ORIGINAL ARTICLE

The Organizational Principles of Online Political Discussion: A Relational Event Stream Model for Analysis of Web Forum Deliberation

Hai Liang

Department of Media and Communication, City University of Hong Kong, Hong Kong

The structure of online political discussion has proven important to deliberative democracy. However, the organizational mechanisms of the structure receive little attention in scholarship. This study employed a random-effects relational event model to differentiate and examine the effects of a set of organizational principles in web forum discussions. By analyzing more than 175,000 forum replies, the study found that cross-ideological debate is an independent organizational mechanism even when accounting for the effects from common interests, opinion congruity, purely structural effects, and conversational norms. These findings differ from the selective exposure thesis and previous incidental claims of political disagreement. In addition, the findings indicate that endogenous mechanisms and opinion congruity could influence the tendency of cross-ideological debate to varying degrees.

Previous studies on political discussion emphasize its effects on deliberativeness and show less interest in how people organize their discussions (Eveland, Morey, & Hutchens, 2011; Ryfe, 2005). Most researchers assume that deliberation ensues when certain structural conditions (e.g., equality and autonomy) hold. By organizing interactions along these lines, they feel free to assume that deliberation takes place, thus allowing them to focus on measuring its effects. In the process, the structure of deliberation itself remains essentially unexamined.

The structure here refers to the autonomous rules of organizing political discussions—who replies to whom. The structure of a discussion network that consists of a set of nodes (discussants) and the collection of edges between them (communication relations) has proven important in and of itself (Gonzalez-Bailon, Kaltenbrunner, & Banchs, 2010), in terms of shaping public opinion (Price, Nir, & Cappella, 2006), and in mobilizing political participation (Leighley, 1990). Still, few studies have sought to answer the question: Why, in a situation of free choice in online settings, does a person choose to talk to one person rather than another (Sykes, 1983)?

Corresponding author: Hai Liang; e-mail: rainfireliang@gmail.com

Ideally, cross-ideological discussions could benefit deliberative democracy (e.g., Mutz, 2006). However, scholars have frequently found that online discussions are organized according to a homophily mechanism, which means that participants are more inclined to interact with or be exposed to politically similar individuals (e.g., Adamic & Glance, 2005; Himelboim, McCreery, & Smith, 2013). These previous studies, however, did not seriously consider other competing mechanisms of organizing human communication and thus might generate biased results. As suggested by Kossinets and Watts (2009), homophily could also be induced by other structural mechanisms. To fill this gap, this study employs a relational event framework (Butts, 2008) to differentiate homophily from other mechanisms and test its relative effect size in organizing online political discussions.

Organizational principles of political discussion

Homophily and cross-ideological debate

One of the most fundamental principles of human communication is that the exchange of messages most frequently occurs between those who are similar (Rogers & Bhowmik, 1970). This phenomenon is always conceptualized as *homophily*, which means that pairs of individuals who interact are similar with respect to certain attributes. Rogers and Bhowmik argued that when senders and receivers share common meanings, attitudes, and beliefs, communication between them is more effective. Among the various attributes, homophily with respect to political ideology receives special attention in research on political deliberation, because deliberative conversation is usually characterized by cross-ideological discussions (e.g., Calhoun, 1988; Delli Carpini, Cook, & Jacobs, 2004; MacKuen, 1990; Moy & Gastil, 2006).

The Internet has been considered an amplifier of the homophily principle in computer-mediated communication (CMC). First, people are freer to select discussion partners with like-minded political views (Sunstein, 2001). Online chat rooms or message boards may actually allow people to reinforce their predilections, because the same features that enable Internet users the potential to connect with dissimilar individuals facilitate their interactions with agreeable associates (Wojcieszak & Mutz, 2009).

Second, research on sociopsychological effects of CMC shows that CMC can lead to enhanced normative pressures under some conditions (Lee, 2007; Spears, Lea, & Lee, 1990). Lack of social cues in text-based communication makes the available cues in virtual spaces (e.g., group membership) situationally salient and results in strong influence of social norms on behavior. This will facilitate online users to converse with like-minded individuals.

Although the tendency toward homophily in social networks is well established, it is not necessarily operating in all online discussion environments. Some users seek reinforcement, but others go online to encounter different opinions. Individual motivations vary and, therefore, so do individual behaviors and ultimately the structures of the discussion networks that emerge from them (Kelly, Fisher, & Smith, 2005;

Stromer-Galley, 2003). The ability to discuss politics online without risking personal relationships may liberate many who would otherwise avoid political disagreement. The participation of even a small number of oppositional debaters in a discussion environment could inhibit political polarization. Politically heterogeneous discussion networks do exist on the Internet. For example, Kelly et al. (2005), in a study of eight politically oriented Usenet discussion newsgroups, showed that individuals often preferred discussing issues with users with whom they disagreed.

The most popular explanation for online cross-ideological debate is the incidental claim: People intend to communicate with similar users and are only incidentally exposed to opposing views (Lev-On & Manin, 2009; Wojcieszak & Mutz, 2009). Brundidge (2010) argued that inadvertency of media may facilitate exposure to diverse discussion networks through (a) less than perfect online selective exposure strategies, (b) nonavoidance of encounters with political difference, and (c) weakened social boundaries between far-flung geographic locations. This incidental claim indicates that, even with less purposive seeking out of political difference, people are still likely to be exposed to it.

This unintentional consequence could occur depending upon platform characteristics and dysfunctional sociopsychological influences. In the offline situation, mass media (Mutz & Martin, 2001) and workplaces (Mutz & Mondak, 2006) are the places for exchanging opposing political views. Similarly, online news websites and nonpolitical discussion spaces (Wojcieszak & Mutz, 2009) are expected to gather people with different opinions and provide chances for cross-ideological debate. Moreover, Lev-On and Manin (2009) argued that although the lack of social cues leads to stronger social pressures, it might facilitate cross-ideological debate when no cues are available on the platform.

Endogenous principles of organization

In addition to homophily, a discussion network could evolve endogenously without knowing any attributes of the discussants. That means individuals could select their discussion partners simply based on the existing communication relations. The first set of these endogenous principles are *structural mechanisms*. Sometimes, they are also referred to as purely structural effects (Ackland, 2013). Reciprocity, transitivity, and preferential attachment (popularity effect) are the three most important examples of these effects. Although some of these are the prerequisite conditions for political deliberation (e.g., reciprocity), they have rarely been examined in the previous studies (see Schneider, 1996).

Reciprocity refers to a tendency that if *A* replies to *B*, there is a good chance that *B* will reply to *A* in later conversations (i.e., $AB \rightarrow BA$). The tendency reflects the degree of true conversation (vs. monologue) among the participants. Himelboim's (2011) study found that the reciprocity ratio is 0.44 on average for political newsgroups and 0.53 for philosophy-related groups.

Transitivity refers to the tendency that if *A* talks with *B*, and *B* talks with *C*, there is a good chance that *A* will talk with *C* in later conversations (i.e., $\{AB, BC\} \rightarrow AC$).

A strong tendency of transitivity indicates group cohesiveness in social network analysis. Although it has long been demonstrated as a powerful predictor of friendship relationships (e.g., see Davis, 1967, 1970; Holland & Leinhardt, 1971, 1981), few studies related it to the selection of discussion partners. Carpenter, Esterling, and Lazer (2004) found that contacts with third-party actors are a crucial determinant of presence or absence of communication between two actors in the context of exchanging political information.

The popularity effect refers to the tendency that people are more inclined to talk to the users who have received many replies. Barabasi and Albert (1999) found that the older nodes in a network are more likely to have collected links than the ones that have just been created. In this way, network growth favors nodes that have been around a long time. Because of network growth and preferential attachment, the "rich get richer."

Purely structural mechanisms can also affect the measure of homophily. The problem for researchers studying homophily is that both opportunity structures and network effects other than homophily can "mask" the true level of homophily in a social network (see Ackland, 2013). The observed homophily could be induced by transitivity and other structural mechanisms (Kossinets & Watts, 2009). Empirical estimation of homophily could be very biased without consideration of these structural mechanisms (e.g., Wimmer & Lewis, 2010). Therefore, these mechanisms are usually treated as essential control variables in social network analysis, even when they are not the focus of the study in question.

Conversational norms governing communicative acts have substantial influences on the structure of communication (Goffman, 1967; Wilson, Wiemann, & Zimmerman, 1984). These norms include restrictions on the number of recognized speakers (Schegloff, 2000) and expectations of reciprocity in turn-taking (Wilson et al., 1984), as well as other constraints on attention and involvement that allow interaction to be maintained over time (Goffman, 1963). In order to quantify the effect of conversational norms on communicative action, Gibson (2003, 2005) proposed quantifying the local temporal dynamics of conversation by counting events known as participation shifts (or "P-shifts").

Gibson (2003) partitioned the participants in a conversation into the roles of speaker (sender A), target (receiver B), and unaddressed recipient (third party X). As the conversation unfolds, occupancy of these roles shifts; such shifts were governed by the norms of conversational interaction, and they are the basis for Gibson's analysis.

Based on Gibson, Butts (2008) integrated P-shifts into a relational event model and tested the hypotheses using real communication data. P-shift effects are formally defined as: turn receiving ($AB \rightarrow BA$, $AB \rightarrow B0$, $AB \rightarrow BY$), turn claiming ($A0 \rightarrow X0$, $A0 \rightarrow XA$, $A0 \rightarrow XY$), turn usurping ($AB \rightarrow XA$, $AB \rightarrow XB$, $AB \rightarrow XY$), and turn continuing ($A0 \rightarrow AY$, $AB \rightarrow A0$, $AB \rightarrow AY$). Turn receiving states that the receiver of the initial event is a potential sender, and the receiver of the initial event is also the sender of the subsequent event. For instance, $AB \rightarrow BA$ reflects the reciprocal relationship, and $AB \rightarrow B0$ indicates a shift in which a sender A directs an event toward another

receiver, *B*, who in turn directs the next action toward a nonsending target. Turn claiming P-shifts involve scenarios in which the recipient of the first action is not a potential sender. Turn usurping P-shifts involves the interruption of a conversation by a new speaker. Finally, turn continuing shifts involve scenarios in which the sender is preserved in each event.

Social selection of discussion partners

In order to investigate the organizational rules of online political discussions, this study situated the hypotheses in web-based discussion forums. There are a number of shared characteristics for web forum discussions: They are public, moderated, and organized by themes and topics; participants can remain anonymous and do not have to be online at the same time (Witschge, 2008). The web discussion forum has long been celebrated for its deliberative potentials (e.g., Himelboim, 2011; Rheingold, 1993), and forum discussions are less constrained by personal relationships than other social media platforms. Discussions in web forums are organized in the form of conversation threads. A thread is a collection of a seed post and its subsequent replies, usually displayed in chronological order.

The major difference between a web forum and a Usenet newsgroup is that the newsgroup automatically delivers every new message to its subscribers. A forum on the other hand, requires users to visit the website and check for updates. In this sense, users of web forums are more selective than newsgroup users. Forums also differ from chat rooms and instant messaging services in their asynchronous communication. Thus, forum users have more time to respond to others deliberatively.

The major purpose of this study is to examine the social mechanisms that drive individuals to select discussion partners and eventually form a discussion network. A major controversy in previous studies is whether online political discussions are based on the principle of homophily. Although cross-ideological debate is presumed an essential requirement for political deliberation, there is much empirical evidence against it in online settings (e.g., Adamic & Glance, 2005; Himelboim et al., 2013). Nevertheless, studies conducted in online forums consistently support that forum users are more inclined to seek politically dissimilar conversational partners (e.g., Kelly et al., 2005; McGeough, 2010). Therefore,

> RQ1: To what extent do users discuss politics with those who are politically dissimilar in web discussion forum?

In addition, previous studies in political communication generally measured cross-ideological debate by calculating the frequency of participating in cross-ideological discussions in ego networks (e.g., Wojcieszak & Mutz, 2009). This actually measures the extent to which individuals are exposed to politically dissimilar perspectives. On the other hand, cross-ideological debate as an organizational principle, which is equivalent to heterophily, is a tendency of selecting politically dissimilar discussion partners.

Simply calculating the proportion of cross-ideological discussions at the aggregate level is both theoretically and methodologically insufficient. First, as found in

Conover et al. (2011), a retweet network (exposure) is highly polarized, whereas the user-to-user mention network (debate) is politically heterogeneous on Twitter. The result implies that users could be exposed to homogenous content but converse with politically dissimilar users.

Second, even if homogenous discussions are more frequent than cross-ideological debates, this finding does not necessarily mean that individuals prefer to converse with those who are similar. Endogenous factors, such as structural mechanisms and conversational norms, are the major competing mechanisms of the homophily principle in generating discussion networks. These mechanisms should be controlled carefully in order to obtain a more reliable estimate of the tendency to select discussion partners in forum discussions. Following the tradition, purely structural mechanisms include popularity, reciprocity, and transitivity. Therefore,

> H1a: Users are more likely to discuss with the individuals who have already received many replies (popularity effect), have posted replies to (reciprocity), or share many discussion partners with (transitivity) the potential senders.

Empirical studies have usually selected parts of the conversational norms as predictors according to the context of communication. For instance, Butts (2008) included $AB \rightarrow BA$, $AB \rightarrow AY$ and $AB \rightarrow XB$, $AB \rightarrow BY$, and $AB \rightarrow XA$ to examine the effects on radio conversations. Reciprocity as a conversational norm is measured in the same fashion as how structural mechanism is measured; the only difference is that they are based on different theoretical perspectives. Therefore, besides reciprocity ($AB \rightarrow BA$), this study includes two additional conversational norms: $AB \rightarrow AY$ and $AB \rightarrow BY$. Because forum users are unlikely to check for updates every minute, they might post replies continuously once they logged in. This phenomenon is characterized by $AB \rightarrow AY$. And $AB \rightarrow BY$ indicates a diffusion-like process in forum discussions: Users who received replies are more likely to post a reply to another user.

> H1b: Users are more likely to post a reply when they have just posted a reply to ($AB \rightarrow AY$) or have just received a reply from ($AB \rightarrow BY$) another user.

> RQ2: Is there any difference in terms of the likelihood of cross-ideological debate when structural factors and conversational norms are taken into account? In other words, do endogenous factors mediate the cross-ideological effects on selection of discussion partners?

Another, more fundamental, controversy is whether and to what extent online cross-ideological debate is an incidental byproduct (Lev-On & Manin, 2009; Wojcieszak & Mutz, 2009) or an independent rule of organizing online political discussions. The incidental argument suggests several factors that increase the chance of cross-ideological debate on the Internet. One of the most relevant is the social context for cross-ideological debate (Mutz & Mondak, 2006). In web forums where discussions are organized according to topics, users sharing interests in certain topics have more opportunities to meet up with each other, and thereby they are more likely to interact with each other. Therefore,

> H2: Users are more likely to discuss with the individuals who share common interests both in political and nonpolitical issues.

Furthermore, because those who share common interests in political and non-political issues might come from different political backgrounds, according to the incidental argument,

H3: The impact of cross-ideological principle on selecting discussion partners is mediated by common interests.

Beyond the dyadic-level organization principles, opinion climate could also influence individual actions. MacKuen (1990) advanced a game theory about the conversational conditions that encourage people to become involved in political talk. Basically, the theory proposes that the likelihood of getting involved in a political discussion depends on the degrees of perceived friendliness of one's conversational environment and, thus, one's expected "pleasure" from the conversation. From a national representative survey, Kim, Wyatt, and Katz (1999) actually found that perceived friendliness of the opinion climate bolsters an individual's willingness to argue with people with different opinions. In web discussion forums, when users find there are many replies for their positions, they are more confident to argue and are willing to convince non-like-minded other users. Therefore,

H4: Users are more likely to discuss with politically dissimilar other users when the opinion climate is congruent with their own.

Method

Data collection

Concerning inclusiveness and representativeness of online deliberation, forum selection was based on popularity of the website (i.e., user traffic, number of participants, and number of posts). *Alexa.com* provides a list of top political discussion forums under the category "All Categories > Society > Politics > Chats and Forums."[1] The top three websites were examined for possible selection. *Politicalforum.com* does not provide any information about users' political orientation, which is the core variable in this study. *Revolutionaryleft.com* is a web forum intentionally designed for radical leftists and, thus, is not a platform for cross-ideological discussions.

Accordingly, this study selected *debatepolitics.com* as the research arena. It is a very popular, active political forum in the United States with 20,053 members who have created nearly 6 million posts. According to *Alexa.com*,[2] more of *debatepolitics.com*'s users are male than female, and users are better educated than Internet users in general. Although the platform guarantees anonymity, it forces members to report their political leanings. The platform claims to have a nonbiased, nonpartisan, and equal discussion environment. It encourages cross-cutting discussions and has strict rules to ensure civil discussions.

As on most web-based political discussion forums, discussion moderators scan the messages, delete inappropriate replies (e.g., flaming, harassment, spamming), and suspend posting privileges of users who break the rules (see Witschge, 2008). Unlike some government-run forums (see Wright, 2006), moderators are not expected to stimulate conversations or provide background information or other services.

Although this forum does not represent all kinds of discussion platforms, it shares the general characteristics of web-based political forums created for nonpartisan and citizen-driven discussions (see Witschge, 2008).

Owing to computational complexity and sampling difficulty in social network analysis, researchers usually focus on a moderate-sized network (Wasserman & Faust, 1994).[3] In doing so, this study collected all threads and messages posted in the section "2012 US Presidential Election" on *debatepolitics.com*.[4] A presidential election is arguably one of the most important political events in the United States, and discussions on it usually concern a wide range of issues (Zhang, Cao, & Tran, 2013). There are 175,960 messages in 2,372 discussion threads in the final record. The total number of authors involved in the discussions is 1,178.

The following information for each post or reply was extracted automatically by a computer program: text of the message, time of the message posted, author name of the message, author name being mentioned in the message, and political leaning of the author. In order to compare political discussions with nonpolitical discussions, the study also collected all discussion threads and messages posted in the nonpolitical subforums.[5] This nonpolitical dataset contains 1,402 participants and 150,949 messages without self-replies. The topics in this subforum include academia, conspiracy theories, leisure activities, philosophy, and religious discussions.

Measures
Political ideology
Political ideology is explicitly provided by the *debatepolitics.com*. The platform forces users to select their political leanings from a list of 18 items. The items are recoded into three categories for formal analysis: Left (recoded from liberal, libertarian-left, progressive, slightly liberal, socialist, very liberal, communist), right (conservative, slightly conservative, very conservative, libertarian, libertarian-right), and centrist (centrist, moderate, independent, other). For those users who selected "private" and "undisclosed" (83/1178), political ideologies were manually coded according to the content of the replies. Of 1,178 participants in the election sample, 22% (261) are left-wing, 30% (358) are right-wing, and 47% (559) are centrist.

Relational event stream
Reply-to relationship is explicitly displayed in the texts. A user can reply to a previous message specifically and the platform will automatically quote this message in the reply. Replies without mentions of other users (17.3% of the population) are assumed in reference to the author of the seed posts (i.e., the first post of a thread).[6] The reply-to relationship could be considered as the edge in a discussion network where the nodes are discussants. Whether to reply to a specific person is determined by the sender rather than the receiver, so edges in a discussion network are directed. There are 171,338 edges excluding self-replies.

Relational event stream is measured according to the reply-to relationship in forum discussions. A relational event is operationalized as an author *A* replies to

author B at time t, which is represented by $a = (A, B, t)$, where $A \in S$ represents the sender of the action ($s(a) = A$), $B \in R$ represents the receiver ($r(a) = B$), and $t \in R$ represents the time at which the action is taken ($\tau(a) = t$). Given a time-ordered set of actions $a_1, a_2, \ldots$, let the set $A_t = \{a_i : \tau(a_i) \leq t\}$ consists of all actions taken on or before time t. A_t is an event stream. Methodologically, the purpose of the study is to predict A_t by the following variables.

The homophily mechanism hypothesizes that a relational event is more likely to occur when the sender and receiver are from the same ideological group. In order to test this tendency, the study constructs two $1,178 \times 1,178$ ideology matrixes to indicate whether any two participants are from different ideological groups (coded as 1) or not (0) and whether one of the two participants is the centrist (1) or not (0). These two matrixes (i.e., cross vs. within, centrist vs. within in Table 2) are included in formal models simultaneously as covariates. This operationalization is analogous to regression models with categorical predictors: The omitted within-ideological matrix is the reference group. The parameter of the cross-ideological matrix represents the likelihood difference between left↔right and (right↔right + left↔left) discussions.

Common interests

The common interest in politics between discussants is measured by the number of threads they had both attended in the past political discussions, whereas the common interest in nonpolitical issues is measured by the number of threads they had attended in nonpolitical discussions. For two participants, if there are 200 political threads in the forum that both of them visited before time t, the common interest in politics is 200 at time t. The final data are contained in a dynamic and symmetric $1,178 \times 1,178 \times t$ matrix. Common interest in nonpolitical issues is measured similarly.

Opinion congruity

Opinion congruity is measured by the consistency between personal opinion and the opinion climate. In previous studies, opinion climate could be measured at different levels. However, Yun and Park (2011) found that only the immediate opinion climate (in a discussion thread) shows a significant influence on public expression in online forums. Therefore, opinion climate is measured by the average opinion of existing replies in the same discussion thread. Because the data do not contain any issue stances directly, political ideologies were recoded as −1 (left), 0 (centrist), 1 (right) to represent users' opinions. The opinion congruity index is constructed by multiplying the own opinion with the opinion climate, yielding a continuous score in which −1 indicates low congruity and +1 indicates high congruity (Ho & McLeod, 2008). In a relational event stream, the opinion congruity for each user is time-varying given the number of replies is growing in discussion threads. To differentiate the impacts of opinion congruity on sending and receiving messages, two variables were introduced in the formal analysis: "CongruentSnd" indicates the sending effect, whereas "CongruentRec" indicates the receiving effect

in Table 2. Accordingly, "CongruentSnd × Cross" indicates the interaction effect between congruity sending effect and cross-ideological tendency.

Endogenous factors

The calculational details of all endogenous factors could be found in Butts (2008). Popularity is measured by the effect of the number of messages that *A* received from other users on *A*'s future receiving rate. Reciprocity effect is measured by the effect of the number of messages that *B* replied to *A* on the probability that *A* replies to *B*.

Transitivity is measured by the effect of the number of shared discussion partners between *A* and *B* on the probability that *A* will talk to *B*. Given the reply-to relationship is directional, transitivity has several variations. "OTPSnd" measures the probability of $AXB \rightarrow AB$, which states that if *A* and *B* share many *X*s (and *A* replies to *X*, *X* replies to *B*), *A* will reply to *B*. "ITPSnd" measures the probability of $BXA \rightarrow AB$, "OSPSnd" measures the rate of $\{AX, BX\} \rightarrow AB$, and "ISPSnd" measures the rate of $\{XA, XB\} \rightarrow AB$.

In addition, this study includes two recency effects to control the autocorrelation effect — the impact of recency of previous actions on future actions. "RRecSnd" measures the recency of how receipts of replies from *B* affects *A*'s future rate of sending to *B*. "RSndSnd" measures the recency of how sending to *B* affects *A*'s future rate of sending to *B*.

Random-effects relational event stream model

In order to differentiate and measure the different mechanisms of organizing forum discussions, this study employed a random-effects relational event stream model (RERESM), which is composed of a relational event stream model and a random-effects model. The event stream model is a sequence of survival models (Butts, 2008) in predicting an event sequence A_t. To understand the dynamics within discussion threads (e.g., the impact of opinion congruity) and reduce computational complexity, the relational event model is repeatedly conducted on each of the 1,833 threads (with more than five replies) instead of estimating them as a whole network. That means the relational event stream is constructed at the thread level rather than at the forum level. Second, the random-effects model is used to estimate the average effect size, statistical significance, and heterogeneities across the discussion threads. In doing so, Hedges and Vevea's (1998) random-effects meta-analysis is employed.

Results

Cross-ideological debate

To answer RQ1, Table 1A presents the number of replies between and within different political ideologies. Generally speaking, users were more inclined to converse with politically dissimilar other users. There were many of messages that were directed to the opposite positions. Only 20.5% of messages were within-ideological replies. The left users were more inclined to reply to the right, and the right users

Table 1 Frequency and Expected Frequency of Cross-Ideological Debate in Political (A) and Nonpolitical (B) Discussions

A

From-to-	Centrist	Left	Right	Row Total
Centrist	**9,915**	12,951	17,906	40,772
	(9,347)	(14,022)	(17,403)	
Left	12,077	**8,103**	38,116	58,296
	(13,365)	**(20,049)**	(24,882)	
Right	17,288	37,872	**17,110**	72,270
	(16,568)	(24,855)	**(30,847)**	
Column total	39,280	58,926	73,132	171,338

B

From-to-	Centrist	Left	Right	Row Total
Centrist	**19,883**	8,565	29,075	57,523
	(19,044)	(9,139)	(29,341)	
Left	12,284	**5,785**	10,075	28,144
	(9,318)	**(4,471)**	(14,355)	
Right	17,807	9,631	**37,884**	65,282
	(21,613)	(10,371)	**(33,298)**	
Column total	49,974	23,981	76,994	150,949

Note: Cell entries are the observed frequencies with expected values in parentheses. Bold values indicate within-ideological discussions. Self-replies are excluded.

were more inclined to reply to their left counterparts. The observed values within left or right were significantly smaller than the expected values (8,103 vs. 20,049, 17,110 vs. 30,847). However, the observed value within centrists was larger than the expected value (9,915 vs. 9,347). Cross-ideological debates always occurred between the left and the right, whereas centrists were more likely to converse with centrists.

Table 1B shows that the observed values within ideologies were larger than the expected (5,785 vs. 4,471, 37,884 vs. 33,298) in nonpolitical discussions. Unlike the incidental argument, users in this forum were more inclined to communicate with the users who share the same ideology about nonpolitical issues. The table shows little evidence that nonpolitical discussions could increase the likelihood of cross-ideological debate.

Organizational rules

Table 2 presents the summarized coefficients of RERESM. In Model 1, the positive coefficient of cross versus within-ideological discussion suggests that users were more inclined to select politically dissimilar users as discussion partners. *B* is 1.05 (*SE* = .03) which could be interpreted as the odds ratio for selecting a dissimilar partner being

Table 2 Random-Effects Relational Event Model for Discussion Organization

	Model 1	Model 2	Model 3	Model 4
Ideology				
Cross versus within	1.05**	0.76**	0.75**	0.80**
	(.03)	(.02)	(.02)	(.02)
Centrist versus within	0.73**	0.55**	0.53**	0.54**
	(.03)	(.03)	(.03)	(.06)
Conversational norms				
$AB \rightarrow BA$		1.93**	1.89**	1.88**
		(.02)	(.02)	(.02)
$AB \rightarrow BY$		1.13**	1.08**	1.08**
		(.02)	(.02)	(.02)
$AB \rightarrow AY$		1.39**	1.34**	1.34**
		(.02)	(.02)	(.02)
Structure				
Popularity		6.86**	6.61**	0.83**
		(.08)	(.08)	(.08)
RRecSnd		2.89**	2.84**	2.86**
		(.02)	(.02)	(.02)
RSndSnd		−0.77**	−0.83**	−0.84**
		(.03)	(.03)	(.03)
OTPSnd		−0.14**	−0.16**	−0.16**
		(.01)	(.01)	(.01)
ITPsnd		0.17**	0.15**	0.15**
		(.01)	(.01)	(.01)
OSPSnd		0.12**	0.12**	0.12**
		(.01)	(.00)	(.00)
ISPSnd		−0.08**	−0.08**	−0.08**
		(.01)	(.01)	(.01)
Common interests				
Politics			0.11**	0.11**
			(.01)	(.01)
Nonpolitics			−0.05**	−0.05**
			(.01)	(.01)
Opinion congruity				
CongruentSnd				−0.14*
				(.07)
CongruentRec				0.45**
				(.06)
CongruentSnd × Cross				0.20*
				(.10)
CongruentSnd × Centrist				0.09
				(.09)

Table 2 *Continued*

	Model 1	Model 2	Model 3	Model 4
Control variables — sending effects				
Author or not	1.09**	0.05	0.12**	0.12**
	(.03)	(.03)	(.03)	(.04)
Left versus centrist	0.07**	0.03	−0.03	−0.08**
	(.02)	(.02)	(.02)	(.02)
Right versus centrist	0.18**	0.12**	0.08**	0.10**
	(.03)	(.02)	(.02)	(.02)
Control variables — receiving effects				
Author or not	2.48**	1.30**	1.41**	1.39**
	(.02)	(.03)	(.03)	(.03)
Left versus centrist	0.16**	0.09**	0.02	0.00
	(.03)	(.02)	(.02)	(.03)
Right versus centrist	0.31**	0.16**	0.11**	0.07**
	(.03)	(.02)	(.02)	(.02)
Number of threads (messages)	1,833	1,833	1,833	1,833
	(170,634)	(170,634)	(170,634)	(170,634)

Note: Cell entries are unstandardized coefficients with standard errors in parentheses. The model fit is calculated by the accuracy of prediction. Model 1 could predict 14.8% of the discussion pairs including both senders and receivers at a specific time. Model 2 is 27.4%. Model 3 is 27.7%. Model 4 is 28.0%.
*$p < .05$. **$p < .01$.

185.8% (i.e., $\exp(1.05) - 1$) higher than selecting a similar one when holding all control variables at a fixed value. The effect size was $Z = 41.67$.

In terms of control variables, authors of threads were more active (sending effect) and attractive (receiving effect) than other contributors (author or not). The odds ratio for authors receiving a reply was 1094.1% higher than the ordinary users ($B = 2.48$, $SE = .02$), whereas the odds ratio for sending a reply was 197.4% higher ($B = 1.09$, $SE = .03$). Both the left and right were more active and attractive than centrists.

Concerning H1 and H2, there were two major competing mechanisms in organizing online political discussions: the endogenous mechanisms and common interests. Model 2 suggests that selecting discussion partners was subjected to purely structural mechanisms as well as conversational norms. Specifically, the significant popularity parameter indicated that users were more inclined to reply to the popular posters. The large effect size ($B = 6.86$, $SE = .08$) suggests that there was a large difference with one more reply in attracting others' replies. This suggests a strongly hub-dominated network, with a few actors receiving most of the replies. Model 1 suggests that the dominant users were the authors of discussion threads. However, when controlling the popularity effect, the effect size of the author decreased from 107.70 to 43.22

in Model 2. This indicates that the domination of the author could be explained in part via a preferential attachment mechanism.

The effect sizes of RRecSnd and RSndSnd were 134.58 and −25.66 respectively, which suggests that users were more likely to discuss with the person who has sent a reply to them recently but less likely to discuss with the one who has received their replies earlier. The effects of OTPSnd, ITPSnd, OSPSnd, and ISPSnd represented the triadic effects in conversations. All effect sizes were relatively small compared to other purely structural effects (−13.47, 18.19, 24.50, and −7.13 respectively). The negative coefficient of OTPSnd indicated that if A replies to X, and X replies to B, then A will be less likely to reply to B. The positive coefficient of ITPSnd indicated that if B replies to X, and X replies to A, A will be more likely to reply to B. The positive coefficient of OSPSnd means that if A replies to X, and B replies to X, A will be less likely to reply to B. Finally, the negative coefficient of ISPSnd indicated that if X replies to A, and X replies to B, A will be less likely to reply to B.

Conversational norms, as expected, showed a strong impact on the organization of political discussion. $AB \rightarrow BA$ models the reciprocal effect of conversation ($B = 1.93$, $SE = .02$, $Z = 82.39$). The odds ratio for B replying to A when B has received B's reply was five times higher than the odds ratio when A did not send a message to B previously. Users were more likely to post replies consecutively, given the effect size of $AB \rightarrow AY$ is 70.06 ($B = 1.39$, $SE = .02$). Users who just received a reply were also more likely to post a reply to another user ($B = 1.13$, $SE = .02$, $Z = 50.29$).

Overall, both purely structural mechanisms and conversational norms showed great impacts on selecting discussion partners. Therefore, both H1a and H1b are supported. Regarding RQ2, the impact of cross-ideological debate did decrease when endogenous factors are included in Model 2. The coefficient of cross-ideological discussion declined from 1.05 in Model 1 to 0.76, which means that the estimated odds ratio—cross/within—declined by 0.72 when controlling endogenous factors. The effect size declined slightly from 41.67 to 40.03, suggesting that there was a mediation effect of endogenous factors on cross-ideological organization of political discussion. Yet, the principle of selecting politically dissimilar partners appeared consistently strong even when accounting for the impact of endogenous mechanisms.

Model 3 includes common interests as edge covariates. As predicted in H2, users were more likely to have a discussion when they shared common interests in political issues ($B = 0.11$, $SE = .01$, $Z = 20.37$). Those users who frequently joined in the same political discussion threads were more likely to select each other as discussion partners. However, common interests in nonpolitical issues showed a negative effect on selecting discussion partners ($B = −0.05$, $SE = .01$, $Z = −7.71$). Therefore, H2 is not supported.

Compared to Model 2, the coefficient and effect size of cross-ideological debate decreased slightly (from 0.76 to 0.75, 40.03 to 39.27 respectively). The mediation effect of common interests proposed in H3 was quite small. To further illustrate the mediation effects of common interests on cross-ideological debate, the correlations between the variables were examined. The results showed that the correlation

between the common interest in politics was positively related to the frequency of cross-ideological discussions ($r = .54$, $p < .01$). The coefficient was lower between the common interest in politics and the frequency of within-ideological discussions ($r = .44$, $p < .01$). The results indicated that common interest in politics increased cross- much more than within-ideological discussions. However, the correlation between the common interest in nonpolitics and the frequency of cross-ideological discussions was not significant ($r = -.01$, $p > .05$). Therefore, H3 is not supported.

According to H4, opinion congruity could influence the formation of a discussion network. Model 4 suggests that users were less likely to express congruent opinions (CongruentSnd, $B = -0.14$, $SE = .07$). However, congruent opinions were more likely to be mentioned by others (CongruentRec, $B = 0.45$, $SE = .06$). The last, a positive interaction effect between congruent opinion climate and cross-ideological debate ($B = 0.20$, $SE = .10$), suggests that users were more likely to select politically dissimilar partners when opinion climate is congruent with their own. Therefore, H4 is supported. In addition, the estimated odds of discussing with politically dissimilar individuals increased by 0.11 from Model 3 to Model 4, implying that some within-ideological discussions were induced by opinion congruity. Further analysis suggests that 61% of within-ideological discussions and 53% of cross-ideological discussions occurred when the opinion climate was congruent.

Discussion

This study employed a relational event model to examine various competing organizational principles in online political discussions. The most striking finding is that cross-ideological debate was an independent principle in organizing political forum discussions. Users were more likely to talk with politically dissimilar users, even when accounting for all other competing mechanisms. Among these, common interest has been shown to be a major competing mechanism by previous studies considering cross-ideological debate as an incidental consequence of meeting people with different ideologies who share an interest in political or nonpolitical discussion (e.g., Lev-On & Manin, 2009; Wojcieszak & Mutz, 2009). Findings of this study did not support this argument. They showed that common interest in politics facilitates cross-ideological discussions, whereas common interest in nonpolitical issues actually inhibited this tendency.

These findings differ from those of recent research on the selective exposure thesis. There could be several reasons for this inconsistency. First, previous studies did not consider the competing organizational principles in political discussions. The incidental argument assumes that selecting similar users is "normal," whereas selecting dissimilar users is considered an induced phenomenon. Actually, both homophily and heterophily could be induced by other competing mechanisms (see Kossinets & Watts, 2009). This study further confirmed this claim. Some competing principles could amplify the estimation of homophily in political discussions. In this study, opinion congruity was one of these factors. In contrast, endogenous factors might

be considered as amplifiers of cross-ideological debate, because the cross/within ratio declined when these factors were controlled. Estimations of homophily without consideration of structural effects could be very biased.

Second, the forms of cross-ideological interaction measured in the studies are different. Previous studies generally examined the extent to which individuals were exposed to homogenous arguments (i.e., cross-ideological exposure). However, this study focused on the extent to which individuals actively debate with politically dissimilar users (i.e., cross-ideological discussion). Being involved in political discussions is fundamentally different from passively viewing web pages. The effort involved in these situations varies and, therefore, so do individual choices.

When users are required to expend much effort to write replies in forum discussions (other than just clicking a rating button on the website), they were more inclined to disagree with existing opinions, because people tend to project a perception of being discriminate when expressions are costly (Wu & Huberman, 2010). In addition, in repeated discussions, composing messages requires much cognitive effort to logically organize personal thoughts, thus making messages senders more open-minded (Pingree, 2007).

This effort argument of cross-ideological interaction is consistent with empirical studies on online polarization. In Twitter studies, the odds ratio of "retweeting" cross-ideological messages (i.e., cross/within) was 0.02 (Conover et al., 2011), whereas the odds ratio of "replying-to" cross-ideological messages was 0.53 in Yardi and Boyd (2010) and 0.50 in Conover et al. (2011). Similarly, Hargittai, Gallo, and Kane (2008) found that the odds ratio of hyperlinking to cross-ideological bloggers was 0.10, whereas the odds ratio of mentioning cross-ideological bloggers in blogs was 0.16. These findings indicate that effortless expressions (e.g., retweeting, clicking a rating button) tend to be more homogenous than effortful political discussions (e.g., posting replies, writing blogs).

Third, the methods of data collection were different. Traditional studies on political disagreement generally used self-reported data. Wojcieszak and Price (2012) found that objective disagreement and perceived disagreement were weakly correlated in online forum discussions. Owing to people's looking-glass perception, individuals may report more agreement than the reality (e.g., Fields & Schuman, 1976). This could be a reason that survey-based studies usually support the homophily hypothesis.

Implications for online deliberation

Talking with politically dissimilar users satisfies the basic requirements of healthy political deliberation (Moy & Gastil, 2006; Mutz, 2006). However, the mere existence of diversity does not sufficiently indicate democratic discussions. Diversity could also imply aggressive debates without mutual understanding or a bipolarization of competing perspectives that lacks communication in discussion threads.

Regarding bipolarization, this study examined the effects of cross-ideological debate on organizing political discussions. Results suggested that there were sufficient (beyond expectation) amounts of discussions across political ideologies.

Even though cross-ideological debate decreased the degree of bipolarization with respect to political ideology, it might cause additional user fragmentation.

The proportion of cross-ideological discussions was negatively correlated with the density of the discussion network ($r = -.72$, $p < .01$). That means cross-ideological discussion networks were less dense (i.e., have fewer connections among participants) than within-discussion networks. Given that cross-ideological discussions were prevalent, in general, this suggests that forum users were more likely to argue with a small subset of opponents and to talk with a broader range of supporters.

Cross- and within-ideological discussions could possibly function differently in organizing a discussion network. Discussions between like-minded individuals could strengthen group identity (Yardi & Boya, 2010), and thus increase group cohesiveness. On the other hand, cross-ideological discussions are more instrumental to achieving consensus or persuading others. This instrumental motivation increases the possibility that discussions occur repeatedly between a pair of discussants. Normatively, cross-ideological discussion is not good for social solidarity and organizing collective actions (Freelon, 2010).

In addition, this explanation applies to why within-ideological discussions were more prevalent in nonpolitical discussions, as this study reported. Discussion contexts vary, and therefore so do individuals' motivations, and eventually interaction patterns. In nonpolitical discussions, such as of music and sports, people are more inclined to talk with politically similar people to reinforce group identity. From a communitarian perspective (Freelon, 2010), this kind of conversation is essential for community sustainability, further participation, and even civil deliberation.

Even if diverse opinions and a sufficient amount of cross-ideological debate exist in political discussions, people may discuss with each other in an insulting way. The well-known flaming phenomenon has long been considered as threatening to democracy (e.g., Papacharissi, 2004). Although this study did not directly test the degree of civility and politeness in cross-ideological debate, a post hoc analysis of angry expressions in cross-ideological replies was conducted following the classic sentiment analysis procedures in Computer Science (see Jurka, 2012). [7] The result showed that 9.2% of cross-ideological replies and 9. 4% of within-ideological replies contained angry expressions. Cross-ideological debates did not arouse significantly more angry expressions ($\chi^2 = 0.97$, $p > .05$), suggesting that the discussion environment was relatively polite and civil.

Several sociotechnical features might make web discussion forums more ideal for online political deliberation than other online discussion platforms. First, the mixed presentation of both proattitudinal and counterattitudinal messages in discussion threads might facilitate cross-cutting exposure and debate. If users are interested in a discussion thread and read it, they will inevitably encounter diverse political opinions. According to Garrett (2009), people have a psychological preference for proattitudinal information without a corresponding aversion to counterattitudinal information. Therefore, forum users are very likely to read and discuss across ideologies when they encounter counterattitudinal messages.

Besides, exposure to diverse opinions in a discussion thread implies that participants with any ideology might have a high level of opinion congruity on average and, thus, they are more inclined to argue with politically dissimilar users as this study found. In this sense, cross-ideological interactions should be fewer in partisan forums (that is designed for a specific party) and discussion platforms organized around individuals (e.g., Twitter, Facebook). Information presented on Twitter and Facebook is restricted by how individuals select their followees. However, the following networks on social media platforms are usually homogenous (Himelboim et al., 2013).

Second, most web discussion forums technically guarantee anonymity, often achieved by the use of pseudonyms in web forums. Even if posters use their actual name, there is no verification process and users are virtually impossible to identify. Although anonymity has the potential to increase uncivil discussions (e.g., flaming), attenuating sociopsychological barriers and increasing the likelihood of cross-ideological debate is important. Anonymity and lack of physical presence, which are common in online forums, may protect people from negative sanctions (see Yun & Park, 2011). However, anonymity is not so common on more recent social media platforms. Users of Facebook and Twitter are not anonymous and are well-connected with offline friends. Discussions on such platforms are more susceptible to social sanctions, and thus are more homogenous.

Third, the presence of moderation in web forums is expected to overcome the negative effects of anonymity. In the current forum, flaming, harassment, and other inappropriate messages in forum discussions are prohibited explicitly. This might explain the low number of angry expressions in cross-ideological posts; all angry expressions were deleted by forum moderators. In addition, moderation rules and guidelines were internalized, and participants self-moderated their content before posting. Dahlberg (2001) argued that high quality of deliberation is more likely to be associated with self-moderation rather than receipt of moderation. In this study, the first reason could be falsified immediately, because few messages were deleted.

Nevertheless, the web forum selected in this study is representative of a wide range of web-based political forums created for nonpartisan, citizen-driven discussions, according to Witschge's (2008) descriptions of web forums. Furthermore, this study provided an important empirical case in which political deliberation was more likely to happen due to the uniqueness of design, if any. It indicates that online political discussion has the potential to be deliberative. From a designer's perspective, future deliberative platforms should incorporate the features of web forums to ensure the high quality of deliberation on the Internet.

Limitations and future research

This study features several limitations that future research might address. To examine the endogenous mechanisms of network formation, this study focused on a single forum to obtain a whole discussion network. This method of data collection could limit generalizability of the current findings. First, users on *debatepolitics.com* are not representative of other populations. They are better educated than users of other

popular U.S. discussion forums according to *Alexa.com*. These users might be politi-
cally sophisticated and, thus, are more capable of arguing or debating with politically
dissimilar users without changing their minds. These arguments and debates should
be differentiated from deliberation, which emphasizes taking opposite ideas into con-
sideration when talking with others (Pingree, 2007).

Nevertheless, the effect of demographic characteristics on deliberative online
conversation is very limited. Representative survey findings indicate that neither
gender nor education have been significantly related to discussion network het-
erogeneity offline (Brundidge, 2010) or to exposure to political disagreements in
online discussion groups (Wojcieszak & Mutz, 2009). This implies that results of
this study might be replicable in other web forums with different demographic
characteristics.

Participating in web forums, however, is a self-selection process per se. Diversity-
seeking users may be more likely to participate in forums encouraging cross-
ideological debate. On the other hand, people who are less open-minded, less
informed, and less interested in politics may select other discussion platforms for
opinion reinforcement. In this sense, online deliberation only exists among a niche
of individuals who are intended to be deliberative. Future research should pay
more attention to the sociotechnical conditions and individual characteristics for
cross-ideological debate on the Internet.

Second, as mentioned above, different sociotechnical features of platforms could
significantly influence the likelihood of deliberation on them. In particular, modera-
tion on *debatepolitics.com* is relatively occasional and light—limited to discouraging
flaming, spamming, profanity, and so on. The forum rules are not intended to rigidly
determine how participants should behave. They are focused on what is not allowed
and how participants should not behave. To maintain civility, certain words are cen-
sored by software in advance. Harassment and hate messages toward a protected group
(e.g., race, gender) are judged by the moderator team. Users who break the rules will
lose their posting privileges.

However, moderation could be very pervasive and active on some discussion
forums. For example, Coleman and Gotze (2001) emphasized the role of moderation
in online deliberation, such as ensuring fair exchanges among the parties, offering
a balanced summary of arguments, and giving feedback to participants. On the
contrary, there are many discussion forums do not have regulation rules, such as the
Usenet political forums in Kelly et al. (2005).

Although forum rules are necessary in shaping the quality of online debate,
how the variation of rules can influence political deliberation remains unclear.
Jensen (2003) argued that a high level of active moderation can lead to an improved
quality of argumentation (e.g., civility, rationality). However, political discussions
without moderation could increase the likelihood of cross-ideological debate due
to the lack of social constrains (Kelly et al., 2005). Rhee and Kim (2009) found that
active moderation (e.g., providing background information) actually decreased the
number of message postings in a field experiment. In this case, whether the rules on

debatepolitics.com actually facilitate or inhibit the tendency to political disagreement is difficult to know.

Furthermore, an aim of *debatepolitics.com* is to facilitate information exchange and debate across political parties, as the forum claimed on the front page. Even if there are no concrete rules to ensure that discussions will be organized in this way, it could be perceived as the norm of interaction on this platform. This norm might increase the possibility of cross-ideological debate. Future studies should address how and to what extent explicit forum rules and implicit norms could alter the likelihood of cross-ideological discussion on the Internet.

Finally, this study was more focused on the structural aspects of forum deliberation. Whether the discussions are civil or rational leads toward mutual understanding is still unclear. The content of political discussion is equally important for evaluating online deliberation. Although the post hoc analysis of angry expressions in this study suggested an optimistic interpretation, some other damaging expressions to effective deliberation (e.g., dismissive and ridiculing expressions) have not been examined (e.g., Papacharissi, 2004). Whether cross-ideological discussions are civil and rational is far from clear. Future research might extend this study by systematically investigating the dynamics between the structure and content of online political discussion.

Regardless of these limitations, this study contributes to the literature about online political discussion in several respects. First, it demonstrated that forum users were more likely to talk politics with politically dissimilar users, which contradicts the selective exposure thesis and incidental arguments of political disagreement. It also sheds light on the design of future deliberative systems.

Second, this study employed a relational event framework to differentiate homophily from other competing mechanisms and suggested that both purely structural mechanisms and conversational norms could structure online political discussions. It demonstrated a new tool for analyzing continuously observed political discussions.

Third, these competing mechanisms could also mediate the observed homophily in real social networks. The study urges that future studies on selective exposure and political disagreement carefully control for these variables in quantitative analyses.

Acknowledgments

Notes

1 http://www.alexa.com/topsites/category/Top/Society/Politics/Chats_and_Forums

2 http://www.alexa.com/siteinfo/debatepolitics.com

3 Moderate-size here means that the number of nodes and the number of edges in a social network are not too large for computational purposes. It depends on the complexity of the algorithms employed in social network analysis. For the relational event model, the estimation process is very time-consuming. This study successfully analyzed a network with 1,178 nodes and 170,634 edges using a statistical machine in 1 week. However, it is nearly impossible to analyze all forum messages (over 6 million) using this model.

4 http://www.debatepolitics.com/us-elections/

5 http://www.debatepolitics.com/non-political-forums/

6 An alternative manipulation is to assign the receivers randomly from the population. Both methods generate similar findings. Tables presented in the results section are based on the first method.

7 First, for each reply, stop-words (e.g., a, an, the), punctuation marks, and white spaces were removed from the raw texts. Second, a naïve Bayes classifier trained on an emotion lexicon was employed to identify automatically whether a reply contained angry expressions or not.

References

Ackland, R. (2013). *Web social science: Concepts, data and tools for social scientists in the digital age*. London, England: Sage.

Adamic, L. A., & Glance, N. (2005). *The political blogosphere and the 2004 U.S. election: Divided they blog*. Paper presented at the Proceedings of the 3rd International Workshop on Link Discovery, Chicago, IL.

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, **286**(5439), 509–512.

Brundidge, J. (2010). Encountering "difference"; in the contemporary public sphere: The contribution of the Internet to the heterogeneity of political discussion networks. *Journal of Communication*, **60**(4), 680–700. doi:10.1111/j.1460-2466.2010.01509.x.

Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, **38**(1), 155–200. doi:10.1111/j.1467-9531.2008.00203.x.

Calhoun, C. (1988). Populist politics, communications media, and large scale social integration. *Sociological Theory*, **6**(2), 219–241. doi:10.2307/202117.

Carpenter, D. P., Esterling, K. M., & Lazer, D. M. J. (2004). Friends, brokers, and transitivity: Who informs whom in Washington politics? *Journal of Politics*, **66**(1), 224–246. doi:10.1046/j.1468-2508.2004.00149.x.

Coleman, S., & Gotze, J. (2001). *Bowling together: Online public engagement in policy deliberation*. London, England: Hansard Society.

Conover, M., Ratkiewicz, J., Francisco, M., Goncalves, B., Flammini, A., & Menczer, F. (2011, July). *Political polarization on Twitter*. Paper presented at the 5th International Conference on Weblogs and Social Media, Barcelona, Spain.

Dahlberg, L. (2001). The Internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. *Information, Communication & Society*, **4**(4), 615–633. doi:10.1080/13691180110097030.

Davis, J. A. (1967). Clustering and structural balance in graphs. *Human Relations*, **20**(2), 181–187. doi:10.1177/001872676702000206.

Davis, J. A. (1970). Clustering and hierarchy in interpersonal relations: Testing two graph theoretic models on 742 sociomatrices. *American Sociological Review*, **35**, 843–851. doi:10.2307/2093295.

Delli Carpini, M. X., Cook, F. L., & Jacobs, L. R. (2004). Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. *Annual Review of Political Science*, **7**, 315–344. doi:10.1146/annurev.polisci.7.121003.091630.

Eveland, W. P., Morey, A. C., & Hutchens, M. J. (2011). Beyond deliberation: New directions for the study of informal political conversation from a communication perspective. *Journal of Communication*, **61**(6), 1082–1106. doi:10.1111/j.1460-2466.2011.01598.x.

Fields, J. M., & Schuman, H. (1976). Public beliefs about the beliefs of the public. *Public Opinion Quarterly*, **40**(4), 427–448. doi:10.1086/268330.

Freelon, D. G. (2010). Analyzing online political discussion using three models of democratic communication. *New Media & Society*, **12**(7), 1172–1190. doi:10.1177/146144480 9357927.

Garrett, R. K. (2009). Politically motivated reinforcement seeking: Reframing the selective exposure debate. *Journal of Communication*, **59**(4), 676–699. doi:10.1111/j.1460-2466. 2009.01452.x.

Gibson, D. R. (2003). Participation shifts: Order and differentiation in group conversation. *Social Forces*, **81**(4), 1335–1380. doi:10.1353/sof.2003.0055.

Gibson, D. R. (2005). Taking turns and talking ties: Networks and conversational interaction. *American Journal of Sociology*, **110**(6), 1561–1597. doi:10.1086/428689.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. New York, NY: Simon and Schuster.

Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behavior*. New York, NY: Doubleday Anchor.

Gonzalez-Bailon, S., Kaltenbrunner, A., & Banchs, R. E. (2010). The structure of political discussion networks: A model for the analysis of online deliberation. *Journal of Information Technology*, **25**(2), 230–243. doi:10.1057/Jit.2010.2.

Hargittai, E., Gallo, J., & Kane, M. (2008). Cross-ideological discussions among conservative and liberal bloggers. *Public Choice*, **134**(1–2), 67–86. doi:10.1007/s11127-007-9201-x.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, **3**(4), 486–504. doi:10.1037/1082-989X.3.4.486.

Himelboim, I. (2011). Civil society and online political discourse: The network structure of unrestricted discussions. *Communication Research*, **38**(5), 634–659. doi:10.1177/0093650210384853.

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on Twitter. *Journal of Computer-Mediated Communication*, **18**(2), 40–60. doi:10.1111/jcc4.12001.

Ho, S. S., & McLeod, D. M. (2008). Social-psychological influences on opinion expression in face-to-face and computer-mediated communication. *Communication Research*, **35**(2), 190–207. doi:10.1177/0093650207313159.

Holland, P. W., & Leinhardt, S. (1971). Transitivity in structural models of small groups. *Comparative Group Studies*, **2**(2), 107–124. doi:10.1177/104649647100200201.

Holland, P. W., & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, **76**(373), 33–50. doi:10.2307/2287037.

Jensen, J. L. (2003). Virtual democratic dialogue? Bringing together citizens and politicians. *Information Polity*, **8**(1), 29–47.

Jurka, T. (2012). *Sentiment*: *Tools for sentiment analysis*. Retrieved from http://CRAN.R-project.org/package=sentiment

Kelly, J., Fisher, D., & Smith, M. (2005, May). *Debate*, *division*, *and diversity*: *Political discourse networks in USENET newsgroup*. Paper presented at the Online Deliberation Conference 2005, Stanford University, Stanford, CA.

Kim, J., Wyatt, R. O., & Katz, E. (1999). News, talk, opinion, participation: The part played by conversation in deliberative democracy. *Political Communication*, **16**(4), 361–385. doi:10.1080/105846099198541.

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, **115**(2), 405–450. doi:10.1086/599247.

Lee, E. J. (2007). Deindividuation effects on group polarization in computer-mediated communication: The role of group identification, public-self-awareness, and perceived argument quality. *Journal of Communication*, **57**(2), 385–403. doi:10.1111/j.1460-2466.2007.00348.x.

Leighley, J. E. (1990). Social interaction and contextual influences on political participation. *American Politics Quarterly*, **18**(4), 459–475. doi:10.1177/1532673x9001800404.

Lev-On, A., & Manin, B. (2009). Happy accidents: Deliberation and online exposure to opposing views. In T. Davies & S. P. Gangadharan (Eds.), *Online deliberation: Design, research, and practice* (pp. 105–122). Stanford, CA: CSLI.

MacKuen, M. (1990). Speaking of politics: Individual conversational choice, public opinion, and the prospects for deliberative democracy. In J. Ferjohn & J. Kuklinski (Eds.), *Information and democratic process* (pp. 59–99). Urbana: University of Illinois Press.

McGeough, R. E. (2010, November). *The market as the forum*: *Amazon.com discussion forums as deliberative spaces*. Paper presented at the annual convention of the National Communication Association, San Francisco, CA.

Moy, P., & Gastil, J. (2006). Predicting deliberative conversation: The impact of discussion networks, media use, and political cognitions. *Political Communication*, **23**(4), 443–460. doi:10.1080/10584600600977003.

Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy*. New York, NY: Cambridge University Press.

Mutz, D. C., & Martin, P. S. (2001). Facilitating communication across lines of political difference: The role of mass media. *American Political Science Review*, **95**(1), 97–114.

Mutz, D. C., & Mondak, J. J. (2006). The workplace as a context for cross-cutting political discourse. *Journal of Politics*, **68**(1), 140–155. doi:10.1111/j.1468-2508.2006.00376.x.

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, **6**(2), 259–283. doi:10.1177/1461444804041444.

Pingree, R. J. (2007). How messages affect their senders: A more general model of message effects and implications for deliberation. *Communication Theory*, **17**(4), 439–461. doi:10.1111/j.1468-2885.2007.00306.x.

Price, V., Nir, L., & Cappella, J. N. (2006). Normative and informational influences in online political discussions. *Communication Theory*, **16**(1), 47–74. doi:10.1111/j.1468-2885.2006.00005.x.

Rhee, J. W., & Kim, E. M. (2009). Deliberation on the Net: Lessons from a field experiment. In T. Davies & S. P. Gangadharan (Eds.), *Online deliberation: Design, research, and practice* (pp. 223–232). Stanford, CA: CSLI.

Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier*. Reading, MA: Addison-Wesley.

Rogers, E. M., & Bhowmik, D. K. (1970). Homophily-heterophily: Relational concepts for communication research. *Public Opinion Quarterly*, **34**(4), 523–538. doi:10.1086/267838.

Ryfe, D. M. (2005). Does deliberative democracy work? *Annual Review of Political Science*, **8**, 49–71. doi:10.1146/annurev.polisci.8.032904.154633.

Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, **29**(1), 1–63. doi:10.1017/S0047404500001019.

Schneider, S. M. (1996). Creating a democratic public sphere through political discussion: A case study of abortion conversation on the Internet. *Social Science Computer Review*, **14**(4), 373–393. doi:10.1177/089443939601400401.

Spears, R., Lea, M., & Lee, S. (1990). De-individuation and group polarization in computer-mediated communication. *British Journal of Social Psychology*, **29**(2), 121–134. doi:10.1111/j.2044-8309.1990.tb00893.x.

Stromer-Galley, J. (2003). Diversity of political conversation on the Internet: Users' perspectives. *Journal of Computer-Mediated Communication*, **8**(3). doi:10.1111/j.1083-6101.2003.tb00215.x.

Sunstein, C. (2001). *Republic.com*. Princeton, NJ: Princeton University Press.

Sykes, R. E. (1983). Initial interaction between strangers and acquaintances: A multivariate analysis of factors affecting choice of communication partners. *Human Communication Research*, **10**(1), 27–53. doi:10.1111/j.1468-2958.1983.tb00003.x.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. England: Cambridge University Press.

Wilson, T. P., Wiemann, J. M., & Zimmerman, D. H. (1984). Models of turn taking in conversational interaction. *Journal of Language and Social Psychology*, **3**(3), 159–183. doi:10.1177/0261927x8400300301.

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. doi:10.1086/653658.

Witschge, T. (2008). Examining online public discourse in context: A mixed method approach. *Javnost – The Public*, **15**(2), 75–92.

Wojcieszak, M., & Mutz, D. C. (2009). Online groups and political discourse: Do online discussion spaces facilitate exposure to political disagreement? *Journal of Communication*, **59**(1), 40–56. doi:10.1111/j.1460-2466.2008.01403.x.

Wojcieszak, M., & Price, V. (2012). Perceived versus actual disagreement: Which influences deliberative experiences? *Journal of Communication*, **62**(3). doi:10.1111/j.1460-2466.2012.01645.x.

Wright, S. (2006). Government-run online discussion fora: Moderation, censorship and the shadow of control. *The British Journal of Politics & International Relations*, **8**(4), 550–568. doi:10.1111/j.1467-856X.2006.00247.x.

Wu, F., & Huberman, B. A. (2010). Opinion formation under costly expression. *ACM Transactions on Intelligent Systems and Technology (TIST)*, **1**(1), 5. doi:10.1145/1858948.1858953.

Yardi, S., & Boyd, D. (2010). Dynamic debates: An analysis of group polarization over time on Twitter. *Bulletin of Science, Technology & Society*, **30**(5), 316–327. doi:10.1177/0270467610380011.

Yun, G. W., & Park, S. Y. (2011). Selective posting: Willingness to post a message online. *Journal of Computer-Mediated Communication*, **16**(2), 201–227. doi:10.1111/j.1083-6101.2010.01533.x.

Zhang, W., Cao, X., & Tran, M. N. (2013). The structural features and the deliberative quality of online discussions. *Telematics and Informatics*, **30**(2), 74–86. doi:10.1016/j.tele.2012.06.001.