our "treatment" condition. But we believe that the evidence we have amassed should increase the reader's credence in the causal effect of exposure to political information on Twitter in the various forms we discuss in the paper.

Furthermore, it is important to consider the quantity that randomized exposure to tweets would actually estimate. Seeing a large number of tweets from a source the subject did not chose to follow might cause changes in beliefs, or it might not. But the large-scale question we aim to answer with this project is, "What is the effect of Twitter use on political knowledge?" The randomized experiment described above is not "Twitter use" per se, but something else, something artificial that does not happen in the real world.[25]

Although we acknowledge the limitations of our research design for causal inference, we think that our empirical findings show what effect social media might be having, and clearly show the difference in change in respondents' beliefs over the course of a campaign for people who follow parties and news media on social media versus people who do not.

Conclusion

In this first-ever analysis combining the content of individuals' social-media feeds and panel survey data over the course of an election campaign, we find evidence consistent with the claim that exposure to politics on social media may lead to a more politically informed mass public. These findings contribute to our cumulative understanding of the extent to which election campaigns inform voters. They also yield specific insights on how political knowledge is affected by social media, which is arguably the most important development in political communication in our

[25]In a recent paper, Leeper (2016) demonstrates the limitations of a randomized trail in estimating this quantity of interest. Calculating the average treatment effect of media exposure over an entire sample can mask significantly heterogeneous treatment effects.

times.

Knowledge of politically relevant facts is an important component of the democratic process, and exposure to information is a precondition for knowledge (Lupia, 2015). However, previous work in this area has been hampered by the twin challenges of measuring the key independent variable of information exposure and cleanly identifying its effects on political knowledge. Our approach overcomes the first challenge by matching survey respondents to measures of their actual exposure to political information on social media. This allows us to avoid the biased media exposure self-reports that often plague this independent variable. It also allows us to conduct individual-level analyses rather than relying on measures of aggregate media exposure. We address the second challenge with a multi-wave panel survey design that permits the estimation of what causes changes in knowledge controlling for stable individual-level characteristics of respondents. The result is a rare opportunity to examine real-world evidence showing that exposure to news media generally improves political knowledge, while messages from parties and politicians have less uniformly positive effects.

With regard to the particular effects of social media on knowledge, our findings allay some of the concerns expressed by those who are pessimistic about the implications of this development for politics. Our findings about the types of knowledge that are subject to social media effects largely concords with previous findings (Barabas et al., 2014). Contrary to the worst fears of some, on balance social media users became more informed about politics during the 2015 U.K. general election campaign. Messages from news media improved recipients' knowledge of relevant political facts; messages from the parties improved knowledge of their relative stances on the campaign's most important issues.

But we also uncover results that should temper any unbridled enthusiasm about the impact of social media on political knowledge; namely that exposure to partisan messages about highly salient issues over the course of a campaign can cause knowledge polarization on those issues. As electorates in Western democracies become increasingly divided on the issues of globalization and open borders, the dual effects of exposure to UKIP's tweets on immigration that we discover are worth particular mention. On the one hand, these messages did not harm aggregate levels of knowledge regarding the number of immigrants to the United Kingdom. But on the other, they led UKIP followers to revise their estimates upward, which in turn aggravated disagreement in the overall electorate about a highly salient, objective political fact. We note that this phenomenon could only be discovered with a research design like ours that measures the source, content and recipient of individual political messages. Knowledge polarization thus may possibly be a (yet-to-be discovered) effect of political messages transmitted via traditional media as well.

Ever-growing numbers of people around the world are turning to social media to get information about politics and public affairs, and our findings demonstrate that this phenomenon is worthy of study on its own terms. In addition, our research indicates that new insights about classic questions in political communication research await discovery by those who take advantage of the unparalleled opportunities provided by social media to precisely measure and assess the effects of political messages on political knowledge.

References

Andersen, Robert, James Tilley and Anthony F Heath. 2005. "Political knowledge and enlightened preferences: party choice through the electoral cycle." *British Journal of Political Science* 35(02):285–302.

Bakker, Ryan, Erica Edwards, Liesbet Hooghe, Seth Jolly, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen and Milada Vachudova. 2015. "2014 Chapel Hill Expert Survey." 2015.1.

Bakshy, Eytan, Solomon Messing and Lada A Adamic. 2015. "Exposure to ideologically diverse news and opinion on Facebook." *Science* 348(6239):1130–1132.

Banducci, Susan, Heiko Giebler and Sylvia Kritzinger. 2015. "Knowing more from less: How the information environment increases knowledge of party positions." *British Journal of Political Science* pp. 1–18.

Barabas, Jason and Jennifer Jerit. 2009. "Estimating the Causal Effects of Media Coverage on Policy-Specific Knowledge." *American Journal of Political Science* 53(1):73–89.

Barabas, Jason, Jennifer Jerit, William Pollock and Carlisle Rainey. 2014. "The question (s) of political knowledge." *American Political Science Review* 108(04):840–855.

Barberá, Pablo, John T Jost, Jonathan Nagler, Joshua A Tucker and Richard Bonneau. 2015. "Tweeting From Left to Right Is Online Political Communication More Than an Echo Chamber?" *Psychological science* .

Bartels, Larry M. 1996. "Uninformed votes: Information effects in presidential elections." *American Journal of Political Science* pp. 194–230.

Bartels, Larry M. 2000. "Partisanship and voting behavior, 1952-1996." *American Journal of Political Science* pp. 35–50.

Bartels, Larry M. 2002. "Beyond the running tally: Partisan bias in political perceptions." *Political Behavior* 24(2):117–150.

Bullock, John G, Alan S Gerber, Seth J Hill and Gregory A Huber. 2015. "Partisan Bias in Factual Beliefs about Politics." *Quarterly Journal of Political Science* 10:519–578.

Carpini, Michael X Delli and Scott Keeter. 1997. *What Americans know about politics and why it matters.* Yale University Press.

Chang, Linchiat and Jon A Krosnick. 2009. "National surveys via RDD telephone interviewing versus the internet comparing sample representativeness and response quality." *Public Opinion Quarterly* 73(4):641–678.

Conover, Michael D, Bruno Gonçalves, Alessandro Flammini and Filippo Menczer. 2012. "Partisan asymmetries in online political activity." *EPJ Data Science* 1(1):1–19.

Curran, James, Shanto Iyengar, Anker Brink Lund and Inka Salovaara-Moring. 2009. "Media System, Public Knowledge and Democracy A Comparative Study." *European Journal of Communication* 24(1):5–26.

DellaVigna, Stefano and Ethan Kaplan. 2006. The Fox News effect: Media bias and voting. Technical report National Bureau of Economic Research.

Evans, Geoffrey and Jon Mellon. 2015. "Working Class Votes and Conservative Losses: Solving the UKIP Puzzle." *Parliamentary Affairs* p. gsv005.

Fieldhouse, E., J. Green, G. Evans, H. Schmitt and C. van der Eijk. 2015. "British Election Study Internet Panel Wave 6.".

Flaxman, Seth, Sharad Goel and Justin M Rao. 2013. "Ideological segregation and the effects of social media on news consumption." *Available at SSRN 2363701* .

Fowler, Anthony and Michele Margolis. 2014. "The political consequences of uninformed voters." *Electoral Studies* 34:100–110.

Gentzkow, Matthew. 2006. "Television and voter turnout." *The Quarterly Journal of Economics* pp. 931–972.

Gentzkow, Matthew, Jesse M Shapiro and Michael Sinkinson. 2014. “Competition and ideological diversity: Historical evidence from us newspapers.” *The American Economic Review* 104(10):3073–3114.

Gilens, Martin, Lynn Vavreck and Martin Cohen. 2007. “The mass media and the public's assessments of presidential candidates, 1952–2000.” *Journal of Politics* 69(4):1160–1175.

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

Gottfried, Jeffrey, Michael Barthel, Elisa Shearer and Amy Mitchell. 2016. “The 2016 Presidential Campaign–a News Event Thats Hard to Miss.” *Washington, DC: Pew Research Center. Available at¡ http://www. journalism. org/files/2016/02/PJ_2016* 2.

Greenwood, Shannon, Andrew Perrin and Maeve Duggan. 2016. “Social media update 2016.” *Pew Research Center* 9.

Guess, Andrew M. 2016. “Media Choice and Moderation: Evidence from Online Tracking Data.” *Working Paper* .

Hope, Christopher. 2015. “And they're off: the 2015 general election campaign officially starts this Friday.” *Telegraph UK* .

Huber, Gregory A and Kevin Arceneaux. 2007. “Identifying the persuasive effects of presidential advertising.” *American Journal of Political Science* 51(4):957–977.

Iyengar, Shanto and Kyu S Hahn. 2009. “Red media, blue media: Evidence of ideological selectivity in media use.” *Journal of Communication* 59(1):19–39.

Iyengar, Shanto, Kyu S Hahn, Heinz Bonfadelli and Mirko Marr. 2009. ““Dark Areas of Ignorance” Revisited Comparing International Affairs Knowledge in Switzerland and the United States.” *Communication Research* 36(3):341–358.

Johnston, Richard, Michael G Hagen and Kathleen Hall Jamieson. 2004. *The 2000 presidential election and the foundations of party politics*. Cambridge University Press.

Kelley, Stanley. 1960. *Political campaigning: Problems in creating an informed electorate*. Brookings Institution.

Koch, Jeffrey W. 2008. "Campaign advertisements' impact on voter certainty and knowledge of house candidates' ideological positions." *Political Research Quarterly* 61(4):609–621.

Lau, Richard R and David P Redlawsk. 2001. "Advantages and disadvantages of cognitive heuristics in political decision making." *American Journal of Political Science* pp. 951–971.

Lau, Richard R, Lee Sigelman and Ivy Brown Rovner. 2007. "The effects of negative political campaigns: a meta-analytic reassessment." *Journal of Politics* 69(4):1176–1209.

Lauderdale, Ben. 2015. "What We Got Wrong In Our 2015 U.K. General Election Model." *fivethirtyeight.com* .

Leeper, Thomas J. 2016. "How does treatment self-selection affect inferences about political communication?" *Journal of Experimental Political Science* .

Lupia, Arthur. 1994. "Shortcuts versus encyclopedias: information and voting behavior in California insurance reform elections." *American Political Science Review* 88(01):63–76.

Lupia, Arthur. 2015. *Uninformed: Why people seem to know so little about politics and what we can do about it*. Oxford University Press.

Martin, Gregory J and Ali Yurukoglu. 2014. Bias in cable news: Real effects and polarization. Technical report National Bureau of Economic Research.

Milazzo, Caitlin. 2015. "Getting it Right When it Counts: Constituency Marginality and Voter Perceptions of British Parties' Policy Positions." *Journal of Elections, Public Opinion & Parties* 25(2):111–136.

Prior, Markus. 2005. "News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout." *American Journal of Political Science* 49(3):577–592.

Prior, Markus. 2007. *Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections.* Cambridge University Press.

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

Prior, Markus. 2013. "The challenge of measuring media exposure: Reply to Dilliplane, Goldman, and Mutz." *Political Communication* 30(4):620–634.

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

Strömberg, David. 2004. "Mass media competition, political competition, and public policy." *The Review of Economic Studies* 71(1):265–284.

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

Tillman, Erik R. 2012. "Support for the euro, political knowledge, and voting behavior in the 2001 and 2005 UK general elections." *European Union Politics* 13(3):367–389.

Tucker, Joshua A and Radoslaw Markowski. 2007. Subjective vs. Objective Proximity in Poland: New Directions for the Empirical Study of Political Representation. 2007 Annual Meeting of the American Political Science Association.

Vavreck, Lynn. 2009. *The Message Matters: The Economy and Presidential Campaigns.* Princeton, New Jersey: Princeton University Press.

Weaver, Matthew. 2015. "Nick Clegg's tuition fees 'debacle' undermined trust, says Norman Lamb." *The Guardian* May 12.

Whiteley, Paul, Harold D Clarke, David Sanders and Marianne C Stewart. 2013. *Affluence, Austerity and Electoral Change in Britain.* Cambridge University Press.

Table 1: Number of Survey Respondents per Wave

	Wave 1	Wave 2	Wave 3	Wave 4	All Waves
SMU respondents[a]	2,574	2,507	2,776	2,490	3,846
Retention, previous wave[b]		68%	79%	90%	1,308 (in all 4 waves)
New respondents		32%	19%	0%	

[a]Cell entries are the number of respondents in each wave.

[b]Cell entries are the proportion of respondents returning from the previous wave. Wave 1 concluded on July 31, 2014; wave 2 on December 11; wave 3 on March 30, 2015; and wave 4 (post-election) on June 17.

Table 2: Descriptive Statistics of Relevant Populations

Panel A: Demographic Characteristics

	SMU	SMU Plus	NR (no Twitter)	BES	2011 Census
Women	45%	43%	53%	50%	49%
15+ Years Education	52%	55%	36%	41%	27%
Median Age	48	48	55	53	40
Median Household Income	£34,200	£37,500	£30,000	£27,500	£21,000
Median Ideology†	5.2	5.2	5.0	4.6	

† Self-reported ideology, left to right; asked on a 0-100 scale in our survey and on a 0-10 scale in the BES. The BES is a nationally representative post-election survey of 30,000 voters.

Panel B: Vote Choice, Post-Election

	SMU	SMU Plus	NR (no Twitter)	Election
Conservative	33	32	44	37
Labour	34	35	31	31
Liberal Democrats	8	9	7	8
SNP	5	5	5	5
UKIP	9	8	10	13
Green	10	11	2	4
Other	1	1	0	3
	100%	100%	100%	100%

Panel C: U.K. Country

	SMU	SMU Plus	NR (no Twitter)	Census
England	84	85	85	84
Scotland	5	5	9	8
Wales	9	9	6	5
Northern Ireland	1	1	0	3
	100%	100%	100%	100%

The demographic, vote choice and geographic vote share of the relevant populations: the Social Media Users sample and the SMU Plus sample (the subgroup who shared their Twitter timeline), and the group of control users taken from the Nationally Representative (NR) sample who did not use Twitter.

Table 3: Top Terms Pertaining to the Topic “Ties to the EU”

Term	Score
brexit	1000
no2eu	44
betteroffout	18
eureferendum	6.7
eu	6.7
euref	5.9
grexit	2.2
scoxit	1.5
stayineu	1.3
flexcit	1.3

Examples of the terms we found to tend to co-occur with our anchor terms for the topic “Ties to the EU.” We used this process to find terms that identify a tweet as pertaining to a topic of interest.

Table 4: Distribution of Tweets On Each Topic Received by Followers of Each Group

	ISIS	EU	Economy	Immigration
Labour (532 respondents)	3%	15%	**49%**	34%
Tory (472 respondents)	3%	25%	**45%**	27%
LibDem (224 respondents)	1%	29%	42%	28%
UKIP (102 respondents)	1%	**36%**	19%	**44%**
Right Media (184 respondents)	4%	25%	38%	33%
Centrist Media (763 respondents)	6%	26%	35%	33%
Left Media (161 respondents)	6%	33%	35%	25%

Cell entries are the mean percentage of tweets about the column topic from the sender listed in the row, out of all tweets received by a respondent about the four topics sent by the sender listed in the row (thus sum to 100 across rows). For example, the bottom right corner says that, among the 161 respondents who received at least one tweet sent by Left Media, the mean percentage of tweets about immigration—among the tweets sent by Centrist Media about one of the four topics under study—in their timelines is 25%. Cells bolded for emphasis.

Table 5: Placement of Parties in Waves 1 and 4 Among Twitter Users

	EU, N= 1,035	
	Correct W1	Incorrect W1
Correct W4	54%	27%
Incorrect W4	4%	15%
Total W1	58%	42%
	Immigration, N= 1,013	
	Correct W1	Incorrect W1
Correct W4	64%	14%
Incorrect W4	10%	12%
Total W1	74%	26%
	Spending, N= 798	
	Correct W1	Incorrect W1
Correct W4	38%	18%
Incorrect W4	16%	27%
Total W1	54%	45%

Cell entries are percentages for each possible combination of correct and incorrect answers across wave 1 and wave 4 of the party placement questions: (C,C), (C,I), (I,C), (I,I). The bottom line shows how difficult each question was showing the percentage correct in wave 1.

Table 6: Placement of Parties in Waves 1 and 4 Among Non-Twitter Users

	EU, N= 471	
	Correct W1	Incorrect W1
Correct W4	39%	29%
Incorrect W4	8%	24%
Total W1	47%	53%
	Immigration, N= 455	
	Correct W1	Incorrect W1
Correct W4	49%	17%
Incorrect W4	11%	22%
Total W1	60%	39%
	Spending, N= 343	
	Correct W1	Incorrect W1
Correct W4	25%	20%
Incorrect W4	16%	38%
Total W1	41%	58%

Cell entries are percentages for each possible combination of correct and incorrect answers across wave 1 and wave 4 of the party placement questions: (C,C), (C,I), (I,C), (I,I). The bottom line shows how difficult each question was showing the percentage correct in wave 1.

Table 7: Factual Knowledge: Comparing Twitter Users and Non-Twitter Users

Panel A: Factual Knowledge Among Twitter Users (N=1,325)

	ISIS		Unemployment		Immigration	
	Correct W2	Incorrect W2	Correct W2	Incorrect W2	Correct W2	Incorrect W2
Correct W3	89%	6%	53%	9%	33%	18%
Incorrect W3	2%	3%	13%	24%	19%	30%
Total W2	91%	9%	66%	33%	52%	48%

Panel B: Factual Knowledge Among Non-Twitter Users (N=1,076)

	ISIS		Unemployment		Immigration	
	Correct W2	Incorrect W2	Correct W2	Incorrect W2	Correct W2	Incorrect W2
Correct W3	85%	6%	50%	12%	33%	14%
Incorrect W3	5%	5%	11%	27%	24%	30%
Total W2	90%	11%	61%	39%	57%	44%

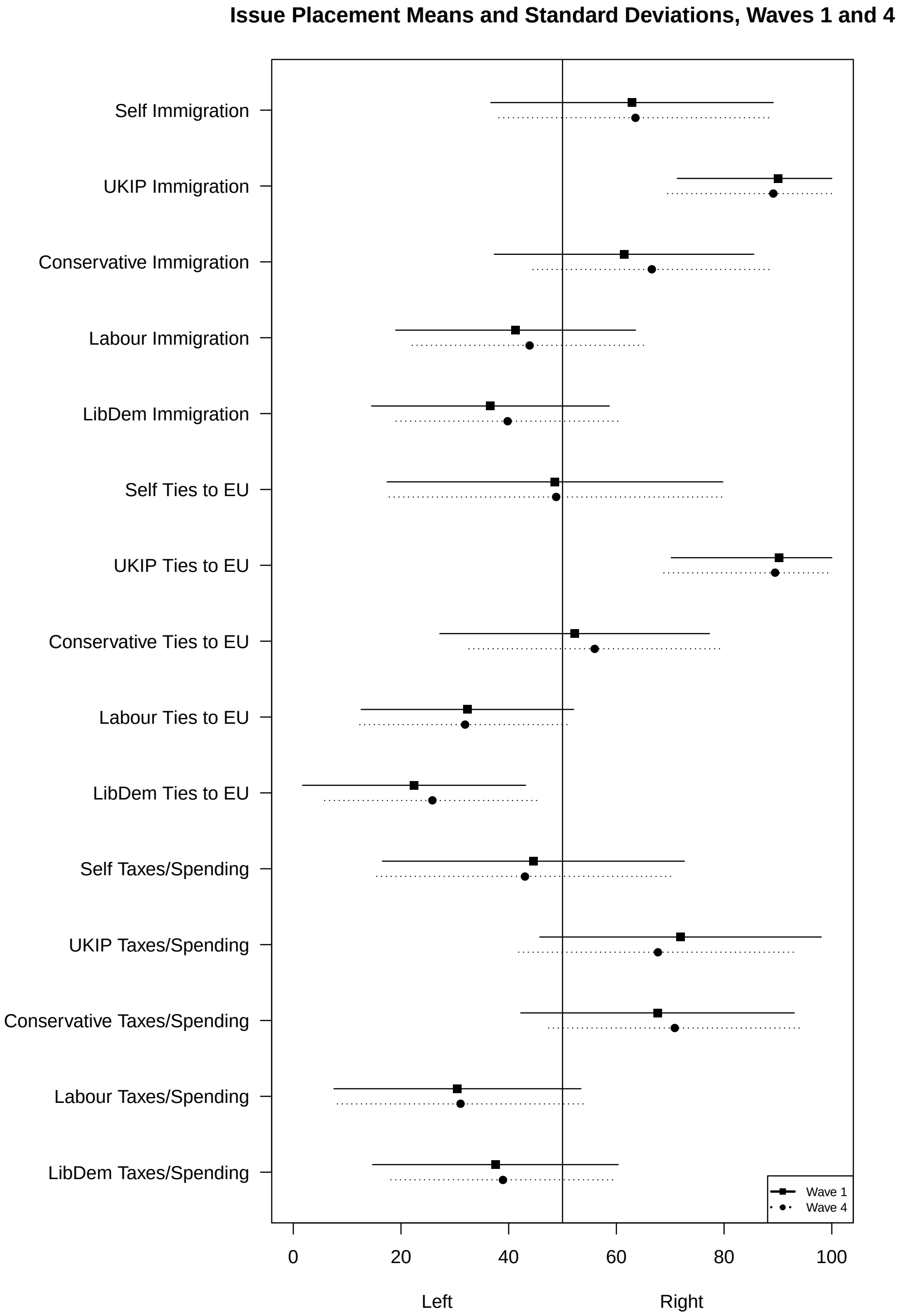
Distribution of Responses to Knowledge Questions: Cell entries are percentages for each possible combination of correct and incorrect answers across wave 2 and wave 3 of the politically relevant knowledge questions: (C,C), (C,I), (I,C), (I,I). The bottom row shows how difficult each question was showing the percentage correct in wave 2.

Table 8: Effect of Tweets on Estimates of Perceived Absolute Levels of Unemployment/Immigration

	Dependent variable:	
	Estimate of Unemployment W3 - Estimate of Unemployment W2	Estimate of Immigration W3 - Estimate of Immigration W2
Labour Tweets	0.091**	−0.040†
(related to topic)	(0.020)	(0.021)
UKIP Tweets	-0.008	0.074*
(related to topic)	(0.041)	(0.030)
LibDem Tweets	0.014	−0.045
(related to topic)	(0.034)	(0.037)
Tory Tweets	−0.044	− 0.001
(related to topic)	(0.029)	(0.029)
Right Media Tweets	−0.31	−0.007
(related to topic)	(0.051)	(0.042)
Center Media Tweets	− 0.033	0.017
(related to topic)	(0.025)	(0.024)
Left Media Tweets	−0.068	0.086
(related to topic)	(0.049)	(0.055)
Demographic controls	✓	✓
Media Use controls	✓	✓
Observations	1,713	1,398
Note:†p<0.1; *p<0.05; **p<0.01		

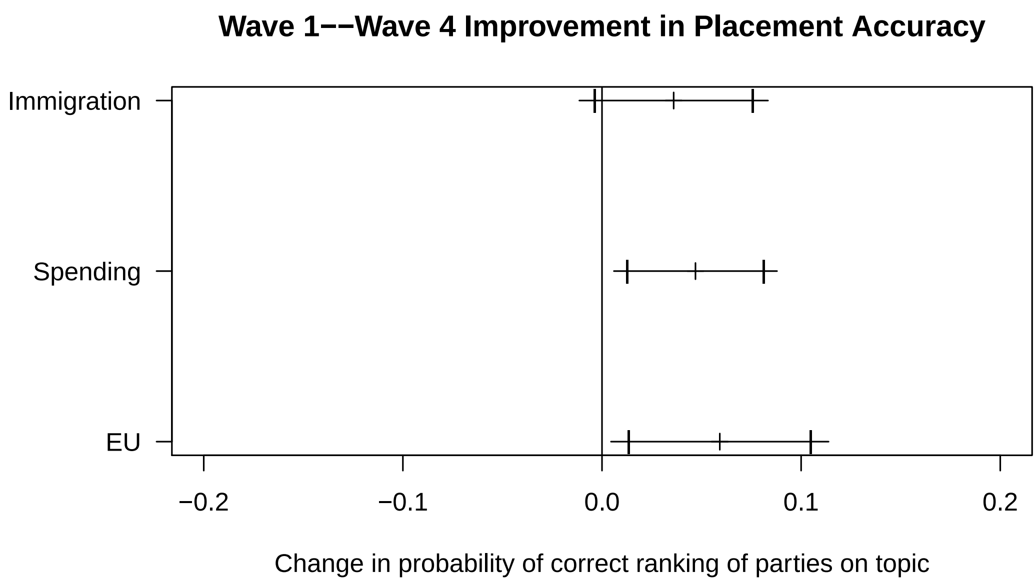
Estimates of the impact of the number of tweets in the respondent's timeline sent by an account affiliated with that party or group of media outlets *and* related to the that topic, calculated from two separate regressions. The dependent variable in each case is an ordinal variable that corresponds to the answer the respondent gave to that factual question in wave 3, estimated with an ordered probit model. Each regression includes demographic and media consumption control variables, as well as a control for the response of the respondent in wave 2.

Figure 1: Respondent and Party Placement on Issues



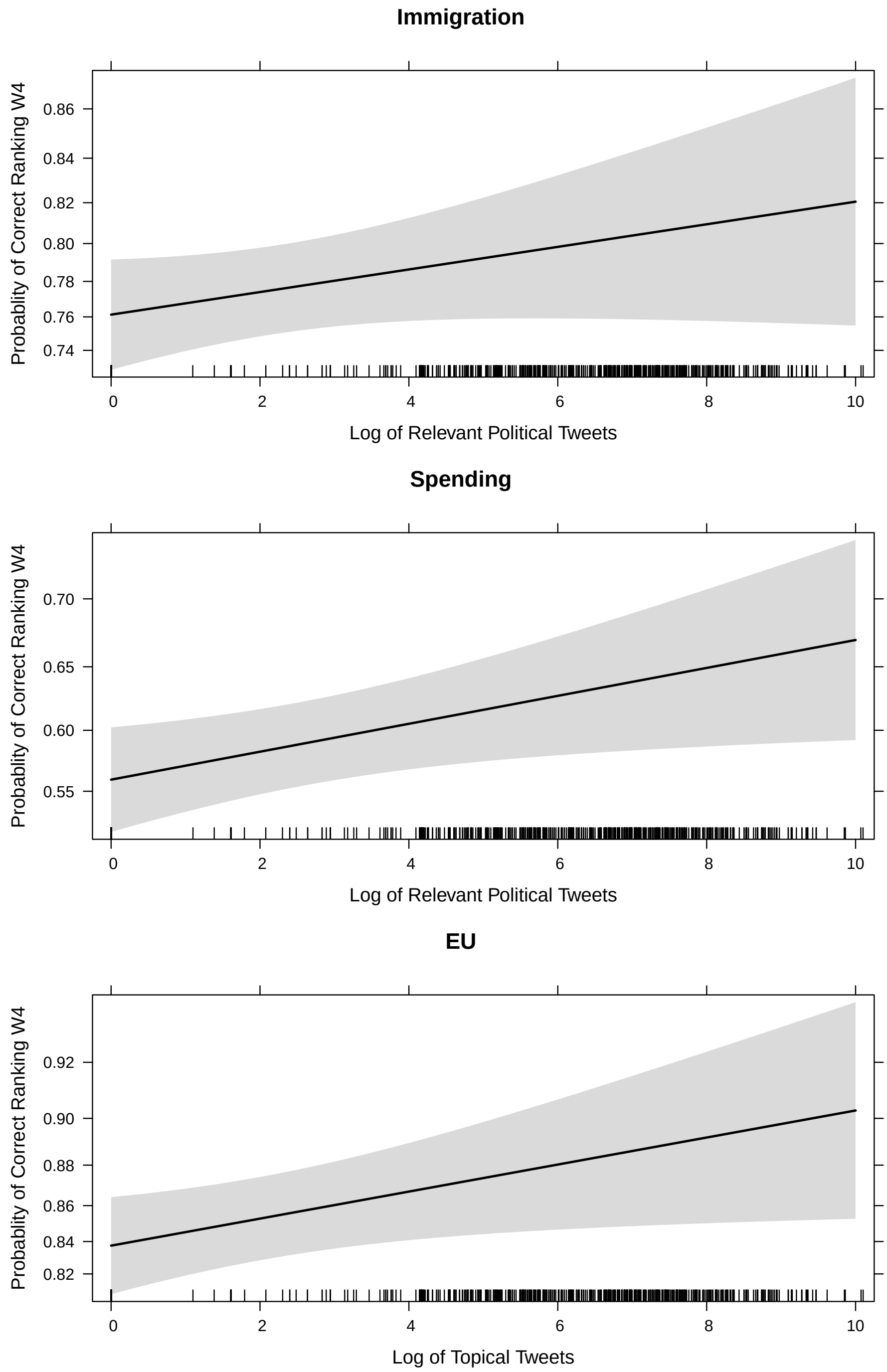
Means and standard deviations of respondents' placements of the four parties and themselves on the three issues under study, at both Wave 1 (the top lines, with squares) and Wave 4 (the bottom, dotted lines, with circles) of the survey. The mean plus standard deviation of UKIP's placement on immigration and the EU exceeded the maximum value of 100, so we truncate them.

Figure 2: Effects of Tweets on Probability of Correctly Ranking Parties



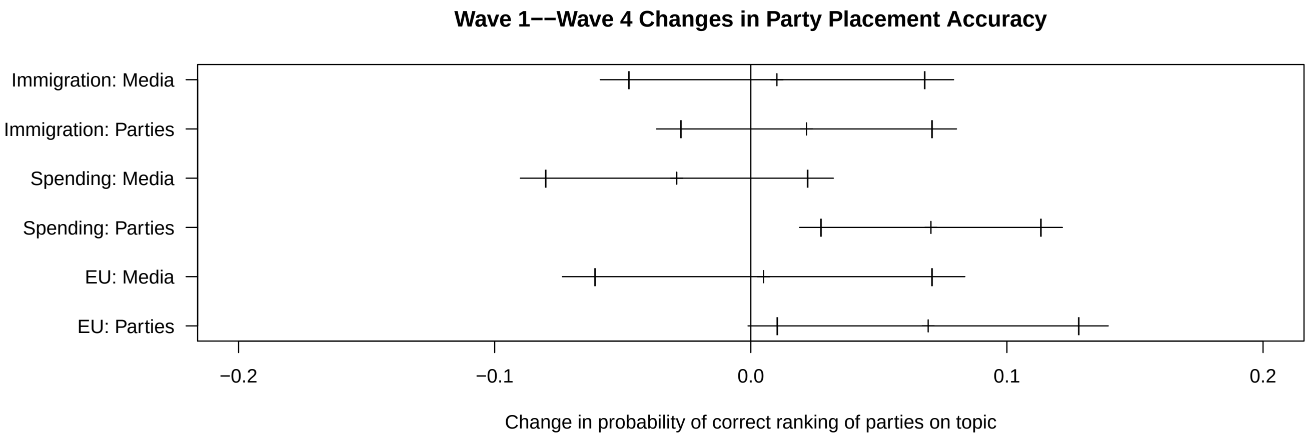
Regression estimates of the effect of topical tweets received on the probability of correctly ranking the parties, by topic. Tick marks are 90% confidence intervals; line segments extend to 95% confidence intervals. The dependent variable in each case is whether the respondent correctly ranked the four parties on that topic in wave 4 of the survey; because this is binary, it is estimated with a logit model. Each regression includes demographic and media consumption control variables, as well as a control for whether the respondent correctly ranked the parties on that topic in wave 1.

Figure 3: Effect of Topical Tweets on Correctly Ranking Parties in Wave 4



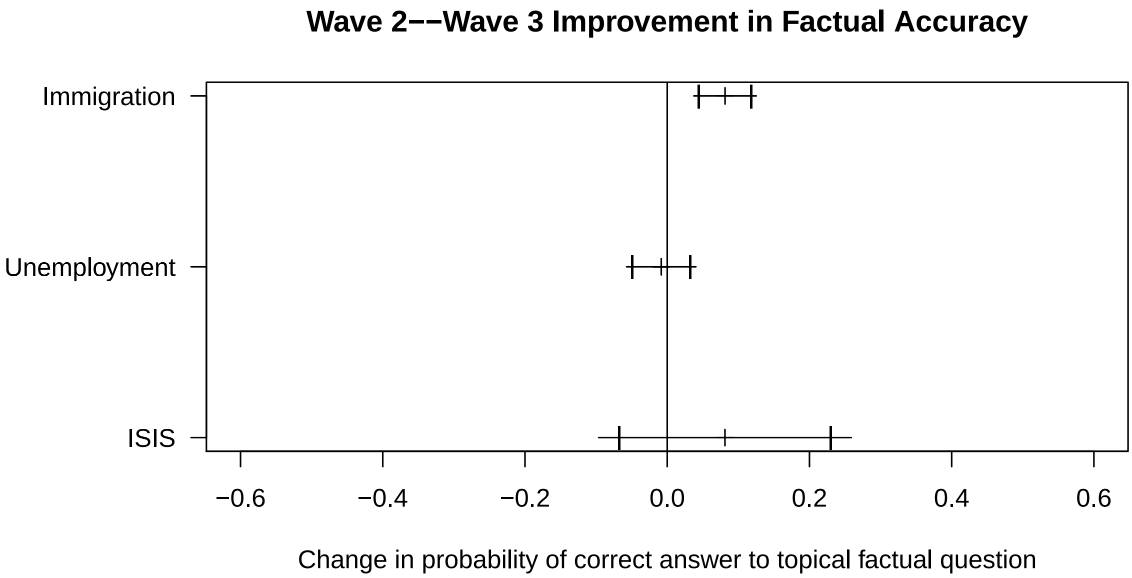
Plots of the estimated effects of topical tweets received on the probability that the subject correctly ranked the four parties on that topic in wave 4 of the survey. This assumes that all other variables, including demographic and media consumption control variables, are fixed at their means.

Figure 4: Effect of Topical Tweets by Source on Correctly Ranking Parties



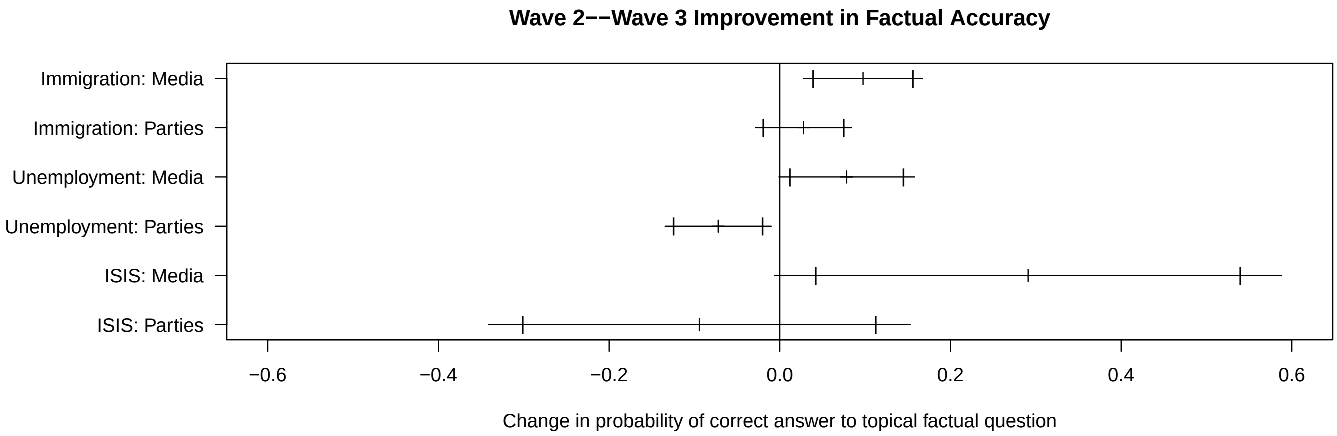
Regression estimates of the effect of topical tweets received from each type of source on the probability of correctly ranking the parties, by topic. Tick marks are 90% confidence intervals; line segments extend to 95% confidence intervals. The dependent variable in each case is whether the respondent correctly ranked the four parties on that topic in wave 4 of the survey; because this is binary, it is estimated with a logit model. Each regression includes demographic and media consumption control variables, as well as a control for whether the respondent correctly ranked the parties on that topic in wave 1.

Figure 5: Effect of Topical Tweets on Factual Knowledge



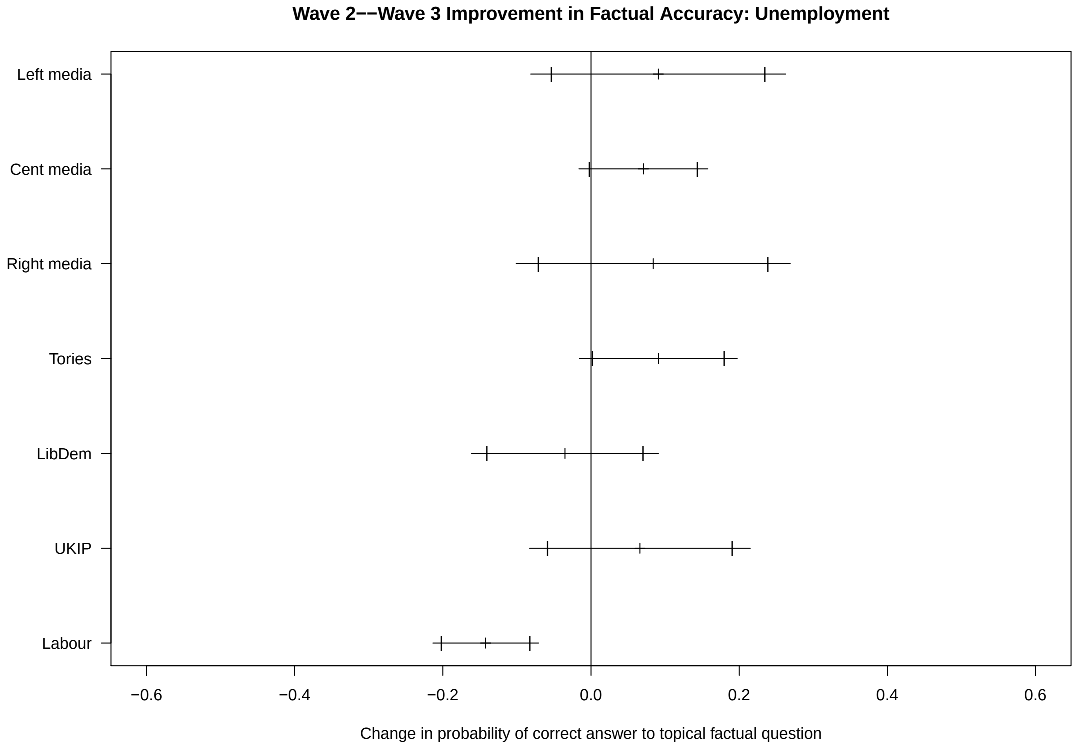
Regression estimates of the effect of topical tweets received on the probability of correctly answering a factual question, by topic. Tick marks are 90% confidence intervals; line segments extend to 95% confidence intervals. The dependent variable in each case is whether the respondent correctly answered the factual question on that topic in wave 3 of the survey, estimated with a logit model. Each regression includes demographic and media consumption control variables, as well as a control for whether the respondent correctly answered the factual question on that topic in wave 2.

Figure 6: Effect of Topical Tweets by Source on Factual Knowledge



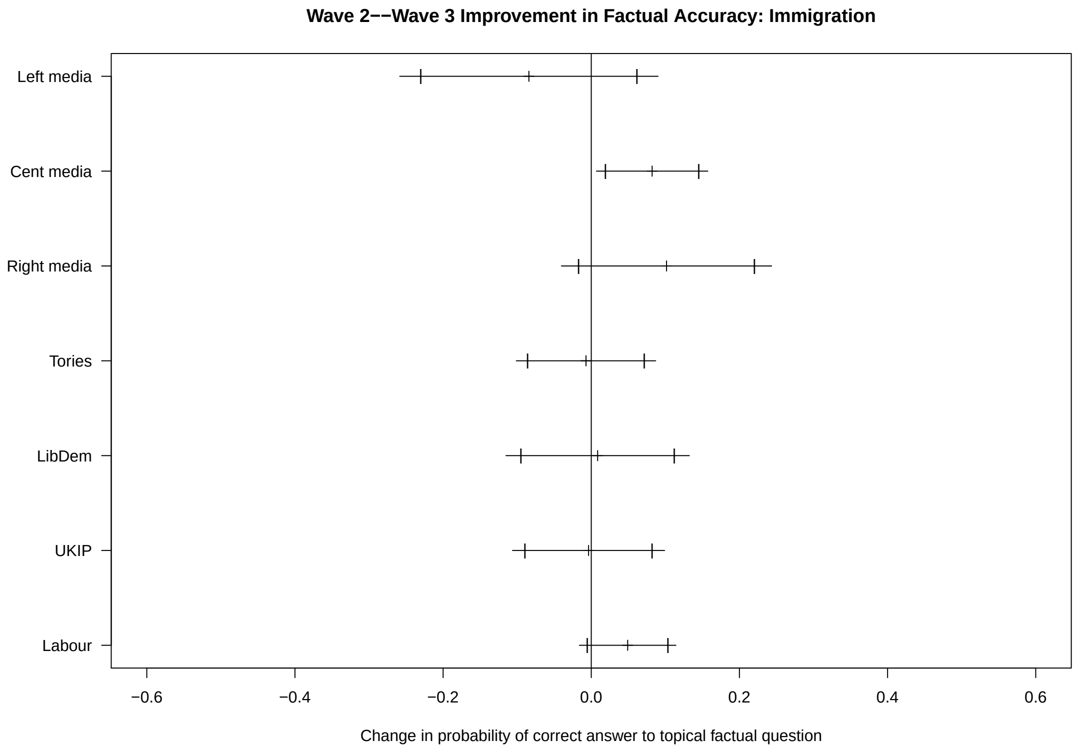
Regression estimates of the effect of topical tweets received from each type of source on the probability of correctly answering a factual question, by topic. Tick marks are 90% confidence intervals; line segments extend to 95% confidence intervals. The dependent variable in each case is whether the respondent correctly answered the factual question on that topic in wave 3 of the survey, estimated with a logit model. Each regression includes demographic and media consumption control variables, as well as a control for whether the respondent correctly answered the factual question on that topic in wave 2.

Figure 7: Effect of Topical Tweets by Source on Factual Knowledge



Regression estimates of the effect of topical tweets received from each type and ideology of source on the probability of correctly answering a factual question, by topic. Tick marks are 90% confidence intervals; line segments extend to 95% confidence intervals. The dependent variable in each case is whether the respondent correctly answered the factual question on that topic in wave 3 of the survey, estimated with a logit model. Each regression includes demographic and media consumption control variables, as well as a control for whether the respondent correctly answered the factual question on that topic in wave 2.

Figure 8: Effect of Topical Tweets by Source on Factual Knowledge



Regression estimates of the effect of topical tweets received from each type and ideology of source on the probability of correctly answering a factual question, by topic. Tick marks are 90% confidence intervals; line segments extend to 95% confidence intervals. The dependent variable in each case is whether the respondent correctly answered the factual question on that topic in wave 3 of the survey, estimated with a logit model. Each regression includes demographic and media consumption control variables, as well as a control for whether the respondent correctly answered the factual question on that topic in wave 2.

Appendix A: Terms Used for Topic Creation

The following are the terms used to create each of the topics analyzed in the paper. If a tweet contained terms from multiple topics, it was labeled as belonging to each of those topics.

ECONOMY: cuts benefits budget welfare vat osborne tax tory disabled tories spending austerity cut reform benefit ids nhs ifs labour disability budget2015 health cameron reforms government

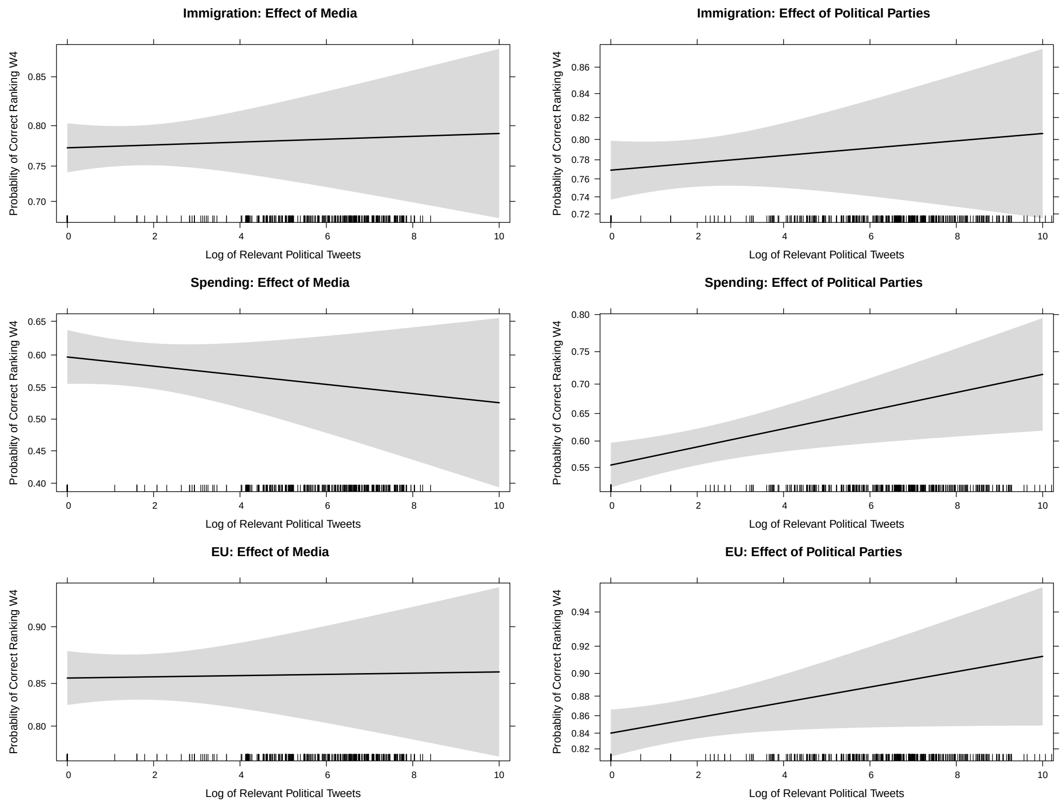
ISIS: isis jihad kobane islam iraq syria fundamentalist iraqi mosul kurds kurdish quran ypg raqqa palmyra islamic twitterkurds fighters ramadi muslim kobani beheading bb4sp beheadings peshmerga

UNEMPLOYMENT: unemployment rate muthafukka youth zerohours nsubsides welfarereform lowest figures toryscum falls jobless employment wages underemployment jobsreport jobs nspain psychocrats massaging longtermplan ngreece satire wca unemployed

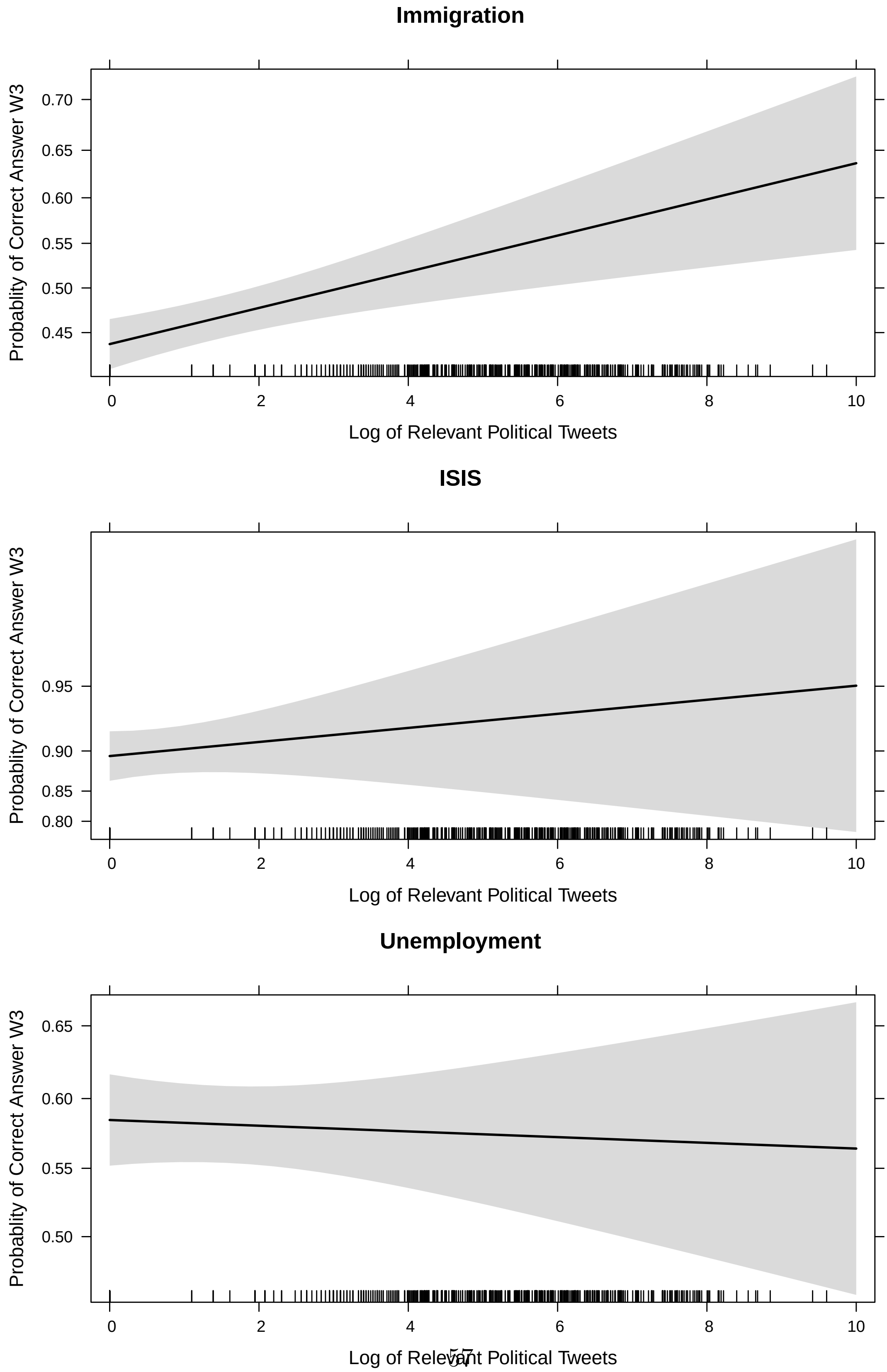
IMMIGRATION: immigration detention uncontrolled ukip obama farage policy controls reform leadersdebate immigrants illegal eu labour yarl mug bbcqt mass bordersecurity nigel ncustoms time4atimelimit noamnesty debate immigrant

TIES TO THE EU: brexit no2eu betteroffout eureferendum eu euref grexit scoxit stayineu flexcit referendum ciuriak yestoeu ivotedukip nothankeu noxi spexit nunelected efta frexit UK scaremongers anually irexit britty

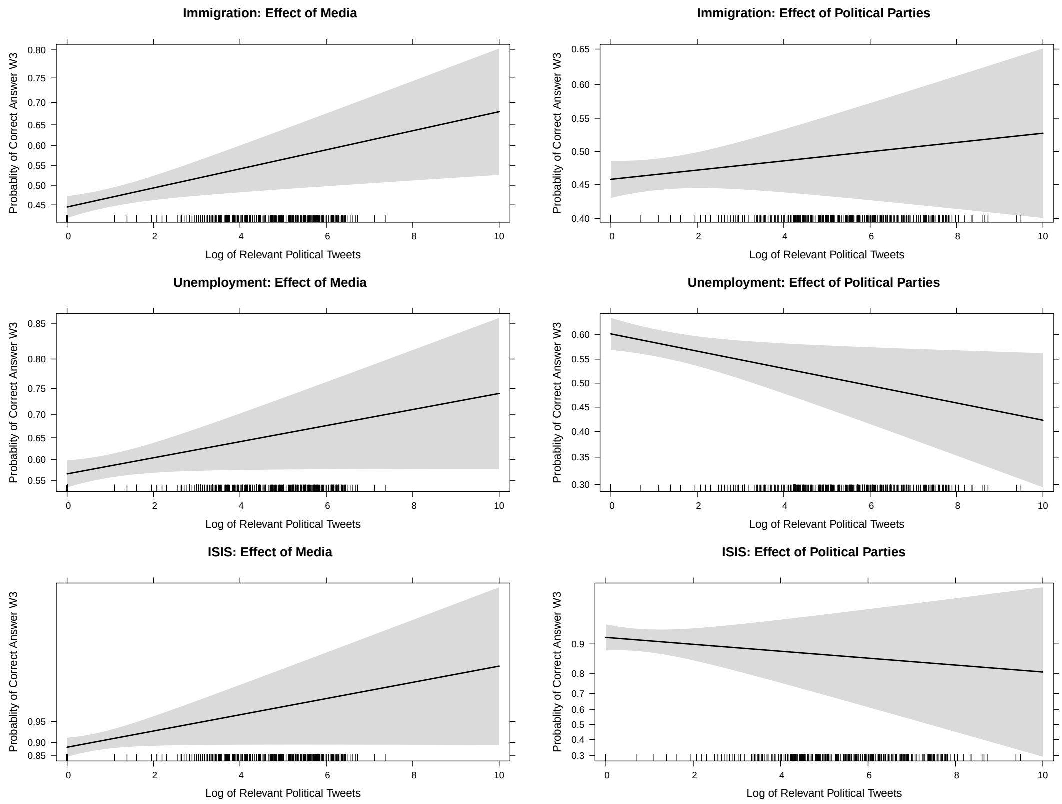
Appendix B: Effects Plots



These plots use the same analysis as those in Figure 4. Effects plot of the impact of the number of tweets in the respondent's timeline related to the that topic by parties or the media on the probability that they correctly ranked the four parties on that topic in wave 4 of the survey. This assumes that all other variables, including demographic and media consumption control variables, are fixed at their means.



These plots use the same analysis as those in Figure 5. Effects plot of the impact of the number of tweets in the respondent's timeline related to the that topic on the probability that they correctly answered the factual question on that topic in wave 3 of the survey. This assumes that all other variables, including demographic and media consumption control variables, are fixed at their means.



These plots use the same analysis as those in Figure 4. Effects plot of the impact of the number of tweets in the respondent's timeline related to the that topic by parties or the media on the probability that they correctly ranked the four parties on that topic in wave 4 of the survey. This assumes that all other variables, including demographic and media consumption control variables, are fixed at their means.

Appendix C: Full Regression Tables

Table 9: Regression Results from Figure 2

	Dependent variable:		
	EU	Spending	Immigration
Answer Previous Wave	1.751***	1.060***	1.596***
	(0.158)	(0.134)	(0.140)
Topical Tweets Received	0.059**	0.047**	0.036
	(0.028)	(0.021)	(0.024)
Twitter Use Frequency	0.0004	−0.001	0.001
	(0.002)	(0.002)	(0.002)
Woman	0.023	−0.044	−0.317**
	(0.148)	(0.141)	(0.143)
Age	0.005	0.008	0.005
	(0.006)	(0.006)	(0.006)
Lower Class	−0.065	−0.143**	−0.058
	(0.055)	(0.055)	(0.055)
Years Education	0.139**	0.179***	0.183***
	(0.057)	(0.054)	(0.056)
Race: Whtie British	0.246	0.416*	0.213
	(0.236)	(0.222)	(0.232)
Married	0.096	−0.291**	−0.095
	(0.147)	(0.139)	(0.144)
Frequency Watch Newsnight	0.005	−0.007	−0.040
	(0.091)	(0.084)	(0.087)
Religious	−0.135	0.237*	0.167
	(0.146)	(0.141)	(0.143)
Frequency Internet Use	0.162	0.191	0.138
	(0.197)	(0.216)	(0.203)
Read Blue Top	−0.309	−0.066	−0.118
	(0.250)	(0.242)	(0.251)
Read Red Top	−0.596**	−0.223	−0.565**
	(0.264)	(0.280)	(0.260)
Read Other Paper	−0.037	0.028	−0.186
	(0.263)	(0.235)	(0.249)
Read No Paper	−0.073	0.188	0.048
	(0.183)	(0.164)	(0.179)
Constant	−1.323	−2.682*	−1.302
	(1.316)	(1.393)	(1.335)
Observations	1,417	1,068	1,376
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 10: Regression Results from Figure 5

	Dependent variable:		
	ISIS	Unemployment	Immigration
Answer Previous Wave	2.611***	2.274***	0.971***
	(0.215)	(0.116)	(0.099)
Topical Tweets Received	0.090	−0.009	0.076***
	(0.090)	(0.025)	(0.022)
Twitter Use Frequency	0.003	−0.002	−0.001
	(0.003)	(0.002)	(0.001)
Woman	0.020	−0.329***	−0.329***
	(0.210)	(0.118)	(0.101)
Age	0.014	0.010**	0.008*
	(0.009)	(0.005)	(0.004)
Lower Class	−0.065	−0.174***	−0.009
	(0.077)	(0.046)	(0.040)
Years Education	−0.016	0.118**	0.066*
	(0.085)	(0.047)	(0.040)
Race: Whtie British	0.033	0.542***	−0.077
	(0.355)	(0.191)	(0.169)
Married	0.431**	0.074	0.100
	(0.212)	(0.120)	(0.102)
Frequency Watch Newsnight	0.425**	−0.039	−0.091
	(0.168)	(0.074)	(0.062)
Religious	0.006	0.358***	−0.167*
	(0.216)	(0.120)	(0.101)
Frequency Internet Use	0.447**	0.419***	0.015
	(0.226)	(0.158)	(0.140)
Read Blue Top	−1.145***	0.739***	−0.258
	(0.434)	(0.215)	(0.171)
Read Red Top	−1.311***	−0.110	−0.289
	(0.446)	(0.227)	(0.199)
Read Other Paper	−0.801*	0.204	−0.003
	(0.483)	(0.215)	(0.181)
Read No Paper	−1.112***	0.128	−0.240*
	(0.356)	(0.148)	(0.126)
Constant	−2.641	−4.119***	−0.564
	(1.650)	(1.062)	(0.934)
Observations	1,892	1,892	1,892
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 11: Regression Results from Figure 4

	Dependent variable:		
	EU	Spending	Immigration
Answer Previous Wave	1.749^{***}	1.054^{***}	1.599^{***}
	(0.158)	(0.134)	(0.140)
Topical Media Tweets Received	0.005	−0.029	0.010
	(0.040)	(0.031)	(0.035)
Topical Party Tweets Received	0.069^{*}	0.070^{***}	0.022
	(0.036)	(0.026)	(0.030)
Twitter Use Frequency	0.0005	−0.0002	0.002
	(0.002)	(0.002)	(0.002)
Woman	0.023	−0.041	-0.320^{**}
	(0.148)	(0.141)	(0.143)
Age	0.004	0.008	0.005
	(0.006)	(0.006)	(0.006)
Lower Class	−0.068	-0.146^{***}	−0.058
	(0.055)	(0.056)	(0.055)
Years Education	0.140^{**}	0.183^{***}	0.184^{***}
	(0.057)	(0.055)	(0.056)
Race: Whtie British	0.248	0.434^{*}	0.219
	(0.236)	(0.222)	(0.232)
Married	0.098	-0.286^{**}	−0.095
	(0.147)	(0.139)	(0.144)
Frequency Watch Newsnight	0.010	0.001	−0.036
	(0.091)	(0.084)	(0.088)
Religious	−0.142	0.230	0.168
	(0.146)	(0.141)	(0.143)
Frequency Internet Use	0.168	0.198	0.147
	(0.197)	(0.216)	(0.203)
Read Blue Top	−0.318	−0.091	−0.128
	(0.250)	(0.243)	(0.252)
Read Red Top	-0.604^{**}	−0.244	-0.581^{**}
	(0.264)	(0.281)	(0.259)
Read Other Paper	−0.038	0.034	−0.196
	(0.263)	(0.236)	(0.248)
Read No Paper	−0.073	0.182	0.041
	(0.182)	(0.164)	(0.179)
Constant	−1.331	-2.720^{*}	−1.337
	(1.316)	(1.393)	(1.335)
Observations	1,417	1,068	1,376
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 12: Regression Results from Figure 6

	Dependent variable:		
	ISIS	Unemployment	Immigration
Answer Previous Wave	2.607***	2.281***	0.969***
	(0.216)	(0.117)	(0.099)
Topical Media Tweets Received	0.298**	0.078*	0.092***
	(0.151)	(0.040)	(0.035)
Topical Party Tweets Received	−0.089	−0.072**	0.025
	(0.126)	(0.032)	(0.029)
Twitter Use Frequency	0.003	−0.002	−0.001
	(0.003)	(0.002)	(0.001)
Woman	0.033	−0.326***	−0.323***
	(0.210)	(0.118)	(0.101)
Age	0.015*	0.011**	0.008*
	(0.009)	(0.005)	(0.004)
Lower Class	−0.054	−0.171***	−0.007
	(0.078)	(0.046)	(0.040)
Years Education	−0.016	0.116**	0.066*
	(0.085)	(0.047)	(0.040)
Race: Whtie British	−0.001	0.539***	−0.077
	(0.357)	(0.192)	(0.169)
Married	0.427**	0.068	0.096
	(0.213)	(0.120)	(0.103)
Frequency Watch Newsnight	0.411**	−0.045	−0.098
	(0.167)	(0.074)	(0.063)
Religious	0.007	0.362***	−0.163
	(0.216)	(0.120)	(0.102)
Frequency Internet Use	0.459**	0.421***	0.021
	(0.227)	(0.158)	(0.140)
Read Blue Top	−1.132***	0.749***	−0.251
	(0.434)	(0.215)	(0.171)
Read Red Top	−1.296***	−0.114	−0.292
	(0.447)	(0.227)	(0.199)
Read Other Paper	−0.807*	0.199	−0.012
	(0.483)	(0.215)	(0.181)
Read No Paper	−1.108***	0.129	−0.238*
	(0.356)	(0.148)	(0.126)
Constant	−2.747*	−4.142***	−0.597
	(1.659)	(1.062)	(0.935)
Observations	1,892	1,892	1,892
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 13: Regression Results from Figures 7 and 8

	Dependent variable:	
	Unemployment	Immigration
Answer Previous Wave	2.579***	2.264***
	(0.216)	(0.118)
Topical Labour Tweets Received	−0.010	−0.142***
	(0.151)	(0.036)
Topical UKIP Tweets Received	11.860	0.066
	(455.522)	(0.076)
Topical LibDem Tweets Received	−0.334	−0.035
	(0.284)	(0.064)
Topical Tory Tweets Received	−0.283	0.091*
	(0.275)	(0.054)
Topical Right-Media Tweets Received	0.298	0.084
	(0.578)	(0.094)
Topical Left-Media Tweets Received	0.427	0.091
	(0.567)	(0.088)
Topical Center-Media Tweets Received	0.250	0.071
	(0.163)	(0.044)
Twitter Use Frequency	0.004	−0.003
	(0.003)	(0.002)
Woman	0.054	−0.314***
	(0.211)	(0.119)
Age	0.016*	0.010**
	(0.009)	(0.005)
Lower Class	−0.050	−0.170***
	(0.078)	(0.046)
Years Education	−0.003	0.115**
	(0.085)	(0.047)
Race: White British	−0.012	0.559***
	(0.361)	(0.194)
Married	0.451**	0.061
	(0.213)	(0.121)
Frequency Watch Newsnight	0.409**	−0.037
	(0.168)	(0.074)
Religious	0.001	0.344***
	(0.217)	(0.122)
Frequency Internet Use	0.459**	0.414***
	(0.227)	(0.158)
Read Blue Top	−1.175***	0.701***
	(0.435)	(0.217)
Read Red Top	−1.306***	−0.111
	(0.447)	(0.227)
Read Other Paper	−0.819*	0.198
	(0.484)	(0.217)
Read No Paper	−1.139***	0.105
	(0.358)	(0.149)
Constant	−2.831*	−4.073***
	(1.660)	(1.066)
Observations	1,892	1,892
Note:	*p<0.1; **p<0.05; ***p<0.01	