Learning (and Unlearning) from the Media and Political Parties: Evidence from the 2015 UK Election[*]

Kevin Munger[†], Patrick Egan[‡], Jonathan Nagler[§]

Jonathan Ronen[¶], and Joshua A. Tucker[‖]

December 27, 2016

[†]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 an ever more prominent role in election campaigns. Does this development educate voters, or mislead them? We explore these questions with a powerful research design measuring change in political knowledge among a panel of voters surveyed during the U.K. 2015 general election campaign. All panelists allowed us to measure 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 tended to increase knowledge of politically relevant facts, and that information received from political parties increased knowledge of party platforms. But in a powerful 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. Extending research on campaign and media effects to the realm of social media, this study shows that campaigns can simultaneously inform and misinform voters. It demonstrates the important role traditional news media continues to play in an era of widespread media disruption.

1 Introduction

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 like 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 addition to being of substantive interest, aspects of social media technology also present 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.

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. Following this six-month campaign, the Conservatives – who before the election had been governing in coalition with the Liberal Democrats under Tory Prime Minister David Cameron – won an outright majority of seats in Parliament. 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 provided us with their account usernames. This allowed us to record in real time the tweets to which panelists were exposed, including the issue-specific content and ideological leanings of the sources that make 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 profound 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 the anti-immigration U.K. Independence Party (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 and Liberal Democrats tended to decrease estimates of the rate of unemployment, but tweets from opposition parties tended to increase these estimates. The fact that the net effect was to inform voter on specific issues, this masks significant heterogenity and the fact that a significant minority of voters became less likely to hold accurate beliefs.

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

5

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.

2   Partisan Information and Learning

There is a general consensus that a more politically informed citizenry is associated with a better-functioning democracy (Campbell et al., 1960; Converse, 1964). On an individual level, political knowledge is a key ingredient in aligning preferences with the political behaviors most likely to realize those preferences. Although there is some scholarly debate about whether and to what extent cognitive short-cuts enable voters who are less knowledgeable to still make the correct voting decision (Bartels, 1996; Fowler and Margolis, 2014; Lau and Redlawsk, 2001; Lupia, 1994), there is little doubt that increased political knowledge contributes to the individual and collective functioning of democracy.

Levels of political knowledge have been improving. Though the topic is less well-studied than in the US, there is evidence of such an increase in knowledge of topical political issues, and of increases in knowledge of party platforms during campaigns in particular, in the UK (Andersen, Tilley and Heath, 2005; Banducci, Giebler and Kritzinger, 2015; Tillman, 2012). Partisan politics in the UK have been undergoing a large shift in the new millennium. Poor economic performance and the unpopular war in Iraq in the mid to late 2000s led to the decline of New Labour, the term for the Labour Party after its move to the right on economic

policy headed by Prime Minister Tony Blair. (Whiteley et al., 2013). The incumbent government in 2015 was a coalition between the Conservatives and the much smaller Liberal Democrats. The coalition government turned out poorly for the Liberal Democrats, who lost support among the left for their cooperation with the Conservatives and especially for their support (contrary to their campaign promises) of an increase in university tuition fees (Weaver, 2015). The distance between the traditional parties and voters on the issues of immigration and the EU similarly hurt the Liberal Democrats and Labour and helped give rise to the nativist UK Independence Party (UKIP) (Evans and Mellon, 2015). This makes the 2015 UK Parliamentary election an excellent case with which to study political learning, as politics is highly salient and information about the parties in flux.

The biggest driver of changes in the flows of political information in society as a whole is the media environment. The development and expansion of new media technologies, from newspapers to broadcast TV to cable, changes both the amount of political information available and distribution of content in a given citizen's media bundle (Prior, 2007). As a given policy topic attracts more media attention, people tend to become more knowledgable about that topic in particular (Barabas and Jerit, 2009). These changes affect people in different ways. Jerit, Barabas and Bolsen (2006) find that newspaper coverage of a topic tends to increase the gap in knowledge between the more and less educated about that topic, while Prior (2005) finds that cable and the Internet increase knowledge among consumers of news media but not among those who prefer entertainment. There is also evidence from cross-national comparisons 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 more market-based system in the US and UK (Curran et al., 2009; Iyengar et al., 2009).

The modern media environment affords people a great degree of choice in media, and there is strong evidence that they tend to self-select into consuming ideologically agreeable media (Iyengar and Hahn, 2009; Stroud, 2008). Even though people are aware of the biased nature of the news they consume, they can still be persuaded by it, as has been shown in the case of cable news networks Fox

News and MSNBC in the US (DellaVigna and Kaplan, 2006; Martin and Yu-rukoglu, 2014). With the advent of social media, the ability of people to further personalize their information environments has caused concern that political communication is increasingly taking place in an "echo chamber." While early studies supported this worry (Conover et al., 2012), more recent work that does not rely on self-selected samples has shown that a large amount of cross-partisan political exchange does take place (Bakshy, Messing and Adamic, 2015; Barberá et al., 2015). Using an approach that is similar to the one we employ in this paper, Flaxman, Goel and Rao (2013) measure the ideological diversity of news and opinion articles in Twitter timelines. They find some ideological segregation, but that this effect is smaller on Twitter than for articles accessed via search engines. This is a crucial point for people's ability to learn during campaigns. If voters, and our respondents, really are isolated in echo chambers then they will never be exposed to information that could cause them to update their views on politically relevant facts or to accurately learn partisan positions during a campaign.

3 Hypotheses

Based on the theories of media effects and motivated cognition discussed above, we have distinct hypotheses about how respondents will learn from different types of political tweets. First, we test the hypothesis that exposure to more tweets about a particular issue will increase knowledge about that issue. Although we believe that characteristics of the source and respondent will mediate this learning, we expect the aggregate effect of tweets from media and politicians to be positive.

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

"Political knowledge" has been measured with a variety of approaches, ranging from questions about long-standing institutions (e.g., "How many members in the House of Commons?") which measure static civil knowledge, to questions about

8

current policy proposals (e.g., "Does Cameron's recently proposed immigration bill aim to increase or decrease the number of immigrants to Britain?") measuring what Barabas et al. (2014) refer to as "surveilance knowledge." In a recent article in the *American Political Science Review*, Barabas et al. (2014) show that exposure to media can have heterogeneous effects on these different measures of political knowledge. Specifically, they find positive effects of media exposure only on "surveillance" knowledge (as opposed to static civic knowledge). And this is our expectation: we do not expect respondents to learn facts about civics or the structure of government institutions during the course of the campaign, rather we expect them to learn about the positions of the parties on issues such as membership in the European Union, and to learn about politically relevant facts – such as how many immigrants enter the UK each year, or what the rate of unemployment is. Barabas et al. (2014) find that these effects are restricted only to general surveillance knowledge, as opposed to policy surveillance knowledge.

We operationalize "knowledge" in two ways. The first is explicitly partisan, and constitutes policy surveillance knowledge. We ask each respondent to place each of the four major parties on a left-right spectrum on three issues: 1) the degree of the UK's ties to the EU; 2) the tradeoff between more social spending and lower taxes; and 3) the level of immigration to the UK. As will be explained in greater detail in the following section, a "correct response" is one that *correctly orders the parties from left to right* on that issue. We assume it is in a given party's interest to attempt to differentiate themselves from other parties on these issue dimensions, and thus the cumulative effect of exposure to Tweets from parties should be a better understanding of where the parties stand *vis a vis* each other on these issues.[1] This expectation concords with the findings in Andersen, Tilley and Heath (2005) that knowledge of party platforms increased during election campaigns in the UK. Furthermore, Banducci, Giebler and Kritzinger (2015) find that exposure to relevant media coverage tends to increase knowledge of party

---

[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 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.

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:

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

We measure general surveillance knowledge about politically relevant facts through multiple-choice questions, examples of which are provided below. These are the types of questions for which Barabas et al. (2014) find positive media effects. Here we expect that parties might very well have an incentive to distort information about facts to better serve their partisan interests, and that the aggregate effect of exposure to information from the parties will be negligible:

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.*

While we do not generally expect tweets from political parties to drive *correct* knowledge of political facts, we can generate more topic-specific hypotheses. Certain facts are inherently political, in that they reflect on the competence of the incumbent parties or the gravity of other highly salient issues. For example, voters evaluate incumbent parties to a large extent through economic performance (Lewis-Beck and Paldam, 2000). As unemployment had been steadily falling in the years leading up to the 2015 UK election, we assume that the incumbent parties (the Conservatives and the Liberal Democrats) would want this fact to be widely known and the opposition parties (Labour and UKIP) would want to obscure this fact (Vavreck, 2009). Our hypothesis is for changes in different directions for different parties on this issue: we expect that exposure to tweets about unemployment sent by incumbent parties will increase knowledge about the

10

true level of unemployment (that it had decreased), and we expect that exposure to tweets about unemployment by challenger parties will decrease knowledge of this fact.

Another important issue in contemporary British politics is legal immigration from the EU to the UK. 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. 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' estimate of the total number of immigrant.

Hypothesis 4 *Exposure to information sent by certain political parties on strategically advantageous topics will increase knowledge of the facts associated with those issue: (H4A) Tweets from incumbent parties will increase knowledge about changes in unemployment; (H4B) Tweets from opposition parties will decrease knowledge about changes in unemployment; (H4C) Tweets from UKIP will increase knowledge about the correct rate of immigration to the UK; (H4D) Tweets from UKIP will increase belief in the number of immigrants coming to the UK.*

4 Data

4.1 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

---

[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

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. 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 each, 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. Overall, there were 1,293 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,

---

selected for their use of Twitter or Facebook. They recently changed the name of the sample to YouGov Social.

[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.

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 UK 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 UK'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.

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,

---

[4]There might be a concern that these median values mask some over-represention of particular

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.

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.

As a "control" group, we drew respondents from another YouGov sample–the "Nationally Representative" (NR) sample. These respondents were entirely separate from the SMU group, but received an identical 4-wave panel survey. Because this sample was representative of the UK 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.

Instead, we only 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).

_____

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.

4.2 Tweets

The "SMU with tweets" 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 followed.

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.[7] 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.[8] Self-reported measures of media use are fraught with measurement

---

[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.

[7]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.

[8]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.

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.[9]

To determine the impact of information seen on Twitter on respondents' preferences we want to curate tweets in respondents' timeline on distinct topics that we measure their opinions on. And we want to be able to aggregate those tweets based on the sources they come from. We chose to examine three key issues we felt to be relevant to the UK election: UK taxing/spending policy; the UK's ties to EU; legal immigration to the UK; and 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 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"; not the most comprehensive list, but 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 these terms, along with their scores can be seen in Table 3. "Brexit" seems to have been an excellent

[9]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.

choice, whereas "euroskeptic" was fairly uncommon, and more appropriate terms expressing the same sentiment included "no2eu" and "betteroffout." [10]

[**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 UK 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. A summary of the distribution of the tweets in the respondents' timelines is shown in Table 4.

[**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

_____

[10]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-occuring words. For a full list of terms, see Appendix A.

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.

5   Results

5.1   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 left-right scale on three major issues in the 2015 election: UK taxing versus spending policy; the degree of UK's ties to the EU; and the level of legal immigration to the UK. 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.[11]

---

[11]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

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."[12] 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.

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 UK'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 UK'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**

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

---

results from waves 2 and 3 in our analysis.

[12]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.

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; the correct orderings and the percentage of respondents identifying them can be seen in Table 5. 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.[13] Overall, ranking the parties on spending was the most difficult, with only 55 percent of respondents in wave 4 answering 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.

The results of the statistical tests of H_1, that exposure to information about a political topic on twitter will increase knowledge about that topic, are presented in Figure 2. Each of the three horizontal lines is the logit coefficient (extended to the 95% confidence intervals) of the logged number of tweets in the respondent's timeline related to the that topic between waves 1 and 4.[14] The dependent variable in each regression is whether the respondent correctly ranked the four parties on that topic in wave 4 of the survey; because this is a binary variable, it is estimated with a logit model. In order to estimate the *change* in knowledge, we control for whether they correctly ranked the four parties on that topic in wave 1. Each regression includes a number of other demographic control variables.[15]

---

[13]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.

[14]Throughout the analysis, we use the log of (1.0001 plus) the number of tweets in respondent's timelines because of the highly skewed nature of these distributions; for brevity's sake, we refrain from saying "log of" in the rest of the paper.

[15]Standard demographic controls are gender, age, class (using the British 5-category system), years of education, race (a binary variable for "white British" or not), marital status, religiosity (binary). Specific control variables for other patterns of media consumption we add are frequency of watching long-running news program Newsnight and frequency of using the internet, (both

We see in Figure 2 that all three of the effects are positive, and that 2 are significant at $p < .05$, while the effect on ranking the parties on immigration is just shy of significant at $p < .10$. These findings support our hypothesis that repondents 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. This approach 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 was exposed to e^{10} tweets about immigration instead of 0, their predicted probability of ranking the parties correctly increases from .56 to .67.

[**Figure 3**] **here**

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 disaggregate these topical tweets by the type of source–if the tweets were sent by a political party (or related politician), *or* by a media organization (or affiliated journalist.) Figure 4 demonstrates that the analysis of Figure 2 is 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). Tweets from the media do not have a statistically significant effect on correctly ranking the parties. These findings lend credence to H_2.

[**Figure 4**] **here**

---

on a 5-point, ordinal scale), and dummy variables for Newspaper Type. In the UK, different types of newspapers are significant signifiers of group identity and carry different kinds of news, so reading "Red Tops" (tabloids like the The Sun or The Daily Mirror) or "Broadsheets" (The Guardian or The Telegraph) is an important measure of media exposure.

5.2 Factual Knowledge

The other way we operationalize political knowledge is through factual questions about 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?[16]

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

Table 6 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." Overall, there is little difference in either the absolute levels of knowledge or in changes in knowledge between the two samples. We can see in Panel A that 89% of frequent Twitter users gave the correct answer on ISIS in both waves 2 and 3, while 6% 'learned' the answer between waves (giving the wrong answer in wave 2 but the right answer in wave 3), and 3% gave the incorrect answer in both waves, and 2% actually became less-informed: giving the correct answer in wave 2, but the 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. Put another way, about as many people got the question right in wave

---

[16]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.

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.[17]

[**Table 6 Here**]

Our analysis here uses the same specification as for the party placements. 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 independent 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.[18] Figure 5 plots three horizontal lines that represent the logit coefficient (with standard errors) of these primary independent variables.[19]

[**Figure 5 Here**]

Figure 5 shows a positive and significant relationship between the number of tweets and respondents' knowledge of immigration, but not for unemployment and ISIS, lending weak support to H_1. However, the standard errors for the estimate on ISIS are quite large. This is due to the comparatively few tweets on that topic (see Table 4, Panel B). The estimate is positive, but small: 88% of respondents got the question correct in both wave 2 and wave 3, so there is little change in knowledge to be explained.

We again disaggregate these tweets by their source, 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. 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.

---

[17]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.

[18]We include full results in an online appendix.

[19]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.

[**Figure 6 Here**]

We expect that some political parties have strategic incentives to emphasize different aspects of the same topic, and that these divergent emphases might have contrasting effects on change in knowledge, our H_4. When discussing the economy, for example, the incumbent parties (the Conservatives and Liberal Democrats) should want to play up the fact that the UK 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 perform the same analysis as above to determine the impact of tweets about a topic (seen by) a respondent increases knowledge of that topic, but now with 7 independent tweet count variables. The results are shown in Figure 7.[20]

[**Figure 7 Here**]

The results of our analysis of the Unemployment question 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 have a slightly negative effect, contrary to our expectations of a positive effect. Our post-hoc explanation is that the Liberal Democrats suffered greatly from their alliance with the Conservatives, in which they helped support an increase in tuition fees that enraged their constituency. As a result, they actually tried to distance themselves from the coalition govern-

---

[20]We do not plot the results for the analysis of the ISIS question. 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 relative to the other topics that the standard errors of our estimates of the disaggregated tweet counts are large.

ment, which would explain our contrary findings, although given our low levels of statistical confidence in any effect here we do not want to make too much this finding.

Turning to immigration, 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 truve value of the number of immigrants each year, UKIP may have been exageratting the number

To explain these results, 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 categories to the relevant multiple choice question.[21]

Table 7 displays the results of these ordered probit regressions. We use 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 7 Here**]

Column 2 of Table 7 explains the null results from the analysis of immigration

---

[21]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 DV takes a value of -1.

knowledge in Figure 7. 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.

These findings do not concord with H_4(C), but they do support H_4(A, B and D), and note that (C) and (D) are incompatible (contingent on respondents' prior beliefs). We thus find 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.

Overall, we find moderate support for our hypotheses, and explain the cases in which support is lacking. The primary effect of exposure to tweets related to certain topics is to increase knowledge of those topics. 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.

5.3   Causality

In the above discussion, we assume that observed correlations between exposure to tweets about a certain topic and changes in knowledge of that topic are causal. This is intuitively plausible, but we lack a crucial element to estimate 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 approach poses for causal inference but 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 effect of exposure to tweets on changes in knowledge is omitted variable bias (OVB). For example, consider our analysis of the topic of immigration. We estimate in Figure 7 that exposure to tweets about immigration tends to increase subjects' estimate of the number of immigrants to UK. It is possible that an omitted variable, "interest in the topic of immigration," is causing both "assignment to treatment" (interest in immigration causes people to follow accounts that tweet about immigration) and changes in the outcome (interest in immigration causes people to raise their estimate of the number of immigrants to the UK). If this is the case, our estimate of the effect may be biased upward.

This line of argument is harder to maintain when we consider the dependent variable not as changes in true or false beliefs but rather changes in absolute estimates of national statistics–specifically, the unemployment rate and the rate of immigration to the UK. We find in Table 6 that exposure to tweets by Labour causes subjects to increase their estimate of the unemployment rate in the UK. If we are again concerned about OVB, we need to maintain that interest in unemployment (or some other omitted variable related to the issue of unemployment) causes subjects to self-select into following Labour politicians *and* causes them to raise their estimate of the unemployment rate. Given that the true answer to the multiple choice question we use to measure this belief in unemployment rate takes three values (decreased, stayed the same, increased) and that the correct answer is the lowest (decreased), this entails a decrease in belief accuracy (see Figure 9).

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.

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 usage on political knowledge?" The randomized experiment described above is not "Twitter usage" per se, but something else, something artificial that does not happen in the real world.[22]

Although we acknowledge the limitations of our research design for causal inference, we believe that we are asking the right question and that our findings bring us closer to an answer.

6   Conclusion

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, research in this area is hampered by the challenge of measuring both the independent variable (information exposure) and the dependent variable (political knowledge) in this relationship. The approach used in this paper offers a new way to circumvent these issues. By using a 4-wave panel survey design and focusing on changes in knowledge (i.e., learning) rather than levels, we control for stable individual-level characteristics of respondents. Further, by matching survey responses to actual measures of exposure to political information on social media, we are able to avoid the biased media exposure self-reports that often plague this independent variable. The net result is "real-world" evidence of the effect of social

---

[22]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.

media exposure on political knowledge.

Our findings deal only with social media users, a group that is not representative of the UK population as a whole. However, social media users represent a large and growing fraction of the electorate, worthy of study on their own terms. Our findings about the types of knowledge that are subject to media effects largely concords with previous findings (Barabas et al., 2014), suggesting that research on traditional media can complement analysis of the effect of social media.

Contrary to the worst fears of some, there was very little evidence of a negative relationship between social media use and changes in political knowledge in the 2015 British election campaign. We found that exposure to UKIP tweets during the election campaign did lead to an upward revision in an individual's belief in the number of immigrants coming to the UK in the past year, but this did not on balance lead to a decrease in the number of respondents correctly identifying the number of immigrants to the UK. And indeed, across the whole study, we did not simply witness those who received more political tweets learning more "incorrect" knowledge about politics. In fact, this almost never happened, and the one case in which it did – the main opposition party (Labour)'s tweets obscuring the fact that unemployment had fallen – is the one which decades of research on economic voting suggests should not have surprised us. Exposure to tweets about politics generally, tweets from political parties, and tweets from the media seem to have increased political knowledge in at least some areas some of the time. Thus in the first analysis – to our knowledge – to have combined respondents' Twitter feeds with panel survey data over the course of an election campaign, we find evidence consistent with the idea that exposure to politics on Twitter may actually be contributing to a more politically informed mass public.

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.

Campbell, Angus, Philip E Converse, Warren E Miller and E Donald. 1960. "The American Voter.".

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.

Converse, Phillip. 1964. "The Nature of Belief Systems in Mass Publics. In Ideology and Discontent, ed. David Apter. New York: Free Press.".

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.

Jerit, Jennifer, Jason Barabas and Toby Bolsen. 2006. "Citizens, knowledge, and the information environment." *American Journal of Political Science* 50(2):266–282.

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* .

Lewis-Beck, Michael S and Martin Paldam. 2000. "Economic voting: an introduction." *Electoral studies* 19(2):113–121.

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: Covariates

	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: UK 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.

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%

Table 4: The mean percentage of tweets about each topic sent by each source received by respondents. For example, the bottom right corner says that, among the 161 respondents who received at least one tweet sent by Centrist 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: 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.

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%

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%

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%

Table 6: 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 knowledge questions: (C,C), (C,I), (I,C), (I,I). The bottom line shows how difficult each question was showing the percentage correct in wave 2.

Table 7: 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: 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.

Issue Placement Means and Standard Deviations, Waves 1 and 4
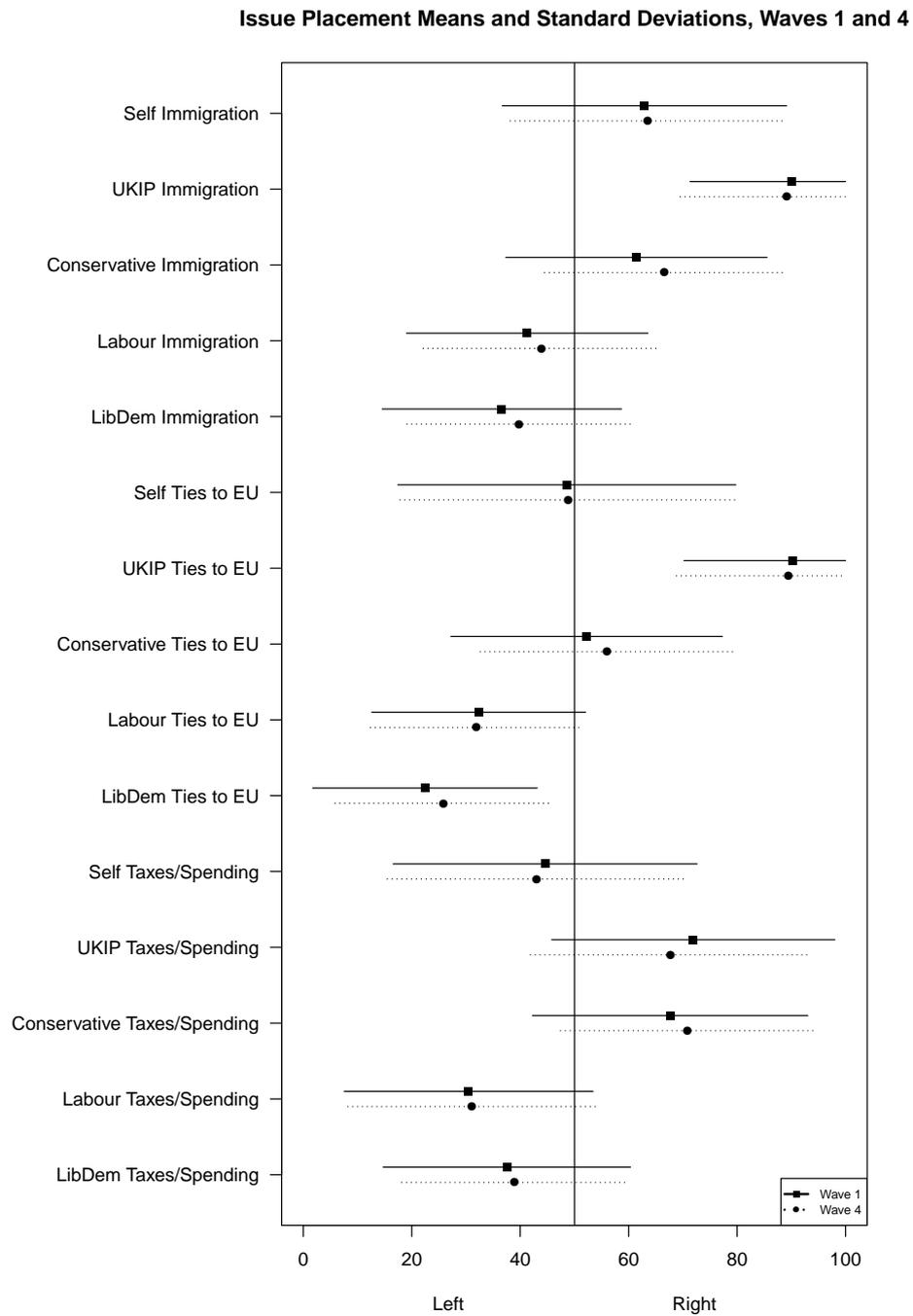
Wave 1--Wave 4 Improvement in Placement Accuracy
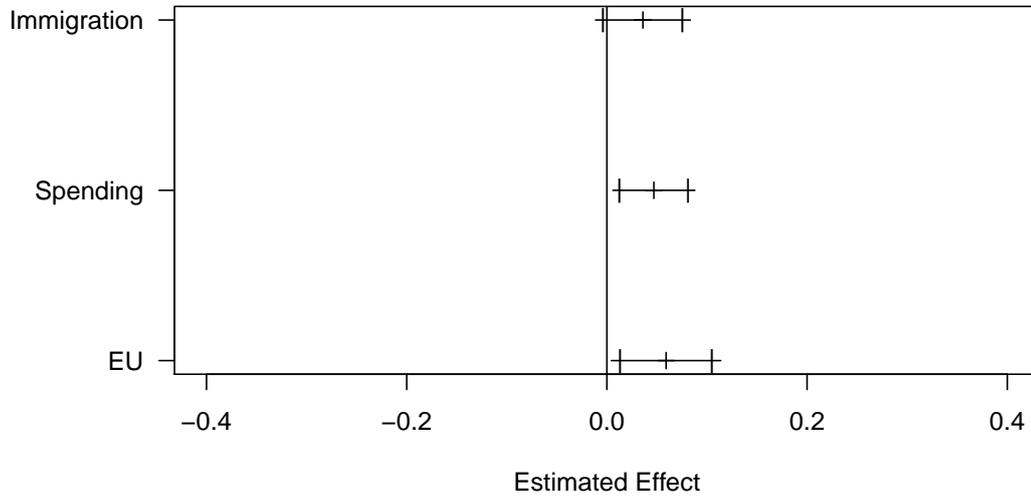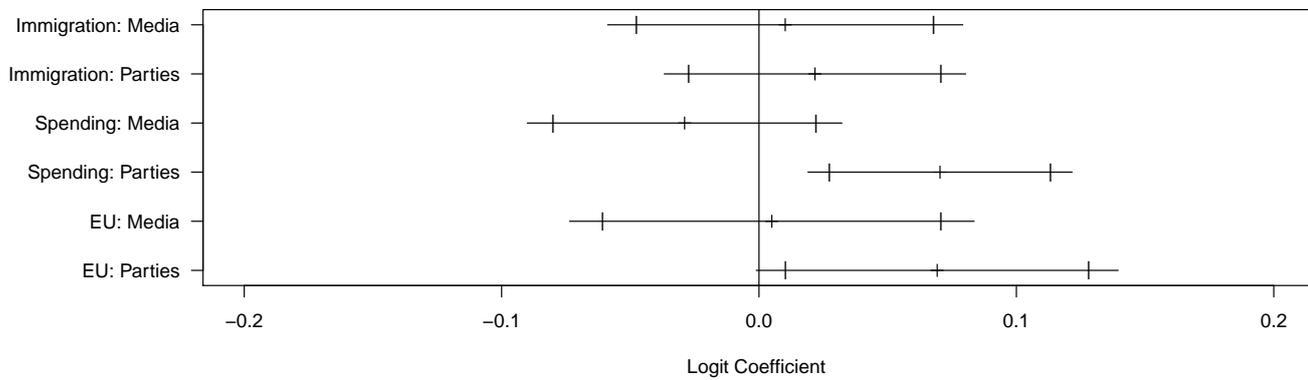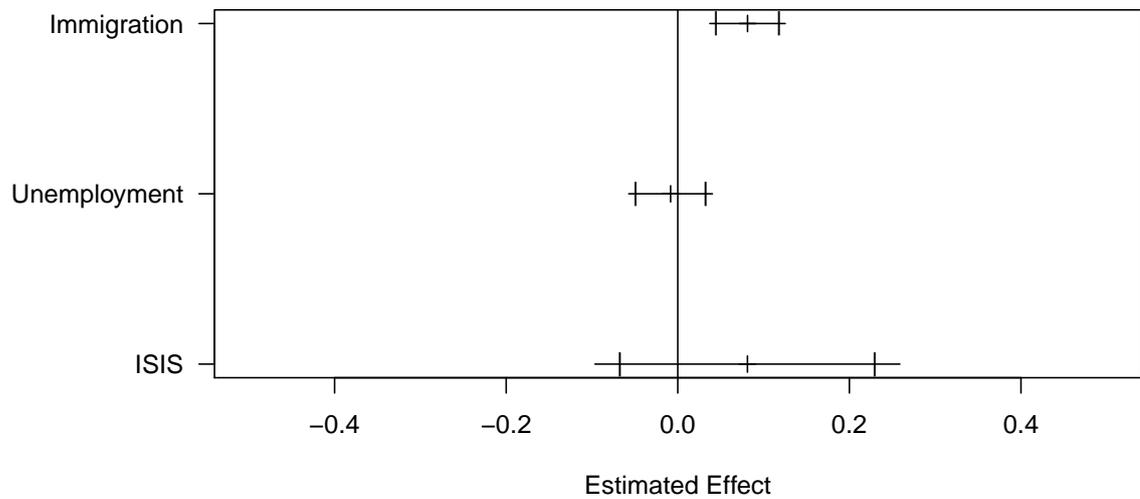

Figure 2: Estimates of the impact of the number of tweets in the respondent's timeline related to the that topic, with three separate regressions. Tick marks are 90% confidence intervals; line segments extend to 95% confidence intervals. The dependent variable in each case is whether the respondent correctly 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.
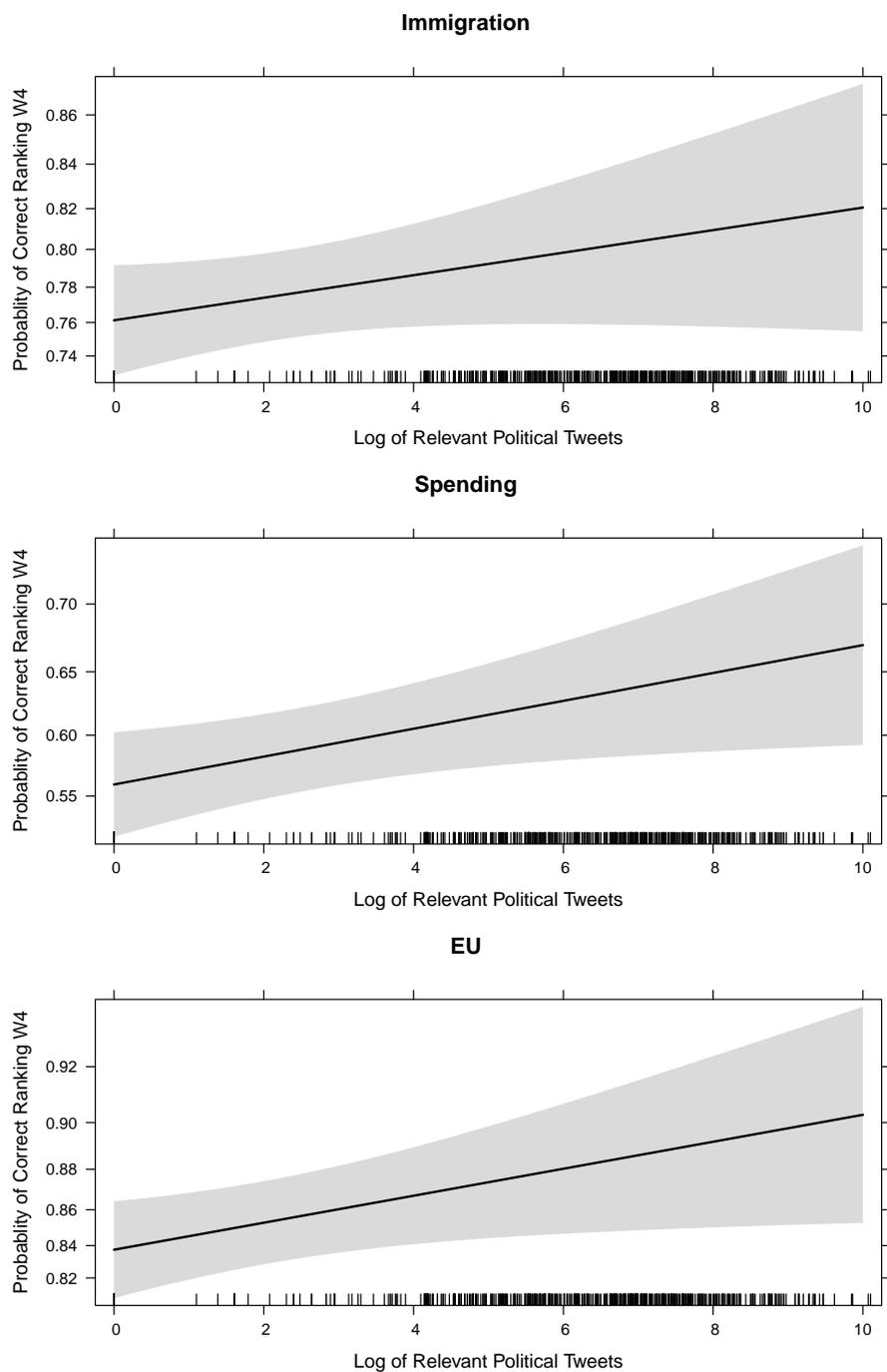
Figure 3: Effects plot of the impact of the number of tweets in the respondent's timeline related to the that topic on the probability that they correctly ranked the four parties on that topic in wave 4 of the survey. This assumes that all other variables, including demographic and media consumption control variables, are fixed at their means.

45

Effects of Tweets on Changes in Wave 1--Wave 4 Party Placement Accuracy
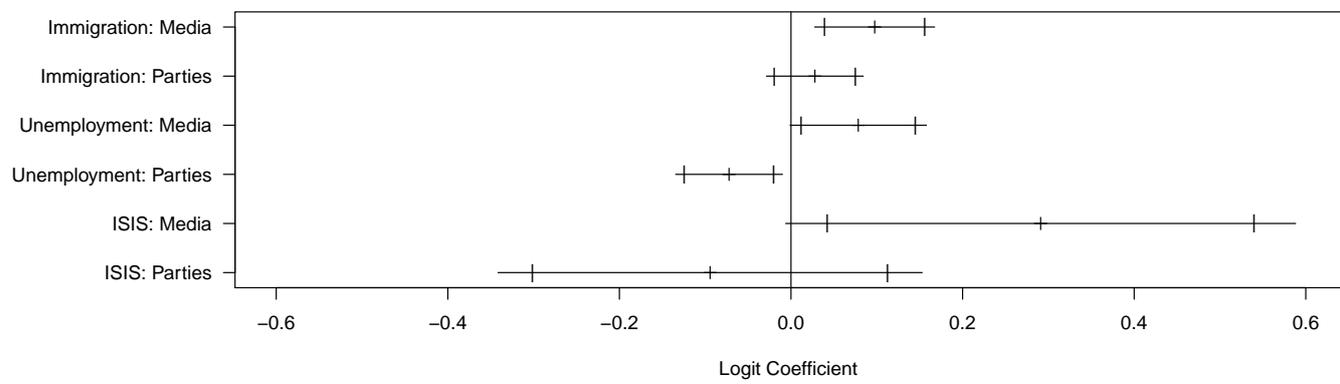

Figure 4: Estimates of the impact of the number of tweets in the respondent's timeline related to the that topic disaggregated by source, with a separate regression for each of the three topics. Tick marks are 90% confidence intervals; line segments extend to 95% confidence intervals. The dependent variable 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.

Wave 2––Wave 3 Improvement in Factual Accuracy

Figure 5: Estimates of the impact of the number of tweets in the respondent's timeline related to the that topic, with three separate regressions. 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.
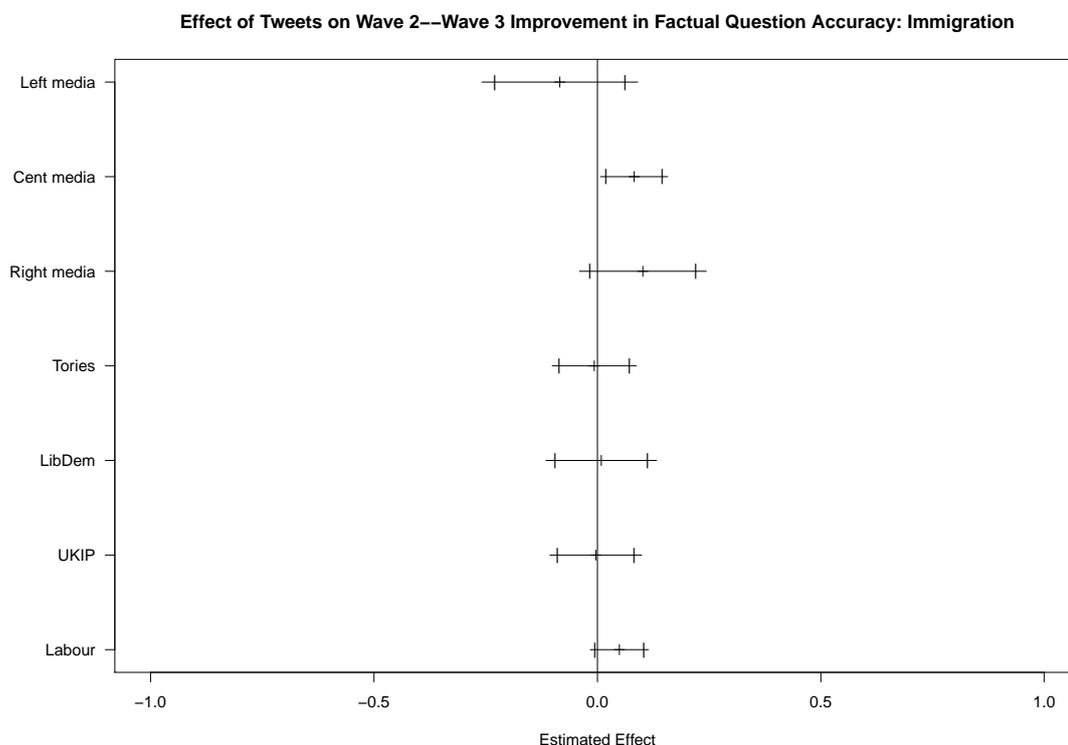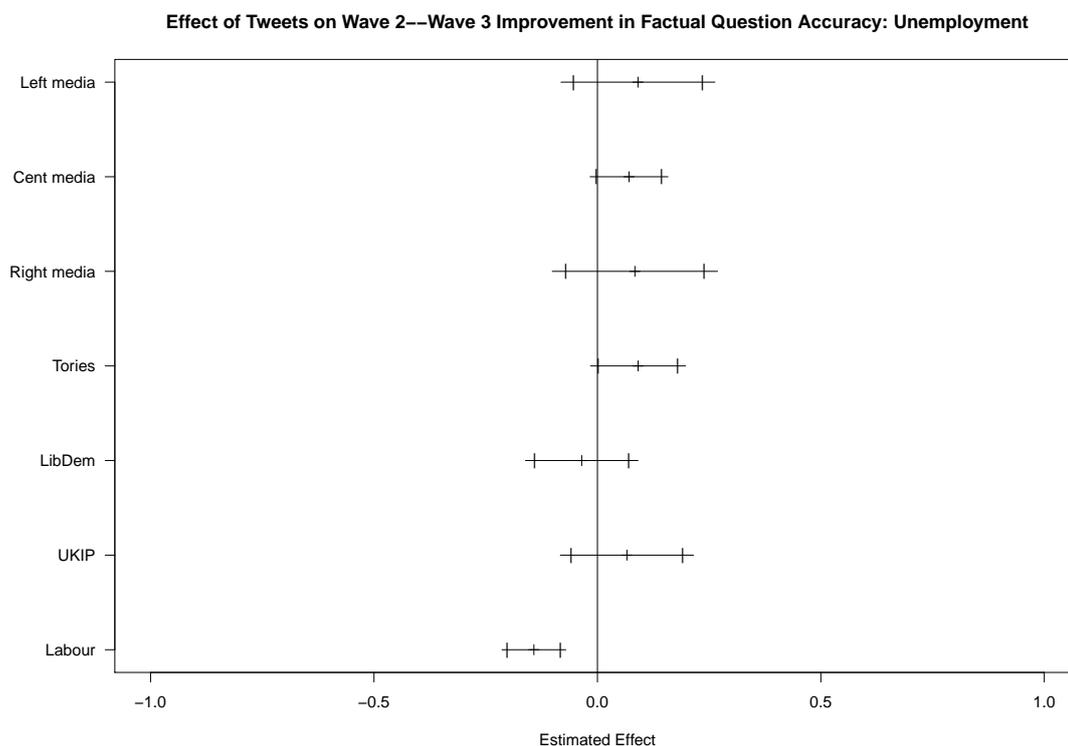
Effects of Tweets on Wave 2––Wave 3 Changes in Factual Question Accuracy

Figure 6: Estimates of the impact of the number of tweets in the respondent's timeline related to the that topic disaggregated by source, with three separate regressions. 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.

Effect of Tweets on Wave 2--Wave 3 Improvement in Factual Question Accuracy: Unemployment

Effect of Tweets on Wave 2--Wave 3 Improvement in Factual Question Accuracy: Immigration

Figure 7: Estimates of the impact of the number of tweets in the respondent's timeline related to the that topic disaggregated by source, with two separate regressions. 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
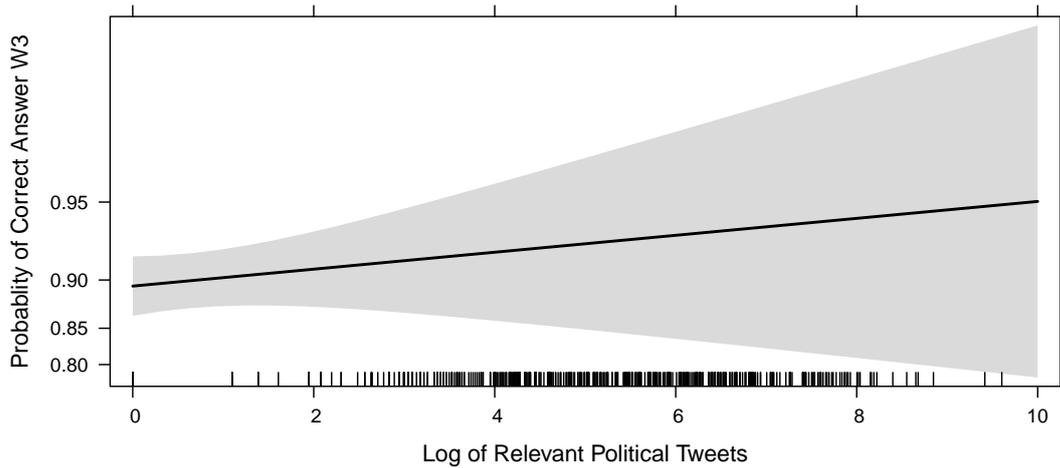
Appendix B: Effects Plots
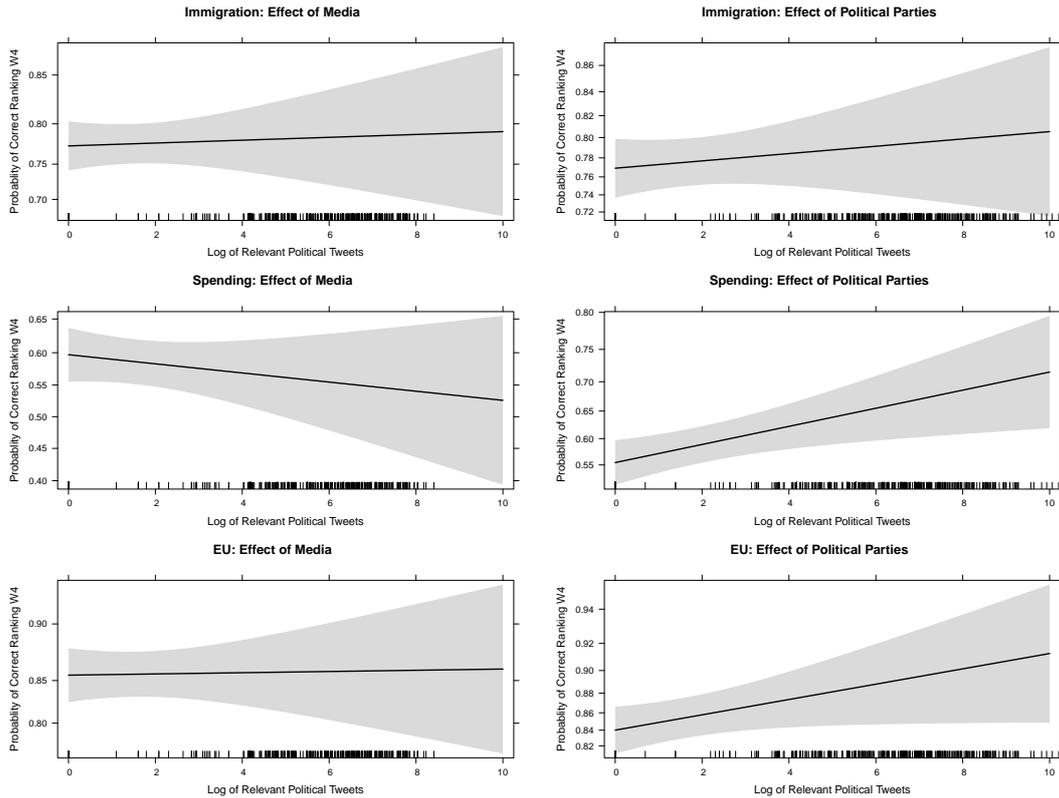
These plots use the same analysis as those in Figure 4. Effects plot of the impact of the number of tweets in the respondent's timeline related to the that topic by parties or the media on the probability that they correctly ranked the four parties on that topic in wave 4 of the survey. This assumes that all other variables, including demographic and media consumption control variables, are fixed at their means.
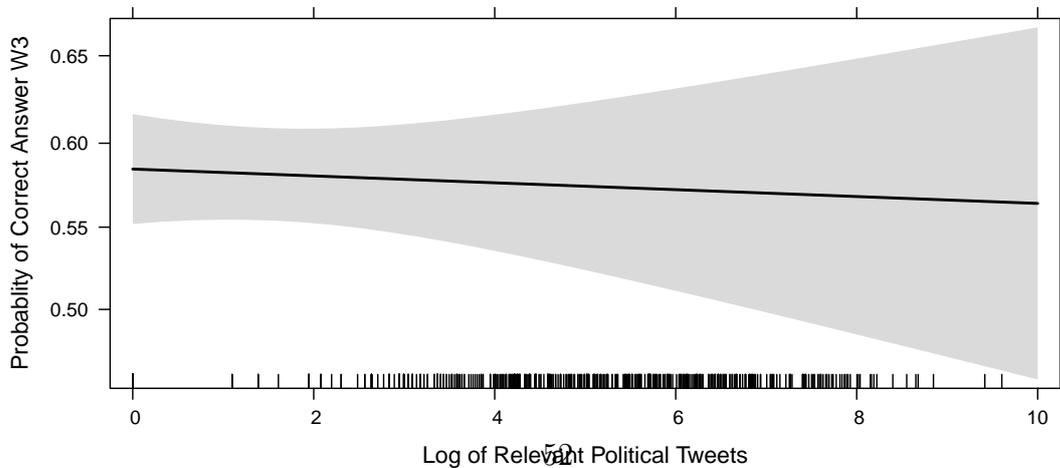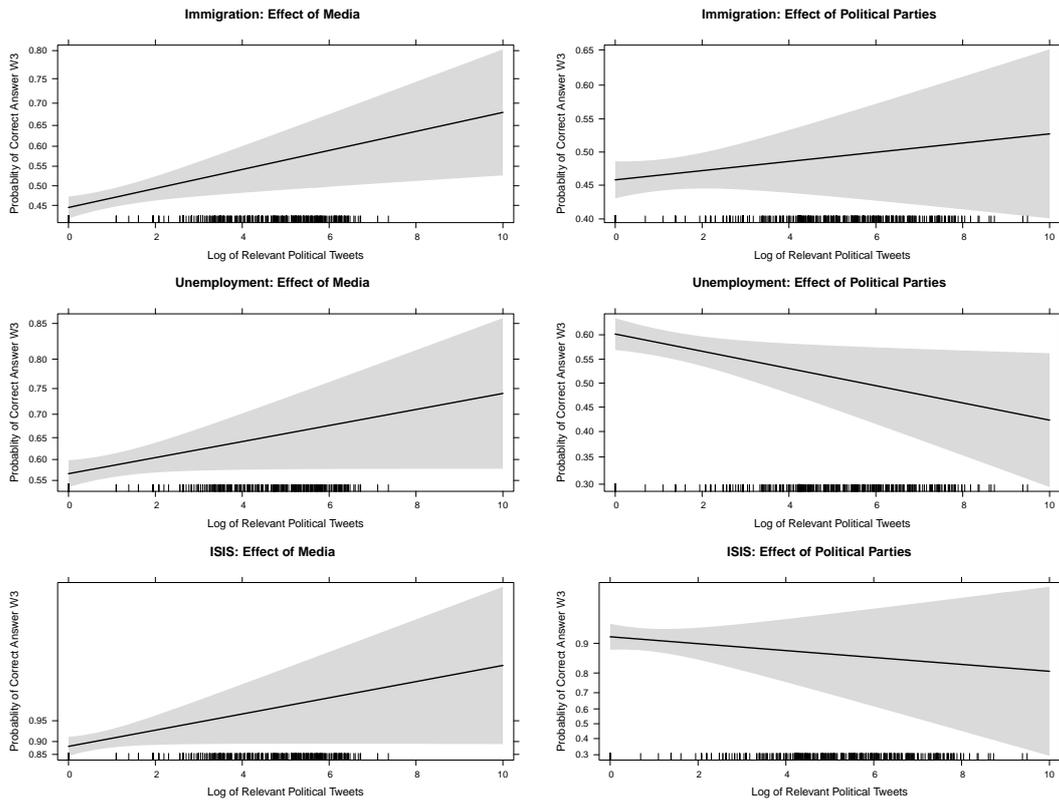
Immigration
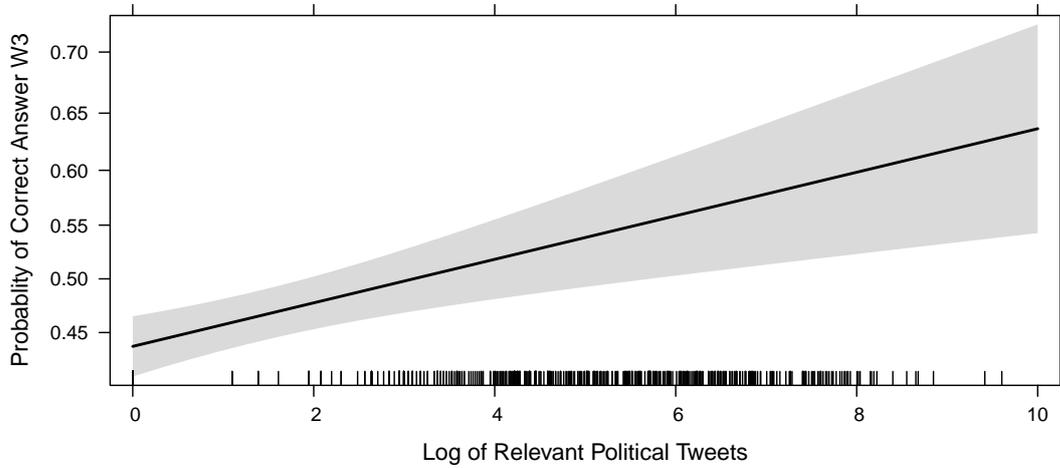
ISIS

Unemployment

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.