Political Psychology

Political Psychology, Vol. xx, No. xx, 2016
doi: 10.1111/pops.12337

The Social Structure of Political Echo Chambers: Variation in Ideological Homophily in Online Networks

Andrei Boutyline
University of California, Berkeley

Robb Willer
Stanford University

We predict that people with different political orientations will exhibit systematically different levels of political homophily, the tendency to associate with others similar to oneself in political ideology. Research on personality differences across the political spectrum finds that both more conservative and more politically extreme individuals tend to exhibit greater orientations towards cognitive stability, clarity, and familiarity. We reason that such a "preference for certainty" may make these individuals more inclined to seek out the company of those who reaffirm, rather than challenge, their views. Since survey studies of political homophily face well-documented methodological challenges, we instead test this proposition on a large sample of politically engaged users of the social-networking platform Twitter*, whose ideologies we infer from the politicians and policy nonprofits they follow. As predicted, we find that both more extreme and more conservative individuals tend to be more homophilous than more liberal and more moderate ones.*

KEY WORDS: political homophily, ideology, motivated cognition, Twitter

We draw on research on personality differences across the political spectrum to develop and test the prediction that people with different political orientations will exhibit different levels of political homophily, the tendency to choose to associate with others similar to oneself in political ideology. Ideological groups with greater political homophily possess political networks with more ties among their members and fewer ties with individuals possessing different ideologies. Thus, greater political homophily is associated with decreased chances of politically diverse interactions and increased rates of interactions with ideologically similar others that tend to reinforce individuals' views and enhance their commitment to their ideological group. These outcomes are in turn likely to increase the polarization of public opinion and promote participation in political collective action.

Since at least John Stuart Mill (1859), political theorists have argued that dialogue across lines of political difference is a key pre-requisite for sustaining a democratic citizenry. Mill held that political disagreement enables individuals to develop skills for critically assessing political claims and provides the challenge necessary for determining if one's own ideas are justified. Hannah Arendt similarly argued that debate "constitutes the very essence of political life" (Arendt, 1961, p. 241), irreplaceable for forming enlightened political opinions that reach beyond the limits of one's own subjectivity to

Published by Wiley Periodicals, Inc., 350 Main Street, Malden, MA 02148, USA, 9600 Garsington Road, Oxford, OX4 2DQ, and PO Box 378 Carlton South, 3053 Victoria, Australia

incorporate the standpoints of others. Empirical work on consequences of disagreement has echoed many of these points. Existing research shows that individuals without exposure to cross-cutting discourse are less likely to see opposing viewpoints as legitimate and less able to provide rationales for their own political decisions (Huckfeldt, Mendez, & Osborn, 2004; Price, Cappella, & Nir, 2002). Such individuals are also more likely to hold extreme attitudes about candidates consisting of entirely positive or negative assessments (Huckfeldt et al., 2004). Moreover, the lack of personal ties to those with different political views is likely to have detrimental effects on political tolerance (Mutz, 2002a). Increased political homophily, and decreased cross-cutting contact, are therefore likely sources of polarization and political discord.

Conversely, political homophily creates dense clusters of within-group ties, which prior work shows reinforce behavioral norms and increase social pressure to take part in costly or risky activities (Centola, 2013; Centola & Macy, 2007). Politically homophilous networks have significant advantages for diffusing political behaviors that require normative pressure or social confirmation—including behaviors like turning out to vote, attending political protests, and engaging in potentially contentious political speech (González-Bailón, Borge-Holthoefer, Rivero, & Moreno, 2011; Kim & Bearman, 1997; Knoke, 1990; Romero, Meeder, & Kleinberg, 2011). At the same time, political homophily may also insulate individuals from exposure to false or offensive information.

Further, a relative dearth of cross-cutting ties is itself a likely resource for collective action, as exposure to dissent can undermine commitment to the group and the extent to which the group's beliefs are taken as facts. Experimental and observational evidence suggests that heterogeneous ties increase ambiguity, which has a demotivating effect on political participation and engagement (Eveland & Hively, 2009; Mutz, 2002b; Visser & Mirabile, 2004)—an effect that has been shown to hold across national settings and in both online and offline networks (Liu, Dai, & Wu, 2013; Mutz, 2006; Valenzuela, Kim, & Zúñiga, 2012). Campbell (2013) summarizes this work by pointing out that strength of preferences, such as identification with a political cause, "does not exist in a vacuum; it is reinforced by a social network of like-minded politicos" (p. 41).

Recently, scholars have sought to qualify this effect by examining variation in consequences of cross-cutting exposure. For example, Jang (2009) found that, while cross-cutting ties are often demotivating, they also motivate participation among the most politically alienated individuals by increasing their understanding of differences between competing positions. Klofstad, Sokhey, and McClurg (2013) also found that effects of disagreement vary between kinds of contact and measures of participation and engagement, but they are overwhelmingly negative. Campbell's (2013) review of literature on networks and participation similarly suggests that, though the effect of cross-cutting ties may not always be negative, it is rarely positive.[1] Thus, while its effects are not monolithic, political homophily on average appears to be an asset for many kinds of collective action.

But how might political homophily vary by ideology? Two bodies of research show that people at different points in the ideological spectrum exhibit different levels of desire for clarity, certainty, stability, and familiarity—a cluster of traits we refer to as *preference for certainty*. First, a long line of work from political psychology finds that more conservative individuals exhibit greater preferences for certainty than more liberal ones (Jost et al., 2003a). Second, research on group identity hews that individuals on either ideological extreme possess greater preferences for certainty than more moderate ones (Greenberg & Jonas, 2003; Hogg, 2007). These findings suggest that more conservative or more extreme individuals may exhibit higher levels of political homophily, as they might be expected to place greater value on encountering concurring opinions and avoiding dissenting ones. As individuals with greater preferences for certainty seek it through social contact, their networks may come to resemble "echo chambers," providing them with reaffirmation and shielding them from disagreement.

[1] Campbell (2013) also highlights research showing that cross-cutting ties may have different effects on participation than other sources of exposure to disagreement (see also Nir, 2005).

These intuitions are difficult to test with traditional political network surveys, which face well-documented methodological challenges, including a substantial prohomophily bias in respondents' recall of their alters' political orientations and difficulties establishing "baseline" rates of homogeneity expected from random mixing (DiPrete, Gelman, McCormick, Teitler, & Zheng, 2011; McPherson & Smith-Lovin, 1987). Here, we address these problems by using network data from Twitter, an online service used by 12% of adult Americans (Smith & Brenner, 2012). Employing a recently validated technique for ideological measurement of Twitter users (Golbeck & Hansen, 2014), we infer users' ideology from the ideological positions of members of Congress and policy nonprofits with whom they initiate ties. We then test our hypotheses by examining 238,943 ego networks from across the political spectrum. The Twitter data are not a representative sample of U.S. voters or any other offline population, which precludes direct statistical generalization of our results to offline phenomena. At the same time, the size and diversity of the Twitter population as well as the observability of Twitter activity bring novel advantages that help overcome long-standing problems common to more traditional data on political networks.

Uncertainty and Threat

In developing our claims relating ideology and political homophily, we draw upon the substantial literature on personality and political attitudes in social psychology and political science. Beginning with *The Authoritarian Personality* (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950), a central argument in this literature has been that individuals' political ideologies and behaviors are partly rooted in chronic personality traits (Jost, Federico, & Napier, 2009). Among the most robust results in this work is the finding that more conservative individuals typically exhibit a cluster of traits reflecting greater orientations towards certainty. Classic studies show that, compared to liberals, conservatives have a preference for reasoning that is dichotomous or based on clear categories to qualified or probabilistic reasoning, a greater tendency to experience threat or anxiety when faced with uncertainty, a lower desire for new experiences, and a higher desire to reach firm conclusions quickly (see the review in Jost et al., 2003a). The *uncertainty-threat hypothesis* (Jost, Glaser, Kruglanski, & Sulloway, 2003a) proposes that the common thread uniting these findings is differences in responses to unknown, uncertain, or threatening situations, which we refer to as "preference for certainty."

Social-scientific treatments have frequently identified traditionalism and opposition to change as fundamental aspects of conservative ideology (e.g., Huntington, 1957; Jost, 2006). Both aspects appear related to preferences for a more stable, certain, and familiar world. In contrast, liberalism is associated with a more positive view of change. For this reason, the uncertainty-threat hypothesis predicts that individuals with stronger preferences for certainty should tend towards conservatism over liberalism. This hypothesis has found strong and consistent support across 50 years of research (Jost et al., 2003a).

Uncertainty and Identity

Another line of research suggests that individuals on the ideological extremes, both left and right, show stronger preferences for certainty than more moderate individuals. This view of the political "true believer" (Hoffer, 1951) suggests that the motivational needs of managing uncertainty and threat are addressed through rigid adherence to extreme ideologies (Greenberg & Jonas, 2003; Hogg, 2007).

According to *uncertainty-identity theory* (Hogg, 2007), group identification reduces uncertainty by providing individuals a clear sense of self and prescriptions for behavior based on prototypical group characteristics. Uncertainty was found to increase the strength of party identification among both conservatives and liberals (Hohman, Hogg, & Bligh, 2010). Since more extreme groups provide greater contrast between members and nonmembers and thus clearer behavioral prototypes and membership criteria (Hogg, 2004), uncertainty-identity theory predicts that individuals with greater needs for certainty may be drawn to more extreme ideologies. Consistent with this, individuals have been shown to identify with more extreme ideological groups when their level of uncertainty was experimentally increased (Hogg, 2004). This is also consistent with the notion that uncertain economic times often coincide with the rise of extreme ideologies. This mechanism could operate at the same time as the one proposed by the uncertainty-threat hypothesis, and a mixed model of the two has found some empirical support (Hogg, 2007; Jost, Glaser, Kruglanski, & Sulloway, 2003b).

From Motivation to Action

We expect that ideological groups whose members hold greater preferences for certainty will exhibit greater levels of homophily. Homophilous contact can confirm worldviews and reinforce ideologies, while heterophilous contact threatens to seed uncertainty and doubt. Thus, it stands to reason that those seeking greater certainty should do so in part via political homophily.

Past research supports this reasoning. Heightened desire for cognitive closure is associated with homophilous preferences such as favoritism for members of one's ethnicity and greater identification with partners in ad hoc groups (Shah, Kruglanski, & Thompson, 1998), and people with higher sensitivity to threat hold more hostile attitudes towards outgroups (Hatemi, McDermott, Eaves, Kendler, & Neale, 2013). The desire for heterophilous contact, however, is associated with traits typical of low desires for certainty, such as sensation seeking and openness to experience (Gerber, Huber, Doherty, & Dowling, 2012; Mehrabian, 1975). Past research also confirms that heightened uncertainty leads to a greater affinity for groups of homogenous, similar others (Jetten, Hogg, & Mullin, 2000).

Summary of Claims

We argue that individuals higher in preferences for certainty will seek social confirmation and avoid disagreement, making them more likely to form homophilous ties. Drawing on the research reviewed above, we propose two hypotheses:

H1: Ego networks on the ideological right will exhibit greater political homophily than those on the left.

H2: Ego networks on the ideological extremes will exhibit greater political homophily than those at the center.

Measuring Political Homophily. Our investigation of political homophily builds on a long research tradition. Early sociometric surveys provided evidence of the political homogeneity of core networks by asking respondents about their closest contacts (Knoke, 1990; Laumann, 1969). These "strong-tie" surveys could not speak to the homogeneity of broader ego networks, as stronger ties are markedly more homogeneous than weaker ones (Granovetter, 1973). Evidence of political homogeneity in broad acquaintanceship networks came from a recent General Social Survey (2006), which specifically measured both weaker and stronger ties (DiPrete et al., 2011).

However, as DiPrete and colleagues point out, measures derived solely from respondents' descriptions of alters capture only *perceived* homophily, which may greatly exaggerate its true levels. For example, studies that interviewed both respondents and their alters found that respondents frequently overestimated their political similarity (Goel, Mason, & Watts, 2010; Huckfeldt, Beck, Dalton, & Lavine, 1995; Laumann, 1969). Out of seven respondent-provided alter characteristics that Laumann (1969) verified via interviews with the alter, party identification was the least accurate, with reported and true identification correlated at $r = .51$. The mistake rate was correlated with ideology, indicating a potential confound.

Another empirical challenge comes from the difficulty of distinguishing network homogeneity produced by homophilous tendency—“homophily proper” (Wimmer & Lewis, 2010) or “choice homophily”—from homogeneity due to other mechanisms. If groups of potential homophilous partners differ in size, random tie creation would lead the majority group to have more homogeneous ties than the minority, even without any homophilous tendency (Blau, 1977; Feld, 1982). This kind of “baseline” homophily (McPherson & Smith-Lovin, 1987) is difficult to rule out with survey data, as the availability of potential homophilous partners in a social environment is generally unknown. While studies of complete face-to-face networks within bounded settings can estimate baseline rates with relative ease, their homogeneity and small scale makes observation of political homophily difficult. For example, in a study of networks between master’s students in a public policy school, Lazer, Rubineau, Chetkovich, Katz, and Neblo (2010) did not find evidence of significant homophily on the basis of either politics or gender, attributing this lack of political homophily to an artifact of their demographically homogeneous sample.

Thus, measuring political homophily involves three major difficulties. First, to measure discrepancies from baseline levels of homogeneity expected from random mixing, the relative availability of potential homophilous partners must be known. Second, information on alters’ political orientations should be drawn from sources other than the ego’s report. And finally, the network data should cover a broad sample of respondents and a range of alters beyond the closest “strong-tie” core. To our knowledge, no published work meets all three criteria.

We also know of no work that examines whether rates of political homophily differ across the political spectrum in interpersonal networks. Such difference was, however, noted in an innovative study of political blogs, with the weblink structure between conservative blogs appearing denser than between liberal ones (Adamic & Glance, 2005). Though a follow-up reanalysis of the data failed to replicate this finding (Ackland & Shorish, 2009), the results still pose a provocative question about possible asymmetries in rates of political homophily. Barberá’s (2015) finding that conservative Twitter users forward messages from other conservatives at greater rates than liberals do from other liberals similarly suggests this possibility.

Method

To test our claims, we examined the Twitter networks of roughly a quarter million politically engaged Americans. Using a procedure recently validated by Golbeck and Hansen (2014), we located these individuals by identifying the Twitter accounts of major U.S. political actors with previously measured political orientations (159 members of Congress and 33 policy nonprofits). We used these as a proxy for the orientations of their followers. We then calculated homophily measures for the ego networks of these followers and analyzed the resulting dataset via multivariate regression with cluster-adjusted standard errors.

Research Site

Twitter is both a social networking service and media platform. Users post short messages (called “tweets”) to their profile. Immediately, everyone subscribing to their account (their “followers”) receives copies of those messages. About 90% of all Twitter accounts are public, meaning that anyone can subscribe (or “follow”) them, view their posts, or examine their ego networks (Takhteyev, Gruzd, & Wellman, 2012).[2] The entire stream of public tweets can also be searched by keyword, allowing users to locate accounts that interest them. In contrast to offline networks, where the choice of partners

[2] From this point, we use “ego network” to refer to the set of the ego’s Twitter ties, the users those ties point to, and the sets of ties belonging to those users.

is often highly restricted by geography, competent Twitter users who wish to create new homophilous ties can thus do so with ease and on a practically limitless scale. The resulting network is composed of directed and often asymmetric ties of *attention* and so features high-degree "hub" nodes belonging to major journalists, celebrities, politicians, and other popular content producers. Such hubs form the basis of our sampling strategy.

Between 2010 and 2012, the percentage of adult Americans using Twitter increased from 5% to 12% (Rainie, 2010; Smith & Brenner, 2012; Smith & Rainie, 2010).This broad and quickly growing user base, combined with the unparalleled observability of online social activity, make services like Twitter a valuable resource for social research. However, these data also introduce some important limitations. Like most large complete-network datasets, our dataset is a single cross-sectional snapshot, precluding many approaches to causal inference. Additionally, we lack demographic covariates for our sample. We thus cannot rule out the possibility that homophily on an unobserved trait is responsible for the homogeneity of ties we observe. Furthermore, our sample is not representative of Americans: The Twitter user base is younger, more female, more educated, higher income, and features higher rates of racial and ethnic minorities than the overall population (Smith & Rainie, 2010). The higher average education of Twitter users in particular might make them more opinionated and thus more politically homophilous than the American public. Our analysis is therefore best viewed as an unusually large and diverse case study rather than a snapshot of the American electorate, leaving open the possibility that the effects we observe are limited to this self-selected, albeit large, population.

On the other hand, Twitter data have important advantages relevant to the methodological challenges detailed above. Since our dataset contains all public Twitter accounts, we can calculate the total number of potential homophilous partners for any given user, which in turn allows us to control for the baseline homophily rates we would observe under random mixing. Equally important, Twitter network data derive from observation rather than self-report, thus avoiding the well-documented prohomophily bias faced by most survey studies. Finally, while a user's Twitter ego network is by no means the same as their offline ego network, its size and geographical distribution are suggestive of a broad mixture of online and offline contacts as well as stronger and weaker ties (see Takhteyev et al. 2012 on Twitter geography). Thus, while Twitter data bring unfamiliar challenges, they also solve many familiar problems, making them a valuable complement to more traditional data.

While not directly generalizable to offline populations, there are nonetheless good reasons to study Twitter for insight into U.S. political networks. First, Twitter is a significant political communication platform in its own right, as evidenced by the range of major political actors who use it. Twitter use is ubiquitous among U.S. social-movement organizations (Obar, Zube, & Lampe, 2012), who use it to disseminate information and mobilize collective action. As of this writing, virtually all members of Congress have Twitter presences (Hemphill, Otterbacher, & Shapiro, 2013). Twitter users' attention to these politicians tracks offline behavior: For example, the volume of Twitter mentions of a congressional candidate predicts her electoral performance, even net of key covariates including incumbency and media coverage (DiGrazia et al., 2013).

Second, studies demonstrate that Twitter networks share many properties and processes with offline phenomena. For example, Dunbar, Arnaboldi, Conti, and Passarella (2015) show that ego networks on both Twitter and Facebook have strikingly similar distributions of degree and tie strength to offline networks, leading them to conclude that "the structure of online social networks mirrors those in the offline world" (p. 39). Geographic distances, national borders, and frequency of air flights also affect ties in ways that resemble networks offline (Takhteyev et al., 2012). In their study of social movements on Twitter, González-Bailón and colleagues (2011) find evidence that both protest recruitment and informational diffusion occur over Twitter ties. They also find that online political behavior diffusion is consistent with the same "complex contagion" dynamics (Centola & Macy, 2007) thought to describe the diffusion of behavior in offline networks, as do Romero, Meeder, and Kleinberg (2011).

The parallels between Twitter and offline electoral politics are also illustrated by Barberá (2015), who shows that the ideological positions of members of the 112th Congress can be estimated solely from Twitter ties among their followers. Barberá treats shared followers similarly to how roll-call ideal-point scaling techniques interpret shared votes. The resulting estimates nearly perfectly recreate roll-call measures of the politicians' ideological positions ($r > 0.94$), yielding a distribution of ideal points for ordinary users that approximates this distribution offline. They also closely track survey and demographic measures of citizen ideology when agglomerated at the state level ($r > 0.87$).

Sample Selection

To create our sample, we first searched Twitter for members of the 111th U.S. Congress, locating 31 active accounts belonging to senators and 128 belonging to representatives (30% of both chambers).[3] For robustness, we also gathered a sample of U.S. policy nonprofits. Our search for 50 such organizations most frequently cited in major U.S. media (Groseclose & Milyo, 2005) produced 33 accounts, consisting of think tanks such as the RAND Corporation and policy groups such as the Sierra Club.

Research shows that the perceived partisanship of news media has a strong effect on who consumes it, with audiences generally preferring news media that is consistent with their views (Iyengar & Hahn, 2009; Stroud, 2008). Similarly, we expect that, on average, we can infer the political orientation of a user from the ideological positions of the hubs they follow. Golbeck and Hansen (2014) validated this approach by examining the Twitter postings of users who follow members of Congress, finding that the ideological scores of politicians reliably predicted their Twitter followers' presidential election vote choices and preferences for ideological news. As we discussed above, Barberá (2015) also showed that the Twitter tie structure between followers of legislators closely reflects the relative ideological positions of these politicians.[4] Thus, our hypotheses suggest that audiences of more conservative or extreme political hubs may follow one another at greater rates than those of more liberal or moderate ones.

Data

Our primary data come from a publicly available Twitter dataset created by Kwak, Lee, Park, and Moon (2010), which contains a complete snapshot of the publicly visible Twitter network from June 2009 (over 40 million nodes and 1.47 billion ties). The dataset consists of only the network structure itself, with no information about the nodes beyond their Twitter account numbers. We linked our hubs to their offline identities via data retrieved from Twitter servers and calculated all network measures via custom *MySQL* routines.

We use archival data from 2009 because it crucially predates Twitter's "Who to Follow" feature. Since July 2010, this feature has encouraged users to follow the same accounts as their alters, thus nudging them towards greater homophily. As of May 2013, it was responsible for the creation of over

[3] Like Golbeck and Hansen (2014), we excluded John McCain, as his recent candidacy for president gave him a categorically different Twitter presence. We also dropped hubs with less than 100 followers, since they were not prominent enough to be properly considered hubs.

[4] The validity of using ideological positions of legislators to proxy those of their constituents has been the subject of recent critiques which point out that legislators tend to be more ideologically extreme than members of the general public and that nonideological factors affect electoral outcomes (e.g., Bafumi & Herron 2010; Enns & Koch 2013). However, we note that, while any individual has little control over who represents her in Congress, she can choose whether to follow any legislator on Twitter and can follow any number of legislators. This greatly increased freedom of choice sets Twitter-based measures apart from those criticized in the literature.

a million Twitter ties *per day* (Gupta et al., 2013), rendering Twitter data gathered after 2010 less suitable for studying homophily.

We draw on a number of further sources for additional information on political hubs. For legislators, we use the Congressional Committee Assignments dataset (Stewart & Woon, 2011) and election results from the CQ Press Voting and Elections Collection (2010). For their constituencies, we use state- and congressional district-level information from the U.S. Census Bureau's American Community Survey (2009), Current Population Survey (2009), and the decennial census (2000). For nonprofits, we use publicly available 2010–11 tax returns filed with the Internal Revenue Service (IRS Form 990) and background information from *GuideStar* nonprofit reports (2015).

Measures

Political orientation. For our measure of legislator ideology, we utilize DW-Nominate scores computed from voting rolls for the 111th U.S. Congress (Carrol et al., 2011; Poole & Rosenthal, 2007). The primary dimension of these scores closely corresponds to the liberal-conservative dimension in U.S. politics (Poole & Rosenthal, 2007), ranging from roughly -1 (liberal) to 1 (conservative). We divide this dimension by its standard deviation.

We use the ideological scores of senators and representatives as proxies for the orientations of their followers. Like Golbeck and Hansen (2014), for the 31% of users in this sample who follow two or more hubs, we average these hubs' scores.[5] Since we have separate hypotheses concerning Left-Right and Center-Extremes differences, we decompose the scaled score into its magnitude and direction components:

$$score = |score| * (-1)^{\mathbb{1}(score>0)} = magnitude * (-1)^{direction}.$$

We define the *ideological extremity* variable as the magnitude of the ideological score, which is equal to $|score|$. It captures how far a given politician is from the ideological center. We then define the *ideological conservatism* variable as the direction component of the score, which is $\mathbb{1}(score > 0)$. It represents whether the hub is conservative (1) or liberal (0).

For nonprofits, we use ideological scores which Groseclose and Milyo (2005) computed by counting the number of times each nonprofit was cited in floor speeches by members of the 103rd through 107th Congresses and then averaging the ADA ratings of the citing members. We linearly translate this measure to range from -1 to 1, thus placing it on roughly the same scale as DW-Nominate scores to make results visually comparable across both datasets.[6] We then follow the above procedures to create our conservatism and extremity measures.

Homophily. Twitter is both a social-networking platform and a media service (Kwak et al., 2010). Its network structure consists of ties of attention. Some ties are interpersonal, connecting regular users with one another, and are thus akin to ties of friendship or acquaintanceship. Others connect regular users to hubs belonging to major public figures or organizations and thus more closely resemble connections between audience members and celebrities or media sources. This makes the observed Twitter network essentially two-mode. Though the gradient-like nature of Twitter popularity sometimes makes these modes difficult to distinguish in practice, we can imagine setting aside hub nodes and examining the interpersonal (user–user) network apart from the audience (user–hub) network.

[5] We tested three alternate coding schemes to rule out potential biases from averaging. Because means cancel extremes, we tested the extremity score of the *most extreme* and of a *randomly chosen* hub a user follows. Since means are affected by outliers, we also used the median of the hubs they follow. Results supported the same substantive conclusions we report here.

[6] The two scores, however, remain distinct enough that we do not combine the populations in any of our regression analyses.

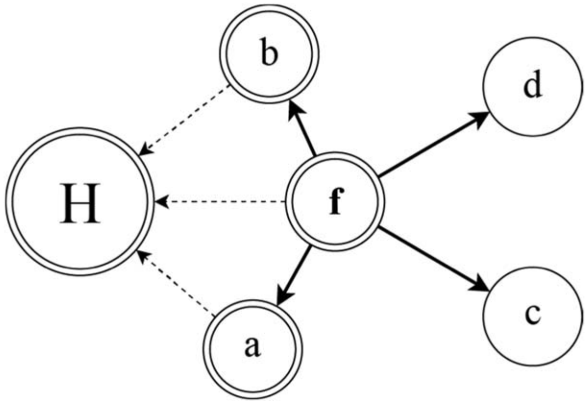

Figure 1. In addition to the tie connecting follower f to the hub H, f has a total of four outgoing ties ($f \rightarrow a,\ f \rightarrow b,\ f \rightarrow c,\ f \rightarrow d$). Since nodes a and b also follow hub H, while c and d do not, the homophily of f with regard to H is $\boldsymbol{o}(f, H) = 100\% * 2/4 = 50\%$.

Our homophily measure follows this reasoning. For any follower f of political hub H, we define her "homophily with regard to H" as the percentage of other accounts followed by f that in turn also follow H. A user's homophily with regard to a political figure they follow is thus equal to the percentage of other users they follow who also follow this figure (see Figure 1). If V is the set of all public nodes on Twitter, and $T_{ij} = 1$ if directed tie i→j exists and 0 otherwise, this equals

$$\boldsymbol{o}(f, H) = \frac{\sum_{g \in V} T_{fg} T_{gH}}{\sum_{g \in V} T_{fg} - 1} * 100\%$$

This measure ranges from 0 to 100%. As this measure is undefined for nodes with no outgoing ties except the tie to hub H, we drop such nodes $(N = 2,665)$ from our sample.

This homophily measure has two advantages. The first is conceptual. If hub H posts a political message on Twitter, 50% of f's alters would also receive a copy of this message. Since users frequently forward the messages they receive to their own followers (Kwak et al., 2010), this measure captures the ceteris paribus chances of f receiving additional exposures to H's message, which fits with the conception of homophilous ties as building blocks of echo chambers. Recent analyses of diffusion on Twitter confirm that such multiple exposures promote the spread of political messages, as expected from diffusions which spread through normative influence (Romero et al., 2011). Thus, $\boldsymbol{o}(f, H)$ should capture a property of network structure that directly corresponds to its ability to sustain normative diffusions.

The second advantage of $\boldsymbol{o}(f, H)$ is methodological. We can enumerate all of the followers of hub H to determine the exact count of f's *potential* homophilous partners. This gives us the rare opportunity to control statistically for differences in the baseline rates of network homogeneity expected under random mixing, which are different for populations of different sizes.

Reciprocity. The measure $\boldsymbol{o}(f, H)$ has a potential to confound homophily with reciprocity. Imagine the following scenario: Followers of hub A are as likely as followers of hub B to initiate new homophilous ties with one another; however, followers of B are more likely to reciprocate homophilous ties once they are initiated by another follower. As a result, followers of B have higher average homophily scores than followers of A in spite of not actually being more homophilous in the sense we use here. To deal with this confound, we introduce a reciprocity measure to capture an actor's tendency to reciprocate incoming homophilous ties from other hub followers. For each follower f of hub

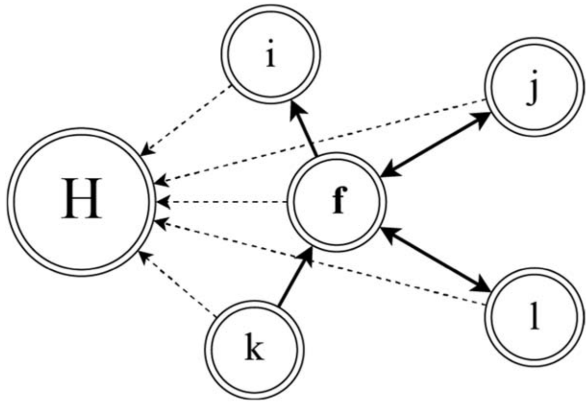

Figure 2. In addition to follower f, hub H has four followers i, j, k and l. Follower f receives homophilous ties from j, k and l, and sends reciprocal homophilous ties to j and l. The reciprocity of f with regard to H is thus $r(f, H) = 2/3 \approx 0.67$.

H, this measure equals the count of ties received by f from other followers of H that f reciprocates by sending a tie back, divided by the total count of such ties received by f (see Figure 2):

$$\boldsymbol{r}(f, H) = \frac{\sum_{g \in V} T_{gH} T_{fg} T_{gf}}{\sum_{g \in V} T_{gH} T_{gf}}.$$

We standardize $\boldsymbol{r}(f, H)$ and set it to 0 for all nodes with no incoming homophilous ties.

Characteristics of members of Congress. Legislators differ not only in their ideological orientation, but also in their prominence, characteristics of their constituencies, and various other ways that may influence the rates at which their followers follow one another. Since geographical proximity significantly increases the probability of a social tie between two people both offline (Butts, 2003) and on Twitter (Takhteyev et al., 2012), factors that affect the physical dispersion of followers make important potential confounds.

The geographic concentration of the legislator's constituents is perhaps the most obvious such factor. We expect that, the higher the constituency's population, the larger portion of a legislator's followers are constituents. Higher population densities and rates of internet usage should result in increased proximity between online constituents, which may in turn yield more densely connected followers. Controlling for these three quantities should account for much of this confound.[7]

Conversely, while many members of Congress may be known largely within their constituencies, those occupying more prominent positions within the party or chamber have more diverse and geographically dispersed audiences, resulting in lower tie densities. We control for a number of major sources of differences in prominence with terms for chamber ("senator"), leadership ("chamber/party" and "committee"), and number of years in chamber ("seniority").

Differences in mobilization present a different potential confound, as highly mobilized users may participate in more political activities offline as well as online, and may thus encounter more opportunities to form politically homophilous ties. In June 2009, the emerging Tea Party movement could have had this effect on some conservatives. This mobilization also resulted in the creation of Tea Party caucuses, formed during the 111th Congress in the House and at the start of the 112th Congress

[7] To estimate district-level Internet usage from 2009 state-level data, we use each district's rural/urban populations to take a weighted average of the state's rural/urban internet usage rates.

in the Senate. Since the Tea Party itself has no clear member rolls or membership criteria, we use memberships in these caucuses as a proxy for Tea Party affiliation. To partially account for other differences in mobilization, we also control for electoral victory margin.

Because conservative and liberal legislators belong to two different political parties, large-scale organizational differences between the two parties represent a potential source of confounds for the liberal-conservative effect we estimate in this article. There are a variety of forms these organizational confounds could take. For example, ties between supporters of one party may in part be a product of efforts by campaigns and party organizations to organize and engage their supporters. Since greater political engagement on the part of followers would mean more opportunities to form politically homophilous ties, greater organizational success by one political party could translate to more ties among its supporters. As party and ideology are confounded in this sample, these differences in tie density could be confused for the predicted effect of ideology.

These data do not allow us to fully rule out this confound. We can, however, replicate our analysis on political nonprofit organizations, where this confound is likely to be weaker. Though the nonprofits have ties to the same political milieu as the politicians, they are further removed from the organizational structure of political parties and the social environment of electoral politics. If the relationship between conservatism and homophily were due primarily to differences between the parties, we would expect the association to be weaker for followers of nonprofits than for followers of politicians. Conversely, an equivalent or stronger estimated coefficient for conservatism would suggest that it is unlikely to be accounted for by this organizational explanation.

Characteristics of nonprofits. Like members of Congress, nonprofits vary in ways which may affect ties between their followers. First, since followers of less prominent nonprofits may be a more specialized audience than those of more prominent nonprofits, they may also form ties to one another at higher rates. To control for this confound, we add covariates for total budget and age of the nonprofit. Conversely, each nonprofit's staff likely forms a portion of its online audience, and this portion may be higher for organizations with larger staffs. Since shared workplaces are a likely basis for social tie formation (Feld, 1982), such nonprofits may appear to have more homophilous followers than smaller organizations. To address this, we control for each organization's employee and volunteer counts.[8]

While some nonprofits are single organizations, others have separately incorporated regional or special-purpose affiliates. For example, the NAACP's name is shared by regional chapters, a series of day-care centers, and other specialized affiliates. Organizations with a more dispersed geography and mission could have more diverse, less homophilous audiences. Conversely, Washington, DC's large and active nonprofit sector may grant Washington-based nonprofits densely connected audiences, especially if they engage in lobbying. We therefore also add dummy terms for regional affiliates,[9] Washington headquarters, and ability to engage in unrestricted lobbying activity.[10]

Results

We located hubs belonging to 31 senators, 128 representatives, and 33 nonprofits, yielding a total of 238,943 unique followers (see Table 1). These followers were closely divided between liberal and conservative hubs, with 131,576 unique followers of the former and 133,210 of the latter. But, while

[8] To compensate for a substantial positive skew in budget and volunteer counts, we use their logarithms.

[9] We located affiliates using the Foundation Center's IRS record database (990finder.foundationcenter.org).

[10] All organizations in our sample are 501(c)(3) nonprofits, which can take unlimited tax-deductible donations but can only engage in limited lobbying. Roughly half are simultaneously incorporated as 501(c)(4)/(6) nonprofits, which take taxable donations but face no lobbying limits. For such dually incorporated nonprofits, we use the sum of their 501(c)(3) and 501(c)(4)/(6) budgets. When their staff counts differed, we used the greater of the two.

Table 1. Descriptive Statistics

	Liberal	Conservative	Total Unique
Senator Hubs	14	17	31
Unique followers of senators	53919	51272	96022
Mean ideological score	−0.856	0.984	0.153
House Member Hubs	48	80	128
Unique followers of House members	27846	78314	98597
Mean ideological score	−0.697	1.224	0.504
Political Nonprofit Hubs	18	15	33
Unique followers of nonprofits	63202	45399	103864
Mean ideological score	−0.902	0.953	−0.059
Total unique followers	131576	133210	238943

Note. The "Total Unique" column does not generally equal the sum of the "Liberal" and "Conservative" columns because some Twitter users follow both liberal and conservative hubs.

the total number of conservatives and liberals in the sample was similar, they followed different categories of hubs. There were more followers of Republican than Democratic legislators, but more followers of liberal than conservative nonprofits. Since our substantive findings do not differ between congressional and nonprofit hubs, these different distributions do not appear to affect our results.

Across all followers in both samples, we find an average homophily rate $\boldsymbol{o}(f, H)$ of 11.0%. Therefore roughly one out of nine alters of each user in our sample follows the same political figure as that user. As we demonstrate in Appendix A in the online supporting information, this rate is substantially higher than would be expected under random mixing and is thus consistent with the presence of political homophily in the sample.

Followers of Members of Congress

We employ multivariate linear regression with two-way clustered standard errors to test our hypotheses. Our main dependent variable is homophily, $\boldsymbol{o}(f, H)$. Our primary independent variables are ideological conservatism and extremity. Because the ideologies of legislators and nonprofits are captured with different measures, we analyze their followers separately.

We control for chamber with the "senator" dummy. We also control for our reciprocity measure $r(f, H)$. To account for the levels of homophily expected from baseline mixing in differently sized populations, we add the covariate *hub network size*, which is the count of the hubs' followers minus one. Since random mixing produces baseline homophily rates that are a linear function of the number of potential homophilous partners (see Appendix A), we can use this quantity to control for baseline mixing in a regression framework. There are also potential differences between new versus more experienced users of Twitter. Since Twitter account identifiers were assigned sequentially and the growth of Twitter's user base was roughly exponential through 2009, we expect the age of a Twitter account to be roughly proportional to the logged count of accounts with greater identifiers: $age(id_f) \propto \log(max\ id - id_f)$. We use this to create controls for *account age* and *hub age*.[11] We also include a covariate for *ego network size*. We standardize all nondummy controls.

Since 31% of users in our data follow multiple hubs, and each follower f has a different value of $\boldsymbol{o}(f, H)$ for each hub she follows, our unit of analysis is the user-hub pairing. There are two nonnested levels of nonindependence in these data. Some variables are nonindependent for followers of one hub and others for all observations of one follower. Without accounting for this nonindependence, our standard errors could be significantly underestimated, as the true amount of independent data points

[11] While it is possible to retrieve the exact account creation dates from the Twitter API, the number of nodes in our sample would make that approach extremely costly.

Table 2. Followers of Members of Congress: Effect of Ideology on Homophily

	Model 1	Model 2
Nonindependent Within Hubs ($N = 159$)		
Ideological conservatism	3.802*** (0.764)	3.251*** (0.646)
Ideological extremity	4.533*** (0.701)	4.057*** (0.556)
Hub network size	3.725*** (0.774)	3.138*** (0.688)
Hub age	1.141*** (0.301)	1.142*** (0.283)
Senator	−2.171** (0.708)	−1.244 (0.966)
Number of constituents	—	−0.118 (0.352)
Const. population density	—	−0.203 (0.134)
Const. Internet access	—	−0.252 (0.277)
Seniority	—	−0.454 (0.4)
Committee leadership	—	−0.253 (0.944)
Party leadership	—	2.151 (1.433)
Victory margin	—	0.32** (0.1)
Tea Party caucus	—	1.85# (1.116)
Nonindependent Within Followers ($N = 160{,}452$)		
Account age	0.129 (0.108)	0.234* (0.117)
Ego network size	−1.193**(0.389)	0.145 (0.102)
Reciprocity	0.269* (0.114)	−1.176** (0.382)
(Intercept)	4.768*** (1.054)	4.762*** (0.93)
R^2	0.161	0.168

Note. Results from linear regressions with two-way clustered standard errors (in parentheses). All nondummy predictors have $\sigma=1$.
#$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. (two-tailed.)

for our primary independent variables is closer to the number of hubs ($N=159$) than it is to the number of observations ($N=383{,}292$). We employ multiway clustered standard errors to prevent this underestimation (Petersen, 2009; Thompson, 2011), clustering by hub and follower. This accounts for nonindependence but also results in significance levels heavily constrained by the number of hubs.

The results of this analysis are in Model 1 of Table 2. They show that ideological conservatism was strongly and positively associated with homophily rate ($\beta=3.80, p < 0.001$). Thus, with all other variables held constant, conservatives' homophily rates were 3.8 points higher than those of liberals, offering support for our first hypothesis. We also found that ideological extremity has a strong positive association with homophily ($\beta=4.53, p < 0.001$), indicating that a one standard deviation increase in extremity was associated with a 4.5-point increase in homophily. Since our average observed homophily rate is 11.2%, these differences are both statistically and substantively significant, thus lending support to our hypotheses.

We next repeated this analysis with the addition of further controls for characteristics of legislators and constituencies we described above. Results are in Model 2 of Table 2. Consistent with the above conjectures, we found a marginal positive effect for Tea Party membership ($\beta=1.85, p < 0.10$) and a weak but significant positive effect for margin of election victory ($\beta=0.32, p < 0.01$). However, the previously significant Senator dummy decreased in magnitude and significance ($\beta=-1.24, n.s.$), indicating that these covariates may partially account for the lower homophily rates among followers of senators. Both ideological conservatism and ideological extremity decreased slightly in magnitude but remained highly significant ($\beta=3.25, p < 0.001$ and $\beta=4.06, p < 0.001$, respectively), again supporting our hypotheses.

Followers of Nonprofits

To demonstrate the robustness of our findings and address concerns about the potential confound between partisanship and ideology, we replicated our analyses with followers of major policy

Table 3. Followers of Policy Nonprofits: Effect of Ideology on Homophily

	Model 1	Model 2
Nonindependent Within Hubs ($N = 33$)		
Ideological conservatism	7.456** (2.416)	6.305*** (1.597)
Ideological extremity	4.316** (1.456)	4.332*** (1.105)
Hub network size	0.629 (1.183)	0.923 (1.022)
Hub age	1.346** (0.416)	0.933* (0.452)
Annual Budget	—	0.976# (0.545)
Number of employees	—	0.408 (0.724)
Number of volunteers	—	0.258 (0.697)
Affiliated organizations	—	−7.665** (2.707)
Years since founded	—	−0.329 (0.633)
DC area	—	3.067* (1.506)
Unrestricted lobbying	—	2.452 (2.177)
Nonindependent Within Followers ($N = 103{,}864$)		
Account age	0.29 (0.346)	0.267 (0.328)
Ego network size	−1.02* (0.518)	−1.04* (0.527)
Reciprocity	1.152** (0.388)	1.103* (0.43)
(Intercept)	3.303* (1.367)	5.188* (2.153)
R^2	0.151	0.185

Note. Results from linear regressions with two-way clustered standard errors (in parentheses). All nondummy predictors have $\sigma=1$.
#$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ (two-tailed.)

nonprofits. The results are given in Table 3. Model 1 replicates our initial test of ideological conservatism and extremity effects. Due to the much smaller hub-level population $(N=33)$, cluster-adjusted standard errors are larger than for congressional hubs, reducing the statistical significance of most coefficients. Nonetheless, ideological conservatism is still significantly associated with higher homophily $(\beta=7.46, p < 0.01)$, an effect of greater size than for congressional hubs. Ideological extremity is also still significantly associated with higher homophily at roughly the same magnitude of effect as in the previous sample $(\beta=4.32, p < 0.01)$.

With additional controls for further characteristics of the nonprofits (Model 2 of Table 3), the effect of conservatism decreased slightly in magnitude but increased in significance $(\beta=6.31, p < 0.001)$. The effect of extremity also increased in significance $(\beta=4.33, p < 0.001)$. The analysis of followers of nonprofits hubs thus provides additional support for our hypotheses. Furthermore, the conservatism coefficients in both models were larger in magnitude than those in Table 2, providing some evidence against the alternate organizational explanation for these effects.

Figure 3 illustrates these homophily differences. It depicts ties between the followers of Cato Institute (conservative) and Amnesty International (liberal), two well-known nonprofits with comparable numbers of followers (10,298 and 10,638). Their mean-follower homophily rates are 13.04% and 6.75%, respectively, a difference similar to the estimated conservatism effect.[12] The force-directed layout places connected nodes closer and disconnected ones further apart, leading the cluster of Cato Institute's followers to appear visibly smaller because of its higher homophily.

We performed additional analyses with both samples to further establish the significance and robustness of our results (Appendix B in the online supporting information). First, we confirmed that the same conclusions obtain when our analyses are restricted to stronger ties, which may play a more

[12] Note that this comparison is imperfect, as these organizations necessarily differ on more characteristics than ideology alone. For example, Cato brands itself as "libertarian" rather than "conservative," and its agenda may be narrower than Amnesty's in ways that could encourage political homophily. While we cannot know whether the homophily differences between these particular groups are due to the effects of conservatism we estimated in our models as opposed to some unobserved features unique to these groups, their similar magnitude makes them a reasonable visual illustration.

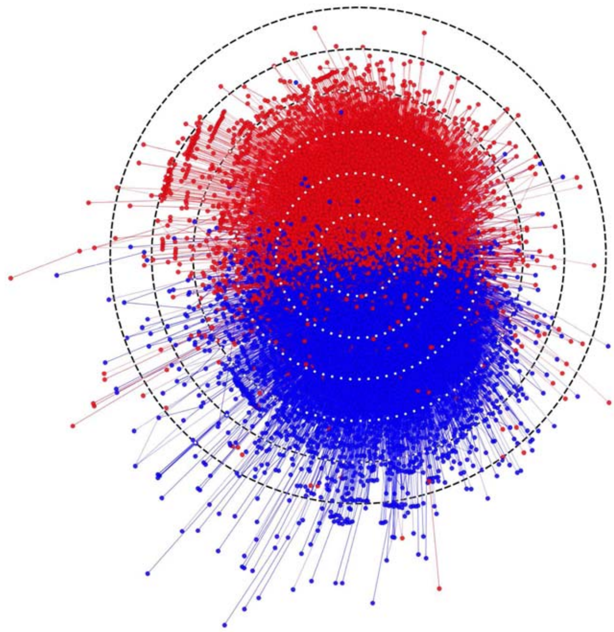

Figure 3. Network of followers of Cato Institute (conservative; top) and Amnesty International (liberal; bottom) that receive at least one tie from another follower in the figure, with isolated node pairs omitted (node colors correspond to followed hubs). The "spring-based" layout places nodes closer if they are connected by a tie, and further apart if they are not, revealing a clear segregation of the network into two parts as expected under conditions of strong political homophily. Though the follower counts in the two networks differ by only 3%, the Cato Institute network is far denser, resulting in a tighter clustering of nodes and thus a visibly smaller size of the densely connected component (compare edge of component against radial grid).

important role than weaker ties for diffusing social influence (Bond et al., 2012; González-Bailón et al., 2011; McAdam & Paulsen, 1993). To ensure robustness to modeling assumptions, we also examined these homophily differences using randomization inference and propensity score matching, finding support for both hypotheses. Additionally, we leveraged the known Twitter account age of all individuals in our sample to test for a possible reverse causal relationship between homophily and ideological extremity (Appendix C in the online supporting information). Results indicated that such a relationship is unlikely to explain the results we report here.

Discussion

We drew on research from political psychology in hypothesizing that more conservative or ideologically extreme individuals would be more homophilous than more liberal or moderate ones. To avoid methodological challenges faced by most survey studies of political homophily, we tested these hypotheses with 238,943 ego networks belonging to followers of political actors on Twitter. These analyses yielded consistent and robust evidence for both hypotheses.

If the same homophily differences we observed with these Twitter networks extend to social networks more generally, the higher homophily rates of more conservative and ideologically extreme individuals could have significant consequences for the emergent dynamics of their respective political networks. These rates should, ceteris paribus, result in networks that embed their members in denser webs of like-minded associations, which could then insulate individuals from the demotivating effects of dissenting views, and may enable political behaviors to spread faster than they would through sparser networks. Our results thus suggest that homophily might provide a structural advantage to the mobilization of right-wing or politically extreme social movements relative to left-wing or moderate ones. We would similarly expect the negative effects of network homogeneity on tolerance and understanding to be unevenly distributed.

While the status of nonelite political polarization in the United States is still debated (e.g., Fiorina & Abrams, 2008), we note that our findings of left-right differences in homophily fit with accounts that find polarization to be asymmetrically tilted towards the political right (e.g., Butler, 2009). Moreover, they recall Sunstein's (2002) warning that social networks may amplify polarization: If homophily is found at higher levels at the extremes of the ideological distribution, the resulting concentration of homogenizing and mobilizing influence could push extreme attitudes ever further away from the center.

ACKNOWLEDGMENTS

This research was supported by fellowships from National Science Foundation Graduate Research Fellowship Program and Interdisciplinary Graduate Education and Research Traineeship Program. We thank Claude Fischer, Michael Hout, Fabiana Silva, Stephen Vaisey, and participants of the Berkeley Mathematical, Analytical and Experimental Sociology working group for feedback on the article. Correspondence concerning this article should be addressed to Andrei Boutyline at Department of Sociology, 410 Barrows Hall, University of California, Berkeley, CA 94720. E-mail: boutyline@berkeley.edu

REFERENCES

Ackland, R., & Shorish, J. (2009). Network formation in the political blogosphere: An application of agent based simulation and e-research tools. *Computational Economics*, *34*(4), 383–398.

Adamic, L., & Glance, N. (2005). The political blogosphere and the 2004 U.S. election: Divided they blog. *Proceedings of the 3rd International Workshop on Link discovery, LinkKDD '05* (pp. 36–43).

Adorno, T., Frenkel-Brunswik, E., Levinson, D., & Sanford, R. N. (1950). *The authoritarian personality*. Oxford, United Kingdom: Harpers.

Arendt, H. (1961). *Between past and future eight exercises in political thought*. Rev. ed. 1968. New York, NY: Viking Press.

Bafumi, J., & Herron, M. (2010). Leapfrog representation and extremism: A study of American voters and their members in Congress. *American Political Science Review*, *104*(03), 519–542.

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. *Political Analysis*, *23*(1), 76–91.

Blau, P. (1977). *Inequality and heterogeneity: A primitive theory of social structure*. New York, NY: Free Press.

Bond, R. M., et al. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–98.

Butler, D. M. (2009). The effect of the size of voting blocs on incumbents' roll-call voting and the asymmetric polarization of Congress. *Legislative Studies Quarterly*, *34*(3), 297–318.

Butts, C. (2003). Predictability of large-scale spatially embedded networks. In R. Breiger, K. Carley, & P. Pattison (Eds.), *Dynamic social network modeling and analysis: Workshop summary and papers* (pp. 313–323). Washington, DC: National Academies Press.

Campbell, D. (2013). Social networks and political participation. *Annual Review of Political Science*, *16*(1), 33–48.

Carrol, R., et al. (2011). DW-NOMINATE scores with bootstrapped standard errors. Retrieved from http://www.voteview.com/dwnomin.htm

Centola, D. (2013). The spread of behavior in an online social network experiment. *Science*, *329*(5996), 1194.

Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, *113*(3), 702–34.

CQ Press. (2010). CQ press voting and elections collection. Retrieved from http://library.cqpress.com/elections/advsearch/election-results.php

DiGrazia, J., McKelvey, K., Bollen, J., & Rojas, F. (2013). More tweets, more votes: Social media as a quantitative indicator of political behavior. *PLoS ONE*, *8*(11):e79449.

DiPrete, T., Gelman, A., McCormick, T., Teitler, J., & Zheng, T. (2011). Segregation in social networks based on acquaintanceship and trust. *American Journal of Sociology*, *116*(4), 1234–1283.

Dunbar, R. I. M., Arnaboldi, V., Conti, M., & Passarella, A. 2015. The structure of online social networks mirrors those in the offline world. *Social Networks*, *43*, 39–47.

Enns, P., & Koch, J. (2013). Public opinion in the U.S. states 1956 to 2010. *State Politics & Policy Quarterly*, *13*(3), 349–372.

Eveland, W., & Hively, M. H. (2009). Political discussion frequency, network size, and "heterogeneity" of discussion as predictors of political knowledge and participation. *Journal of Communication*, *59*(2), 205–224.

Feld, S. (1982). Social structural determinants of similarity among associates. *American Sociological Review*, *47*(6), 797–801.

Fiorina, M., & Abrams, S. (2008). Political polarization in the American public. *Annual Review of Political Science*, *11*(1), 563–588.

General Social Survey (GSS). (2006). [machine-readable data file]. Principal Investigator: T. W. Smith. Co-Principal Investigators: P. V. Marsden & M. Hout, Michael. Sponsored by National Science Foundation. National Opinion Research Center, Chicago, IL.

Gerber, A., Huber, G., Doherty, D., & Dowling, C. (2012). Disagreement and the avoidance of political discussion: Aggregate relationships and differences across personality traits. *American Journal of Political Science*, *56*(4), 849–874.

Goel, S., Mason, W., & Watts, D. (2010). Real and perceived attitude agreement in social networks. *Journal of Personality and Social Psychology*, *99*(4), 611–621.

Golbeck, J., & Hansen, D. (2014). A method for computing political preference among Twitter followers. *Social Networks*, *36*, 177–184.

González-Bailón, S., Borge-Holthoefer, J., Rivero, A., & Moreno, Y. (2011). The dynamics of protest recruitment through an online network. *Scientific Reports*, 1, 1–7.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Greenberg, J., & Jonas, E. (2003). Psychological motives and political orientation—The left, the right, and the rigid: Comment on Jost et al.(2003). *Psychological Bulletin*, *129*(3), 376–382.

Groseclose, T., & Milyo, J. (2005). A measure of media bias. *Quarterly Journal of Economics*, *120*(4), 1191–1237.

GuideStar. (2015). Nonprofit reports and forms 990 for donors, grantmakers, & businesses. Retrieved from http://www.guidestar.org/

Gupta, P. (2013). WTF: The who to follow service at Twitter. Proceedings of the 22nd International Conference on World Wide Web (Pp. 505–514).

Hatemi, P., McDermott, R., Eaves, L., Kendler, K., & Neale, M. (2013). Fear as a disposition and an emotional state: A genetic and environmental approach to out-group political preferences. *American Journal of Political Science*, *57*(2), 279–293.

Hemphill, L., Otterbacher, J., & Shapiro, M. (2013). What's Congress doing on twitter? In Proceedings of the 2013 Conference on Computer Supported Cooperative Work, CSCW '13 (pp. 877–886). New York, NY: ACM.

Hoffer, E. (1951). *The true believer: Thoughts on the nature of mass movements*. New York, NY: HarperCollins.

Hogg, M. (2004). Uncertainty and extremism: Identification with high entitativity groups under conditions of uncertainty. In *The psychology of group perception: Perceived Variability, entitativity, and essentialism* (pp. 401–408). New York, NY: Psychology Press.

Hogg, M. (2007). Uncertainty–identity theory. *Advances in Experimental Social Psychology*, *39*, 69–126

Hohman, Z., Hogg, M., & Bligh, M. (2010). Identity and intergroup leadership: Asymmetrical political and national identification in response to uncertainty. *Self and Identity*, *9*(2), 113–128.

Huckfeldt, R., Beck, P. A., Dalton, R., & Levine, J. (1995). Political environments, cohesive social groups, and the communication of public Opinion. *American Journal of Political Science*, *39*(4), 1025–1054.

Huckfeldt, R., Mendez, J. M., & Osborn, T. (2004). Disagreement, ambivalence, and engagement: The political consequences of heterogeneous networks. *Political Psychology*, *25*(1), 65–95.

Huntington, S. (1957). Conservatism as an ideology. *American Political Science Review*, *51*(2), 454–473.

Iyengar, S., & Hahn, K. (2009). Red media, blue media: Evidence of Ideological Selectivity in media use. *Journal of Communication*, *59*(1), 19–39.

Jang, S.-J. (2009). Are diverse political networks always bad for participatory democracy? Indifference, alienation, and political disagreements. *American Politics Research*, *37*(5), 879–898.

Jetten, J., Hogg, M., & Mullin, B.-A. (2000). In-group variability and motivation to reduce subjective uncertainty. *Group Dynamics: Theory, Research, and Practice*, *4*(2), 184–198.

Jost, J. (2006). The end of the end of ideology. *American Psychologist*, *61*(7), 651–70.

Jost, J., Glaser, J., Kruglanski, A., & Sulloway, F. (2003a). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375.

Jost, J., Glaser, J., Kruglanski, A., & Sulloway, F. (2003b). Exceptions that prove the rule—using a theory of motivated social cognition to account for ideological incongruities and political anomalies: Reply to Greenberg and Jonas (2003). *Psychological Bulletin*, *129*(3), 383–393.

Jost, J., Federico, C., & Napier, J. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology*, *60*(1), 307–337.

Kim, H., & Bearman, P. (1997). The structure and dynamics of movement participation. *American Sociological Review*, *62*(1), 70–93.

Klofstad, C., Sokhey, A., & McClurg, S. (2013). Disagreeing about disagreement: How conflict in social networks affects political behavior. *American Journal of Political Science*, *57*(1), 120–134.

Knoke, D. (1990). Networks of political action: Toward theory construction. *Social Forces*, *68*(4), 1041–1063.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What Is Twitter, a social network or a news media? Proceedings of the 19th International Conference on World Wide Web (Pp. 591–600).

Laumann, E. (1969). Friends of urban men: An assessment of accuracy in reporting their socioeconomic attributes, mutual choice, and attitude agreement. *Sociometry*, *32*(1), 54–69.

Lazer, D., Rubineau, B., Chetkovich, C., Katz, N., & Neblo, M. (2010). The coevolution of networks and political attitudes. *Political Communication*, *27*(3), 248–274.

Liu, T.-P., Dai, S.-C., & Wu, C.-L. (2013). Cross-cutting networks and political participation: Lessons of the 2010 city mayoral elections in Taiwan. *East Asia*, *30*(2), 91–104.

McAdam, D., & Paulsen, R. (1993). Specifying the relationship between social ties and activism. *American Journal of Sociology*, *99*(3), 640–667.

McPherson, M., & Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review*, *52*(3), 370–379.

Mehrabian, A. (1975). Affiliation as a function of attitude discrepancy with another and arousal-seeking tendency. *Journal of Personality*, *43*(4), 582–592.

Mill, J. S. (1859). On liberty. London, United Kingdom: J. W. Parker and Son.

Mutz, D. (2002a). Cross-cutting social networks: Testing democratic theory in practice. *American Political Science Review*, *96*(1), 111–126.

Mutz, D. (2002b). The consequences of cross-cutting networks for political participation. *American Journal of Political Science*, *46*(4), 838–855.

Mutz, D. (2006). *Hearing the other side: Deliberative versus participatory democracy*. Cambridge, United Kingdom: Cambridge University Press.

Obar, J. A., Zube, P., & Lampe, C. (2012). Advocacy 2.0: An analysis of how advocacy groups in the United States perceive and use social media as tools for facilitating civic engagement and collective action. *Journal of Information Policy*, *2*, 1–25.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, *22*(1), 435–480.

Poole, K., & Rosenthal, H. (2007). Ideology and Congress. New Brunswick, NJ: Transaction.

Price, V., Cappella, J., & Nir, L. (2002). Does disagreement contribute to more deliberative opinion? *Political Communication*, *19*(1), 95–112.

Rainie, L. (2010). *Internet, broadband and cell phone statistics*. Washington, DC: Pew Research Center.

Romero, D., Meeder, B., & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on Twitter. Proceedings of the 20th International Conference on World Wide Web (pp. 695–704).

Shah, J., Kruglanski, A., & Thompson, E. (1998). Membership has its (epistemic) rewards: Need for closure effects on in-group bias. *Journal of Personality and Social Psychology*, *75*(2), 383–393.

Smith, A., & Brenner, J. (2012). *Twitter use 2012*. Washington, DC: Pew Research Center.

Smith, A., & Rainie, L. (2010). *8% of Online Americans use Twitter*. Pew Research Center.

Stewart, C., III, & Woon, J. (2011). Congressional committee assignments, 103rd to 112th Congresses, 1993–2011.

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

Sunstein, C. (2002). The law of group polarization. *Journal of Political Philosophy*, *10*(2), 175–195.

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. *Social Networks*, *34*(1), 73–81.

Thompson, S. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, *99*(1), 1–10.

United States Census Bureau. (2000). P2: Urban and rural. *2000 Census*. Washington, D.C.

United States Census Bureau. (2009). B01003: Total population. *2009 American Community Survey*. Washington, D.C.

United States Census Bureau. (2009). Table 1155. Household Internet usage by type of Internet connection and state: 675 2009. Current Population Survey. Washington, D.C.

Valenzuela, S, Kim, Y., & de Zúñiga, H. G. (2012). Social networks that matter: Exploring the role of political discussion for online political participation. *International Journal of Public Opinion Research*, *24*(2), 163–184.

Visser, P., & Mirabile, R. (2004). Attitudes in the social context: The impact of social network composition on individual-level attitude strength. *Journal of Personality and Social Psychology*, *87*(6), 779–795.

Wimmer, A., & Lewis, K. (2010). Beyond and below racial homophily. ERG models of a friendship network documented on Facebook. *American Journal of Sociology*, *116*(2), 583–642.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Appendix A: Baseline Models of Homophily

Appendix B: Additional Analyses

Appendix C: Reverse Causation