Abstract

While internet spaces are believed to expose individuals to a wider array of viewpoints, a worry about self-reinforcing political echo chambers persists. With unprecedented choices online, do these choices lead to cross-cutting exposure or inevitably lead to increasing polarization? Instead of assessing political homogeneity online solely based on written messages, we focus on individuals' underlying motives and mechanisms that drives one's message "reception" decisions, and how such dynamics would manifest through patterns of individuals' message selection behaviors. Using unobtrusively logged behavioral data matched with panel survey responses, a TERGM analysis of online message selection behaviors during 2012 South Korean Presidential election indicates that the impact of *overt* partisan preference was rather limited. Rather, results indicate that various endogenous structural factors are pronounced, coupled with a non-trivial degree of message selection based on similarity of one's candidate evaluative criteria, suggesting that social and utility consideration indeed strongly override overt partisan considerations.

Keyword: Online political discussion, online discussion forum, message selection and exposure, Temporal exponential random graph model

The Dynamics of Message Selection in Online Political Discussion Forums:

Self-Segregation or Diverse Exposure?

Internet-based citizen communication, particularly discussion in online forums, is unique as compared to offline communication as it affords a situation that is relatively free from contextual constraints (Dahlgren, 2005). In online discussion forums, individual choices about what information to encounter and with whom to associate are relatively unconstrained. This unprecedented freedom of choice in an online setting raises the question of whether the enhanced choice leads to cross-cutting communication across ideological divides or homophilic interactions among the politically like-minded. Thus far, empirical endeavors to address this question have produced mixed findings (e.g., Gentzkow & Shapiro, 2011; Messing & Westwood, 2012). Some researchers suggest that online communication platforms, in keeping with the Harbermasian ideal of free and open space for civil society, are a pivotal space in which citizens from diverse backgrounds and with divergent viewpoints can voluntarily connect and interact (Papacharissi, 2004; Stromer-Galley, 2003; Wojcieszak & Mutz, 2009). This ideal cross-cutting space is believed to expose individuals to a wide array of perspectives, fostering quality and richness in citizen deliberation (Dahlgren, 2005). Yet, contrary to the optimistic view, there also is increasing concern that the Internet functions as self-reinforcing political echo chambers (e.g., Sunstein, 2009), which eventually creates ideological segregation and political enmity between partisan groups (Boutyline & Willer, 2017; Colleoni et al., 2014). Despite the increasing scholarly debate it has drawn thus far, how exactly the Internet has changed the landscape of everyday cross-cutting exposure is not yet clearly understood.

Within this context, the present study attempts to advance our understanding of the debate by focusing on *message selection* dynamics in online discussion forums. Although great progress has been made, much of the prior work has primarily been based on

participants' retrospective self-reports (e.g., Stromer-Galley, 2003; Wojcieszak & Mutz, 2009) or based solely on observable posted messages (e.g., Himelboim, 2008; 2011; Graham & Wright, 2014), each of which has inherent limitations. Self-reports provide versatile and flexible data as virtually everything can be measured, yet they are often questionable in term of measurement accuracy especially when it comes to behavioral constructs (e.g., Prior, 2009). Content posted on discussion forums or social network sites can only provide information between visibly connected dyads (i.e., post-reply relations between actors i and j). Thus, exposure to a message, whether cross-cutting or not, is observed only when actor i *replied* after reading actor j's message. This does not fully capture message selection dynamics because selecting (reading) a message does not necessarily result in replying to the message. In a similar vein, some work has explored one's selection behavior in online social networks (Boutyline & Willer, 2017; Colleoni et al., 2014), yet its focus is primarily on "channel" selection (e.g., "following" relations in Twitter), a decision that likely occurs *after* one is exposed to other's (presumably several) messages. Thus, there has been a lack of systematic investigation into individuals' *message selection* decisions – whether one chooses to *read* a given message in a forum – even before they choose to react and reply to the message. This oversight is particularly troubling since key to the debate over the role of the Internet in democracy is whether citizens are indeed exposed to (or they opt for) their fellow citizens' messages containing viewpoints that are different from their own.

Recognizing the limitations in extant research, we direct our attention to individuals' message *selection* (i.e., reading) when participating in an online discussion forum. Informed by past research suggesting that dynamics in communicative interactions "cannot be regressed to mere individuals' predispositions or pure social selection processes based on gender, race, or political viewpoints" (Song, 2015, p. 18; see also Lazer et al., 2010), we approach individuals' message selection behaviors from three different "layers" within which

such dynamics would unfold and manifest: individual motivations, dyadic homophily, and network structural features. Our goal here is to identify whether, and how, citizen's voluntary reading one another's messages results from a purposive search for political similarity or is instead mainly propelled by other motivations (i.e., understanding) and structural features of discussion network, which are less likely to pertain to (or be shaped by) one's overt partisan preference. If the latter is the case, then it indicates the possibility of incidental exposure to cross-cutting messages in online discussion networks. In doing so, we also stress the flip side of the coin – how such dynamics would manifest themselves in terms of whose messages are more likely to be *read* by others. Understanding these issues would shed light on how aggregate exposure patterns – as an end-result of one's message selections – emerge from individuals' message selection behaviors and what role online citizen communication would play in the democratic process.

In what follows, we emphasize two competing explanatory principles – *consistency* and *understanding* – as the two motivational drivers of citizen's online political interactions. Next, we further advance our perspective on how such competing principles could operate in an individual-level and in dyadic setting, and ultimately, how online discussion network structures could recursively influence individual's message selection. We then offer an empirical assessment using novel panel survey data matched with behavioral log data collected during a presidential election period from an online forum in which participants voluntarily posted their messages, and read and replied to messages posted by others. With detailed information of "who" selects "who's messages" and its correlates during a period of heightened attention to politics, our data are aptly suited for disentangling whether online message selection patterns are primarily structured along the partisan lines. Our results from an inferential network-analytic method called temporal exponential random graph model (TERGM) demonstrate that the impact of political preferences in shaping one's message

selection is much more limited than often assumed.

Two Motivational Drivers of Political Discussion: *Consistency* and *Understanding*

A recurring theme in the study of political communication is how much of citizens'

choice about what information to consume is driven by their political beliefs (e.g., Iyengar &

Hahn, 2009; Stroud, 2011) and how much of it is explained by other non-partisan

considerations (e.g., Messing & Westwood, 2012). Underlying these contrasting approaches

is the distinction between two different, albeit co-existing, fundamental human motivations

based on *cognitive consistency* vs. *understanding* principles (Holbert, Weeks, & Esralew,

2013, or *directional* vs. *accuracy* goals in Kunda, 1990). Based on the cognitive consistency

principle, for instance, balance-theoretic frameworks suggest that individuals prefer pro-

attitudinal messages which would lead them to their desired conclusions (e.g., Iyengar &

Hahn, 2009; Stroud, 2011). In contrast, the principle of *understanding* posits that people are

drawn to messages that they believe help them make sense of the situation at hand and reach

accurate conclusions (Holbert et al., 2013). Although these tendencies have been largely

discussed in the context of selective exposure to mass media sources and messages, they

provide a reasonable explanatory framework for message selection in online discussion

settings where interaction is much less spontaneous than face-to-face conversation and

discussants often have strong control over message choice (selectivity) as in the case of

consuming media messages. Within the context of the current study, we discuss how the two

principles of human motivations (i.e., *consistency* vs. *understanding*) shape the dynamics of

message selection in an online discussion forum at both individual (i.e., one selecting others'

messages and one's own message being selected by others) and dyadic (i.e., actor i selecting

actor j's message based on a dyadic characteristic between the actors) levels.

Principles of Consistency and Understanding at the Individual Level

When it comes to *out-going* message selection (i.e., reading others' message), it is

plausible that those who have high understanding motivation seek out messages that they

think are relevant and useful, regardless of whether such messages are pro- or counter-

attitudinal. Because of their motivation to have an accurate understanding, they would extend

more efforts in navigating and sorting through messages available in the online forum, and

this search would less likely be bounded by their political preferences. Thus, their strong

appetite for information would translate into an overall frequency of reading others'

messages. This is indeed in line with the previous findings that need for cognition (Cacioppo

et al., 1996) or accuracy motivations (Valentino et al., 2009) positively predict a host of

information seeking behaviors. On the other hand, those who are high on consistency

motivation would likely avoid messages that they disagree with, and thus their overall

information search would be more selective and narrower. In a non self-selected online forum

where both pro- and counter-attitudinal messages are present, people with strong consistency

motivation would likely engage in a somewhat partial information seeking due to their

political or partisan selectivity (Bennett & Iyengar, 2008) and thus spend less time reading

others' messages in general. This is especially likely when the valence of a new message

cannot be reliably predicted *before exposure* due to the presence of both pro- and counter-

attitudinal messages (e.g., Shook & Fazio, 2009). Recent research, however, suggests that

even strong partisans do not necessarily avoid attitudinally incongruent information *at all*

times (Garrett & Stroud, 2014; Valentino et al., 2009). This speaks to the possibility that

perhaps avoidance is not the only way for those with high consistency motivation to respond

to counter-attitudinal information. If this is the case, consistency motivation does not

necessarily reduce information seeking and message selection behaviors.

For patterns of *incoming* message selection (i.e., being read by others), we expect that

messages posted by those who have higher *consistency* motivations will likely communicate

strong, clear partisan perspectives (Ahn, Huckfeldt, & Ryan, 2014). These partisan cues

revealed in the posted messages would then attract attention from other participants. Given the widespread preferences towards attitudinally congruent information in general public (Garrett & Stroud, 2014), partisan language would function as a trigger for message selection at least by those on the same partisan side. Also, to the extent that one's message reflects their psychological desire for attitudinal consistency, the message might carry more elements of controversy or conflict which likely draw attention in general. In contrast, those with higher *understanding* motivations are less likely to be expressive of their partisan viewpoints, not least they are prone to make more considerate judgements concerning pros and cons from diverse political perspectives (Rudolph & Popp, 2007). Since their messages are less likely to contain strong, one-sided partisan information, they are less likely to be a target of message selection behaviors.

Although we expect that two different motivational tendencies have unique bearing on message selection patterns in online discussion forums, the rationales behind the relationships still remain speculative. We thus propose research questions, rather than hypotheses, as follows:

> **RQ**: How will (a) consistency motivation and (b) accuracy motivation be related to out-going and incoming message selection patterns, respectively, in an online discussion forum?

Principles of Consistency and Understanding at Dyadic Level

Above and beyond its impact at the pure individual-level (outgoing and in-coming selections), the cognitive consistency principle would play a role in message selection dynamics at the dyadic level. The notion of homophily, or the tendency of a given dyad to associate with each other based on their similarities (McPherson et al. 2001), has long been regarded as a powerful determinant of message selection decisions (Garrett & Stroud, 2014; Iyengar & Hahn, 2009; Song, 2015). Based on either the explicit application of political

preferences or on a de facto preference for similarity, research has repeatedly suggested that people can selectively construct their own social environment (Kossinets & Watts, 2009; Lazer et al., 2010; McPherson et al., 2001). Within the present context, this means that ego ("focal respondent") and alter ("potential discussion partner") are more likely to select each other's messages if they share similar political preferences. Therefore, we posit that:

H1: Participants in an online discussion forum will be more likely to select each

other's message when they share similar political preferences.

The understanding principle, on the other hand, would paint a somewhat different picture. We expect that due to human desire to reduce information cost in a decision situation (Downs, 1957; Pietryka, 2016), individuals tend to be inclined to search for information that is deemed to have high "utility" or "relevance," the meaning of which would vary depending on situations. In the midst of election campaigns, citizens' decision priority is on evaluating candidates and making up their mind who to vote for, and these decisions are often based on considerations about various factors including candidates' personal traits and backgrounds, party affiliations, and/or issue positions. Given that each individual voter would have their own evaluative criteria for voting decision, they would find messages containing elements that fit into the evaluative basis they have useful and relevant. In line with this expectation, Ahn and colleagues (2014) suggests that voters often actively glean relevant information from their social networks and they appear to value political expertise even at the absence of political agreement. Similarly, Hart et al. (2009) also shows that disconfirmation bias is substantially reduced for messages with higher informational value. We therefore expect that, in a dyad, two discussants with similar candidate evaluation criteria, whether they are like-minded or not, are more likely to find each other's message to be useful and relevant for their decision about who to support. Based on this rationale, our hypothesis is stated as follows:

H2: Participants in an online discussion forum will be more likely to select each

other's message when they share similar candidate evaluation criteria.

Endogenous Impact of Network Structure

While aforementioned factors are important aspects of message selection dynamics in its own right, they do not operate in a social "vacuum." As such, a theoretical perspective that ignores substantive interdependencies among actors is inevitably incomplete. Below, we attempt to explicate such interdependencies in explaining message exposure patterns.

Reciprocity

Often in online discussion forums, users not only intentionally seek information, but they also spontaneously exchange and respond to others' messages. This may take a number of different forms, yet the most simple and frequent form of such "interaction" may manifest as continuous, interactive message exchange sequences among a set of users. This also implies that such interaction patterns may require a situation in which actor i and actor j mutually choose to view each other's messages and return their attentions to each other – provided that replying to an original message necessitates a responder to actually click and read that message in first place. Based on this expectation, we hypothesize that reciprocity (Wasserman & Faust, 1994) would be a significant and positive predictor of online message selection, as follows:

> **H3**: Participants in an online discussion forum will be more likely to reciprocate
>
> message selection when his or her message has been selected by another participant.

Transitivity, Cyclic Closure, and Local Hierarchy

Transitivity and cyclic closure represent another mechanism of how individuals choose to encounter socially provided messages. Transitivity denotes a situation where node i is more likely to create a tie to node j when they are both connected to another node k. In contrast, cyclic closure denotes a similar but opposite situation where node j forms a tie to node i when they are connected to another node k (Holland & Leinhardt, 1976), as can be

seen in Table 1.

While the most common explanation for transitivity is that it reflects a local spread of social relations (e.g., "friends of my friends are my friends"), such explanation is somewhat less likely within the context of *message selection* in an online discussion forum. That is, a spread of social relationship requires actors to be aware of each other's social relationships in choosing one another to interact. However, within online discussion forum settings, information whether or not k has chosen to view j's messages (which is a prerequisite of a spread of social relations) is generally not available (or not visible) when i choose to view j's messages.

Instead, in light of understanding-based explanation, we propose an alternative possibility of transitivity: that a pattern of transitivity may arise from a hierarchical nature of underlying criteria by which people choose each other's messages. Here, individuals are assumed to pursue a tie with others whose messages exhibit higher status (e.g., argument quality, expertise, trustworthiness, etc.) than themselves. Thus, actor i is expected to seek to create a tie towards a "higher status" actor j (i.e., reading j's message), given i's existing relationship with an intermediate-status actor k who also has a tie to j. In this scenario, i does not necessarily have to be aware of k's tie to j, which often is invisible in online discussion forums. Rather, because of j's high status, j (his/her message) is sought by many individuals in the network including i and k, and if k's status is higher than i but less than j, the k will be sought by i but not by j. When coupled with a negative tendency towards cyclic closure (e.g., j is less likely to form a tie with less prestigious actor i), such a pattern can be interpreted as the local status hierarchy in a given network (Lazega et al., 2012). While this does not necessarily imply that people only purposively seek out higher status individuals based on message qualities at all times, yet evidence indicates that people routinely seek guidance from those who are more versed and politically sophisticated within their social networks

(Ahn et al., 2014; Downs, 1957; Huckfeldt, 2001). Consequently, one possible source of such hierarchical network structuring principles can be an individual's need for having political experts around and choose to view messages of those local experts. Therefore, we predict:

> **H4**: Participants in an online discussion forum will be more likely to select each other's message based on transitivity - a local hierarchy in the discussion forum.

Profile Similarity

Another mechanism that helps us understand the nature of message selection in online forums is the concept of profile similarity (DiMaggio, 1986). In addition to the hierarchical nature in message selection networks, individuals would be more likely to choose to view one another's message when they are both connected to other actors in the network in the same way. For instance, if actors i and j both choose to view the same set of alters ("activity closure"), or i and j are *chosen* by the same set of alters ("popularity closure": see Table 1 below), then the same patterns of incoming and out-going connections shared by i and j signal a common set of properties of the i-j dyad (Block & Grund, 2014; Robins et al., 2009). In such situation, i and j themselves are more likely to see each other's messages. In line with consistency-based explanation, this may be viewed as the structural bases of homophily, in that the formation of ties is driven by the similarity in choices with respect to other actors (DiMaggio, 1986). Therefore:

> **H5**: Participants in an online discussion forum will be more likely to select each other's message when they have similar selection patterns to all other participants (profile similarity).

Preferential Attachment

Many studies indicate that the structure of online social networks tends to follow a power-law distribution, characterized as the skewed distribution of degrees (Barabási & Albert, 1999). While the existence of such a pattern is rather common in online, it appears

that such tendency is also pronounced in online discussion forums. For instance,

Himelboim's (2008; 2011) analysis suggests a sharp inequality in the ability to draw attention

and elicit further engagement with a given message from a large number of users in online

discussion groups. When selecting which messages to click in an online discussion forum,

one often pay attention to certain heuristic cues such as the number of "views" and "likes,"

which signal "utility" based on the popularity of a message. Therefore, a message that has a

large number of engagement cues (such as views or likes) can draw disproportionate

selection behaviors by its self-reinforcing dynamics, leading to a highly imbalanced message

selection distribution. Therefore, we expect:

> **H6**: Participants in an online discussion forum will be more likely to select messages
>
> that are already selected by a large number of others.

Temporal Dynamics in Message Selection Criteria

As elections near, it is reasonable to believe that individuals are more mobilized by

campaign communication (Cho, 2013) and, thus, pay close attention to political messages

both online and offline. Not only is this more likely due to a heightened attention to politics,

but it is also likely because they need more information to reduce uncertainties or anxieties

about their voting decisions (Downs, 1957). While literature generally suggests that strong

partisans and interested voters arrive at their decisions early in the election campaign cycle

(Fournier et al., 2004), the day-to-day dynamics in the campaign environment may prompt

them to seek out confirmatory information. Specifically, an increase in uncertainty about the

election outcome may induce confirmatory information seeking behavior (Carnahan et al.,

2016; Valentino et al., 2009). As changes in the campaign environment (e.g., campaign

competitiveness) *over time* induce more anxiety and uncertainty about the election outcome,

the effect of preference homophily may increase. Therefore:

> **H7**: The effect of preference homophily in message selection increases over time.

Data and Methods

In order to test our predictions, we draw on a unique set of panel data collected during the 2012 South Korean presidential election. The data were collected from an online discussion forum hosted on a research firm's server where participants' posting and viewing activities during a 27 day-period until Election Day (from November 23 to December 19, 2012) were unobtrusively logged. A market research firm invited 400 participants from a nationally representative panel, of which a total of 341 participants completed all three waves of panel surveys. The surveys measured the participants' candidate evaluations and criteria, policy preferences, motivations for using online discussion forums, and other key covariates of interest.[1] Participants were instructed to post freely and read each other's posts about the upcoming election, as they normally would do in other online forums. Activity log data regarding posting and browsing behaviors were later retrieved from the research firm's server and matched with participants' survey responses.

In wave 1, 22 (6.5%) of the 334 participants had no candidate preference. Since candidate preference homophily is a key predictor in our model, we have excluded those who did not have a preferred candidate in all three waves (thus, $N = 312$). Yet an identical model including 22 missing cases with multiple imputation ($N = 5$) on candidate preference did not alter the results and conclusion reported herein.

Construction of Networks and Analysis Strategy

Over the period of data collection, participants on average posted 24.78 messages and read 547.31 postings made by others. Based on the participants' activity logs, we have

---

[1] Although slightly skewed in age (sample median age = 35; population = 38) and sex (sample = 51.9%; population = 49.67%), our final sample closely matches the general population in demographic profile. In addition, the sample's representativeness (the lack of it) is less of a concern because we are taking an inferential network-analytic approach. It is also noteworthy that our sample had enough variability in all of the key covariates (especially for candidate preference), making less likely that our results are biased by the peculiarity of our data.

derived a "message selection" network as a directed actor-actor binary matrix (312 x 312), such that the cell entry X_{ij} is defined as 1 when actor i chooses to view actor j's message and zero otherwise. As such, we also distinguish the direction of ties in this network (i.e., $X_{ij} \neq X_{ji}$). Based on the dates of the three panel surveys (W1 = Nov 27th to 29th, W2 = Dec 11th to 13th, W3 = Dec 21st to Dec 23th), we created longitudinal panel networks of message selection by partitioning log data from the first two waves and matching it to corresponding survey dates (e.g., log data from Nov 27th to 29th were regarded as the 1st wave of the network panel).[2] Since the 3rd wave of the survey was conducted *after* Election Day (which was Dec 19th) but electronic log data were only collected *until* Election Day, we regard the last three days of log data (Dec 17th to 19th) as the last panel in the network.[3] We consider the log data that were available four days prior to the first survey wave (Nov 27th) as well as that collected *between* each survey waves as lagged observations of the respective network panel. Specifically, log data from Nov 23rd to 26th were considered as lagged observations of the first network (Nov 27th to 29th), data from Nov 30th to Dec 10th as lagged observations of the second network (Dec 11th to 13th), and data from Dec 14th to 16th as lagged observations of the last network (Dec 17th to 19th).

Measures

Consistency and understanding motivations. For consistency motivation (Cronbach's α = .86, M = 4.36, SD = 1.03), respondents were asked six items (based on a 7-point scale from "Not at all" = 1 to "Very much" = 7) about whether they visit online discussion forums (including discussion forums other than in the current study) primarily "to

---

[2] We also estimated models with daily slices (t = 26) and found largely the same results with minor discrepancies in estimated coefficients and significance level. Combined with multiple imputation results, our robustness check suggest that our results and conclusions are reasonably robust against potential methodological issues.

[3] Since participants' key characteristics (e.g., candidate evaluations and preferences) are stable across survey waves, we assume participants' characteristics drive the creation of network ties (but not the other way around).

justify my opinion of the issue" or "to confirm that my opinion on the issue is correct."

Understanding motivation (α = .81, M = 5.26, SD = .82) was assessed in a similar manner

using four items (e.g., "to make an accurate and objective assessment of the issue", "to

understand others' opinions", etc.). Since motivations were measured only once at the first

wave of the survey, we regard these characteristics as time-invariant covariates in our model.

Preference homophily and evaluative criteria similarity. We define political

preference homophily (i.e., *consistency* principle) in two ways: (a) candidate choice, and (b)

policy preference. Candidate choice homophily (W1: M = .51, SD = .49; W2: M = .55, SD =

.49; W3: M = .52, SD = .49) was defined in a way that a given dyad was regarded as

homophilous (coded as "1") when they share the same candidate choice.[4] Policy preference

homophily (W1: M = .40, SD = .16; W2: M = .38, SD = .16; W3: M = .39, SD = .16, all range

= 0 to 1) was operationalized with respondents' preferences towards liberal vs. conservative

oriented policy options about economic and North Korea issues.[5] We derived a Euclidean

distance, d, of a given dyadic pair in terms of their dissimilarity in policy preferences, which

was later converted to similarity by taking $1 / (1 + d)$ to make a greater value represent

preference "homophily."

Next, we define candidate evaluation criteria similarity (M = .48. SD = .15, range = 0

to 1) in a similar manner, using a dyadic Euclidean distance d in terms of relative importance

of policy/candidate characteristics (e.g., policy, competence, integrity) versus personal

background (e.g., party affiliation, political career, place of origin, etc.) in candidate

evaluations. Since candidate evaluative criteria were only measured in the wave 1 survey, we

---

[4] Candidate choice (W1: M = .60, SD = .49; W2: M = .66, SD = .47; W3: M = .61, SD = .48) was tapped using a dichotomous measure, where "1" denotes supporting the liberal candidate (Moon Jae-in) vs. "0" denotes supporting the conservative candidate (Park Geun-hye).

[5] Policy preferences were measured three times across panel surveys using four-item measures, based on a 7-point scale from "not at all agree" (1) to "very much agree" (7).

treat this measure as invariant across waves.

Network-endogenous measures. Reciprocity was captured by whether a pair of actors mutually selected each other's messages. For measures tapping a series of triadic configurations (transitive closure, cyclic closure, activity closure, and popularity closure: see Table 1 for details), we relied on the *directed* version of the geometrically weighted edgewise shared partner (directed GWESP) statistics following the model specifications proposed by Snijders et al. (2006) and Robins et al. (2007). The GWESP term models a linear combination of an entire distribution of directed triangles (*i, h, j*) for a given connected dyad (*i, j*) in the network, and this effect is *weighted to produce a decreasing return* following a decay parameter (for a detailed discussion of this measure, see Hunter & Handcock, 2006). As described above, our theory suggests that a series of triadic closure patterns would have a substantial effect on message selection dynamics. Similarly, for measuring in- and out-degree distribution effects, geometrically weighted out-degree and in-degree distribution (GWD-out and GWD-in) terms were used where the parameter estimates for GWD terms represent "evenness" of in- and out-degrees based on message selection activities across the network (for details see Hunter, 2007). We expect these terms to be significant and negative, which would signify differential message selection activities across the network.

Control variables. In order to establish a plausible baseline in our analysis, we control for a host of variables that are known to be related to the extent to which people engage in political discussion. First, we control for participants' socio-demographic factors such as *gender* (1 *being* "female," 48.39%), *age* (in 10-year increment, $M = 3.55$, $SD = .98$), *education* (from "not finished elementary school" = 1 to "currently in post-graduate education or more" = 9, $M = 7.71$, $SD = .97$) and *region of origin* (1 being "Seoul" vs. 0 being "other regions", 40.38% from Seoul). In our analysis, we also controlled for two demographic homophily measures, one based on their gender and the other based on their

regional origin (all coded as 1 if a dyad shares the same gender or regional origin), since

preference homophily may be confounded with demographic homophily (McPherson et al.,

2000). We also controlled for respondents' offline discussion frequency (from "Never" = 1 to

"Always" = 7, W1: $M = 4.50$, $SD = 1.04$; W2: $M = 4.62$, $SD = 1.18$; W3: $M = 4.82$, $SD =$

1.17), news use frequency (measured in *hours*, W1: $M = .76$, $SD = .42$; W2: $M = 1.56$, $SD =$

1.66; W3: $M = 1.65$, $SD = 2.32$), internal discussion efficacy (from "Not at all agree" = 1 to

"Strongly agree" = 7, $M = 4.72$, $SD = .98$), and hedonic motivation ($\alpha = .75$, $M = 4.47$, $SD =$

1.04) for using online discussion forums. News use frequency was defined as the average

exposure in hours to Internet, newspaper and television news about the upcoming election.

Internal discussion efficacy was gauged using a four-item composite measure tapping how

competent and efficacious an individual is in a typical political discussion setting (e.g., "I am

competent at presenting my own opinions in a discussion"). Hedonic motivation was assessed

by a three-item measure, all anchored on a 7-point scale, asking whether they participate in

online forum based on pleasure-seeking motives (e.g., "it is interesting and fun").

Analysis Strategy

Since we aim to properly capture and explain substantive interdependency dynamics

over time, we modeled longitudinally observed message selection networks using a Temporal

Exponential Random Graph Model (TERGM), a time-series extension of the ERGM

framework with the bootstrapping resampling technique described in Desmarais and Cranmer

(2012). It is integral to this approach to model the ties in a given network to be a random

variable ("1" for existence of ties, and zero otherwise) to be explained simultaneously by a

collection of actor covariates and network-endogenous dependencies (Robins et al., 2007;

Snijders et al., 2006), while properly accounting for the non-independence of observations

inherent in network data. The ERGM framework is now regarded as the most versatile yet

flexible method for evaluating the underlying generative properties of a network, as

exemplified in recent applications of the method to various domains (Cranmer et al., 2017).

Since our analytical strategy requires all cell entries to be defined as binary rather

than integers, we opted to dichotomize numbers of selection instances within the same dyad

using the mean number of message selections across all dyadic pairs as a threshold (W1 =

2.5; W2 = 2.9; W3 = 3.1). Therefore, our model only speaks to relatively routine, repeated

message selection dynamics in a given network panel rather than all-inclusive message

selection dynamics, such as accidental, spontaneous selection behaviors. Also, in applying a

longitudinal inferential network analysis technique, we regarded an observation at a given

point in time as depending only on the previous state of the network (i.e., a lagged

observation). In capturing temporal dependencies, we include as additional control variables

a series of lagged endogenous network statistics that might be relevant in messages selection

behaviors, as well as a few additional endogenous network statistics (such as *isolates* and

two-paths) that are necessary for controlling temporal or lower-order effects when estimating

the effect of key parameters. Details of the model specification are provided in online

Supplemental Information. Table 1 below summarizes key model terms and their

corresponding hypothesis, with their graphical depiction and substantive interpretation.

[Table 1 About Here]

Once models were fitted, we assessed goodness-of-fit (*gof*) to identify the model

adequacy by simulating nine hundred new networks (three hundred new networks for each

time step) and compared the network characteristics from the observed vs. simulated

networks (Hunter, Goodreau, & Handcock, 2008). The *gof* results indicate that model

specification is satisfactory (see online Supplemental Information for details). All analyses

were based on maximum pseudo-likelihood estimation with bootstrapped confidence

intervals (Desmarais & Cranmer, 2012), as implemented in the *btergm* package in R (Leifeld

et al., 2017).

Results

Table 2 below reports the key parameter estimates from the final TERGM

specifications along with its 95% confidence intervals based on bias-corrected and

accelerated CIs using 1000 replications (also graphically reported in Figure S1 and in Table

S1 in online Supplemental Information). Relevant to our main interest, the leftmost model

specification ("Final Model") in Table 2 includes the effects of motivation and homophily

controlled for the hypothesized network structural influence, while a series of interaction

models from 2nd to 4th column test whether the effects of various preference homophily

increase over time. Across all models, coefficients can be interpreted as log odds of a tie

conditional on the rest of the network and other model terms.

[Table 2, Figures 1 and 2 About Here]

Our research question asked how consistency and understanding motivations

systematically affect the likelihood of messages *being selected* by other participants, as well

as an individual's selection patterns (i.e., selecting others' messages) within the online

discussion forum. For the final model specification, we found the effect of consistency

motivation to be non-significant in predicting outgoing selection (b = .025, 95% bootstrap CI

= [−.044, .077]), so as to understanding motivations predicting incoming selection (b =

−.052, [−.080, .022]). In contrast, we found a weak but significant tendency for consistency

motivation to predict in-ties (b = .034, [.009, .113]) and understanding motivation to predict

out-going ties (b = .028, [.005, .076]). Empirical patterns indicate that those who strive to

understand the outside world, as opposed to those with low understanding motivation, are

more likely to select and read others' messages in the online discussion forum. At the same

time, on average, people are more likely to select and read messages written by those with

higher consistency motivation.

Concerning our dyadic-level homophily variables, neither candidate choice

homophily ($b = -.032$, $[-.070, .047]$) nor policy preference homophily ($b = -.108$, $[-.212,$ .006]) is related to message selection. Thus, H1 is not confirmed. Such null effects indicate that consistency-driven dynamics (i.e., whether a dyad shares a candidate preference or ideological policy preference) is likely not related to whether people choose to select and view each other's messages. Instead, we find a consistent and substantial effect of similarity in candidate evaluative criteria, such that the more similar a dyad in terms of their candidate evaluative criteria, the more likely they are to expose themselves to another's messages (H2: $b = .407$, $[.399, .415]$). We return to the implications of this finding in the discussion section.

Our next set of hypotheses concerns the endogenous structural effects of network itself. As shown in Table 2, we have found consistent and robust support for these predictions, such that reciprocity (H3: $b = .768$, $[.560, 1.068]$), multiple cyclic closure (H4: $b = -.066$, $[-.076, -.061]$), multiple activity ($b = .035$, $[.033, .043]$) and multiple popularity closure ($b = .113$, $[.083, .232]$, all H5), and preferential attachment (*Popularity spread*, H6: $b = -4.123$, $[-5.343, -3.541]$) were all strongly supported, controlling for the tendency to not have any ties (*isolates*: $b = 1.003$), open triad without closing a triad (*multiple two-path*: $b = .003$, all CIs straddle zero), temporal dependencies, and other motivation and homophily terms.

Among estimated effects, the effect of preferential attachment (or an uneven degree distribution) was the strongest and substantial, as the negative incoming degree distribution parameter indicates (H6: $b = -4.123$). Figure 1 gives a substantive interpretation of the effect, suggesting that, irrespective of time periods, the predicted probability of receiving at least one additional message selection instance from other participants in the forum (excluding who are already connected) sharply increases as a function of the existing in-degree of a node. This suggests that message selection behaviors are largely driven by self-organizing dynamics, consistent with the notion that people are disproportionately drawn to and more

likely to expose themselves to *already popular* messages in a forum (Himelboim, 2008).

[Figure 1 About Here]

In addition to the effect of preferential attachment, participants in the online forum were approximately 2 times (conditional odds ratio = 2.15) more likely to browse others' messages based on a reciprocity effect. Likewise, an individual (ego) is approximately 4 to 12 percent more likely to read another participant's (alter) message for every one person increase in the number of other participants to whom the ego and alter are tied based on outgoing (*multiple activity closure*: conditional OR = 1.035) and incoming connection patterns (*multiple popularity closure*: conditional OR = 1.121). This suggests that when message selection patterns signal latent shared characteristics within a dyad, participants are more likely to select each other's message. Participants in our online forum were also slightly less likely to form a closed three-cycle, suggesting the network has a slight tendency against generalized exchange that returns to lower status individuals. The only exception to this pattern was the multiple path closure term (concerning H4: $b = .057$, $[−.053, .094]$), although the direction of the effect was in the expected direction.

Our last hypotheses predicted that as the election approaches, the impact of preference homophily in predicting message selection dynamics would increase. Among tested interaction terms, only candidate choice homophily is found to significantly interact with time trends (Interaction model I: $b_{interaction} = .051$, $[.038, .071]$). Specifically, the effect of candidate choice homophily is found to increase linearly over time, in a way that message selection in a dyad that shares the same candidate choice is more likely later in the election period, as plotted in Figure 3. Panel B of the Figure 2 gives Johnson-Neyman regions of significance as a function of time trends, additionally revealing that there is indeed a preference *towards heterophily* earlier in the election (as indicated in the negative conditional main effect: $b = −.135$, $[−.211, −.111]$). But this effect gradually disappears as a preference

for the same candidate choice increases. No other interaction terms emerged as significant.

[Figure 2 About Here]

Discussion and Conclusion

Even though prior literature has emphasized the deliberative potential of online

discussions (Papacharissi, 2004; Stromer-Galley, 2003), it is not uncommon to find worries

about self-reinforcing political echo chambers. Since the debate on whether online settings

promote more diverse and balanced exposure to political information is far from resolved, a

more comprehensive understanding of the underlying motivational and structural factors that

drive citizens' everyday discussion is much needed. Against this background, our study

emphasizes *consistency* and *understanding* as the two core explanatory principles of political

discussion online at individual- and dyadic-levels and highlights the role of various

endogenous structural factors that stem from the pattern of online discussion itself as the

crucial determinants of message selection dynamics. This study is among the first to provide

direct evidence that can disentangle the various determinants of message selection decisions

in an online discussion forum setting. Our findings suggest that, while there is some modest

tendency of message selection based on *both* consistency and understanding motivations

(especially at the individual level), the impact of *overt* partisan preference, as measured by

candidate choice homophily and policy preference homophily, is fairly limited. Instead, we

have observed a meaningful pattern of message selection driven by a dyadic similarity in

candidate evaluative criteria as well as robust and consistent effects of endogenous structural

factors on message selection. These results yield significant new insights and add important

nuance to our understanding of how people decide what to read in online discussion

settings.

In particular, we have found that those with higher understanding motivation actively

seek and expose themselves to others' messages. At the same time, those with high

consistency motivation are more likely to be the *target* in message selection dynamics (i.e.,

their messages are more likely to be selected by others). Yet, those with higher consistency

motivations are not necessarily more likely to seek – presumably confirmatory – social

information. Had it been significant, it would have indicated that those with higher

consistency motivations seek and are sought by mostly like-minded individuals, providing

support for the notion of ideological or partisan selectivity in online discussion settings.

However, our results seem to be more in line with Garrett (2009; also see Garrett et al., 2013)

or Bakshy et al. (2015), who find that balanced exposure is more common than is often

assumed. While our results also show that the preference for opinion-reinforcing information

is real (as indicated by the significant effect of consistency motivation predicting *incoming*

ties), this does not necessarily mean that people *only* seek confirmatory information.

　　　　More direct evidence supporting this perspective comes from the results of three

dyadic-level effects. That is, overt partisan homophily – either based on concrete candidate

choice or on abstract policy preferences – does not play a substantive role in message

selection dynamics. Although we have found that the impact of candidate choice homophily

had linearly increased over time, the magnitude of such effect was still limited throughout the

course of the election period. Instead, the similarity in candidate evaluation criteria – in other

words, judgmental criteria by which citizens evaluate candidates – have a substantial effect

throughout all models. It is a particularly noteworthy finding given that such similarity in

judgmental standards is not necessarily shaped by ideological or partisan like-mindedness.

Rather, consistent with understanding motivation, the result suggests that a utility

consideration – that is, specific information they can use in candidate evaluations *irrespective*

of its potential valence and partisan views – is one of the crucial factors determining message

selection. Such interactions (through message selection) between discussants who share a

similar set of evaluative criteria would likely have a potential for exposure to diverse

opinions and thus deliberative outcomes. Therefore, the results at the dyadic level as a whole

provide evidence that counters the notion of selectivity toward ideologically like-minded

messages *at the expense of* cross-cutting messages in online discussion settings. [6]

Across our analyses, preferential attachment emerged as the strongest predictor of

messages selection dynamics, corroborating recent studies about online (Himelboim, 2008;

2011) and offline political discussion (Song, 2015). Compared to studies of readily "visible"

interactions, such as post-reply relationships (Himelboim, 2008; 2011), our behavioral log

data examine selection behaviors that are not necessarily observable to participants. This

suggests that the global-level message selection dynamics are likely to be, at least partly,

driven by aggregate popularity cues (such as the number of "views" or "likes") that enable

participants to identify messages of higher social and informational utility. However, these

aggregate popularity cues do not necessarily give a sign to participants whether a given

message contains politically congenial messages. Considering that the magnitude of this

preferential attachment effect is nearly ten times greater than any of the homophily factors in

our model, we believe social and utility considerations strongly override overt partisan

considerations, echoing Messing and Westwood's (2014) recent finding on selective

exposure dynamics on social networking sites.

Of our findings on the structural factors, the results -- especially those about triadic

configurations -- warrant further discussion. Notably, we found significant and positive, yet

weak, "shared activity" and "shared popularity" effects. The patterns suggest that a pair of

participants who viewed the same set of messages, or whose messages were viewed by the

same set of people, are likely to see each other's messages. Particularly within the context of

---

[6] Although the effect of evaluative criteria similarity was more substantial between those who share the same candidate preference ($b_{interaction}$ = .324, [.039, .466]), the similarity in evaluative criteria had a significant and positive impact on the probability of message selection even among individuals who have different candidate preferences.

our study, cues indicating similar message selection patterns between two individuals are not available unlike in message – reply relationships. Therefore, our setting – which models "low visibility" message selection behaviors – make it particularly unlikely that these effects are driven by characteristics other than actual similarities in participants' criteria guiding their message selection (and/or message writing). At the same time, the extent of the similarities in *profiles* (i.e., message selection patterns) does not necessarily exclude the possibility that each discussant may have selected messages posted by different others for different reasons. Because of this, the notion of profile similarity predicts that eventually those who have a similar "overall" pattern of message selections would have a similar "set" of characteristics. As such, if a given dyad i and j choose to see each other's messages based on their similarities of message selection patterns to/from all other actors k, then it implies that i and j have lots of common (i.e., "homophily") in their attributes. This raises the possibility that people may choose to engage with each other based not on just a single characteristic (such as candidate preference) but some balance (or a sum) of multiple characteristics ("multidimensional homophily": Block & Grund, 2014). Yet it should be acknowledged that, although this enables multitudes of attributes to be simultaneously involved in consideration of "homophily," the exact underlying nature of such homophily still remain elusive as to whether it is driven by overt partisan considerations or by other incidental factors. Even if we interpret these patterns as the support for consistency-driven dynamics, the substantive magnitudes of such effects appear fall short of other "understanding" driven factors.

In consideration of this study's findings, we conclude with a few caveats. First, although the coefficient was in the expected direction, we did not find the expected transitive closure effect. While we do not have any definitive explanation for this, it may be that local-level, hierarchy-based dynamics (as measured by transitive closure effect) become non-significant when there is a strong influence of global-level hierarchies produced by

preferential attachment (again, it has an almost ten-fold increase in its impact than transitive closure). In ERGM, since both triadic closure and degree distributions lead to local clustering while they tend to be highly correlated (Levy et al., 2015), strong global hierarchies produced by a degree-related effect may leave almost no room for a weak, local-level triadic closure term in explaining the emergence of a hierarchical structure of the network.

Second, following our theoretical focus, this study has operationalized "links" among participants as directed message "reading" behaviors. While this is an important addition to the existing literature that focuses largely on self-reports (e.g., Stromer-Galley, 2003; Wojcieszak & Mutz, 2009) or written (posted) messages (e.g., Himelboim, 2008; 2011), our model did not consider the characteristics of the messages themselves. Indeed, it is conceivable that individual's message selection behaviors were at least partly driven by some textual cues available in thread titles (as the first textual cue that respondents would encounter in selecting other's messages), or by some other interactions between network dynamics identified here and message characteristics. While it is arguably an important avenue for future research that would add nuance to our results, it requires consideration of how latent textual topics and observed message quantities are probabilistically generated and how such factors would further condition the observed network dynamics in a stochastic fashion. To our knowledge, a proper probabilistic model addressing such issues is only recently beginning to be developed (e.g., Kim et al., 2017).

Third, regarding the interaction effect between candidate choice homophily and time trends, we also acknowledge that such patterns may have been driven by participants' "learning effects" than campaign competitiveness effect. That is, based on their continued interactions in the forum, they could have learned about others' partisan orientations, which makes them better able to discern partisan leanings of socially provided messages over time. While our finding could be explained by this alternative explanation, it should be also noted

that, ultimately, the impact of candidate choice homophily never exceeded that of other understanding-based effects. As such, the overall results speak to the conclusion that individuals' message selection (thus exposure) patterns are not necessarily self-segregated along the overt partisan lines.

Lastly, we close by also recognizing that our single-country, single election data approach may not generalize to other contexts. Yet, given how similar our results are to those of other online (Himelboim, 2008; 2011) and offline (Song, 2015) political discussion studies from a considerably different geographical and electoral context, we see little reason to expect that the basic underlying mechanisms identified in our study would not be applicable in different times and contexts.

Throughout the paper, we highlighted the notion that online discussion settings do not necessarily create polarized message selection patterns because fundamental human motivations – *consistency* vs. *understanding* – are expected to play important roles in structuring the way people decide to read others' messages, the very first step that could initiate a communicative interaction. Consistent with previous evidence (Bakshy et al., 2015; Garrett & Stroud, 2014; Messing & Westwood, 2014), we find that individuals do not organize their message selection based solely on overt partisan considerations. Instead, message selection patterns in line with understanding goals have been observed. Further, echoing evidence by Lazer et al. (2010) and Song (2015), our results also demonstrate that the endogenous structures of an online discussion network, which have less to do with individuals' overt partisan preferences and directional goals, systematically shape individuals' message selection behaviors, thus having the potential to accidentally expose discussants to cross-cutting political messages. While a self-isolation from different perspectives is still possible online, this study suggests it is not an unavoidable consequence of conscious individual choice.

Reference

Ahn, T. K., Huckfeldt, R., & Ryan, J. B. (2014). *Experts, activists, and interdependent citizens: Are electorates self-educating?* New York: Cambridge University Press.

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

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

Block, P., & Grund, T. (2014). Multidimensional homophily in friendship networks. *Network Science, 2*(2), 189–212.

Boutyline, A., & Willer, R. (2017). The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political Psychology, 38*, 551-569.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*(2), 197.

Carnahan, D., Garrett, R. K., & Lynch, E. K. (2016). Candidate vulnerability and exposure to counterattitudinal information: Evidence from two US presidential elections. *Human Communication Research*, *42*(4), 577-598.

Cho, J. (2013). Campaign tone, political affect, and communicative engagement. *Journal of Communication, 63*(6), 1130-1152.

Cranmer, S. J., Leifeld, P., McClurg, S. D., & Rolfe, M. (2017). Navigating the range of statistical tools for inferential network analysis. *American Journal of Political Science*, *61*(1), 237-251.

Dahlgren, P. (2005). The Internet, public spheres, and political communication: Dispersion and deliberation. *Political Communication, 22*, 147-162.

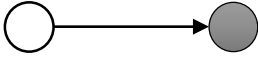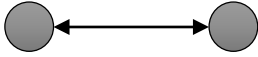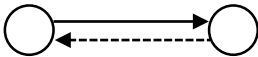Desmarais, B. A., & Cranmer, S. J. (2012). Statistical mechanics of networks: Estimation and
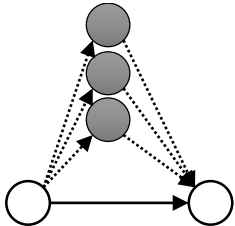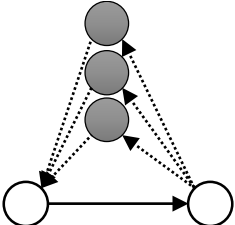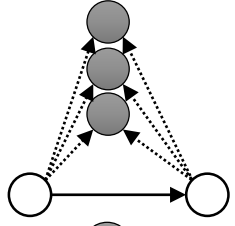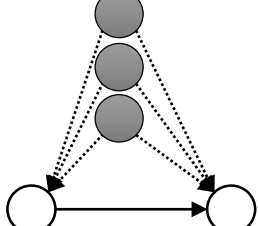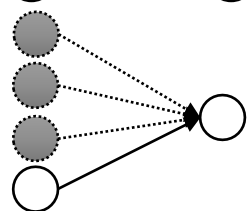
uncertainty. *Physica A*, *391*(4), 1865-1876.

DiMaggio, P. (1986). Structural analysis of organizational fields: A blockmodel approach.

    Research in Organizational Behavior, 8, 335-370.

Downs, A. (1957). *An economic theory of democracy*. New York: Harper.

Fournier, P., Nadeau, R., Blais, A., Gidengil, E., & Nevitte, N. (2004). Time-of-voting

    decision and susceptibility to campaign effects. *Electoral Studies, 23*(4), 661-681.

Garrett, R. K. (2009). Politically motivated reinforcement seeking: Reframing the selective

    exposure debate. *Journal of Communication*, *59*(4), 676-699.

Garrett, R. K., & Stroud, N. J. (2014). Partisan paths to exposure diversity: Differences in

    pro-and counterattitudinal news consumption. *Journal of Communication*, *64*, 680-

    701.

Gentzkow, M., Shapiro, J. M. (2011). Ideological segregation online and offline. *The*

    Quarterly Journal of Economics, 126(4), 1799-1839.

Graham, T., & Wright, S. (2014). Discursive equality and everyday talk online: The impact

    of "superparticipants". *Journal of Computer-Mediated Communication*, *19*(3), 625-

    642.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009).

    Feeling validated versus being correct: a meta-analysis of selective exposure to

    information. *Psychological Bulletin, 135*(4), 555-588.

Himelboim, I. (2008). Reply distribution in online discussions: A comparative network

    analysis of political and health newsgroups. *Journal of Computer-Mediated*

    Communication, 14(1), 156-177.

Himelboim, I. (2011). Civil society and online political discourse: The network structure of

    unrestricted discussions. *Communication Research, 38*(5), 634-659.

Holbert, R. L., Weeks, B. E., & Esralew, S. (2013). Approaching the 2012 U.S. Presidential

election from a diversity of explanatory principles: Understanding, Consistency, and

Hedonism. *American Behavioral Scientist, 57*(12), 1663-1687.

Holland, P. W., & Leinhardt, S. (1976). Local structure in social networks. *Sociological*

Methodology, 7, 1-45.

Huckfeldt, R. (2001). The social communication of political expertise. *American Journal of*

Political Science, 425-438.

Hunter, D. R., & Handcock, M. S. (2006). Inference in curved exponential family models for

networks. *Journal of Computational and Graphical Statistics*, *15*(3), 565-583.

Iyengar, S., & Hahn, K. S. (2009). Red media, blue media: Evidence of ideological

selectivity in media use. *Journal of Communication*, *59*(1), 19-39.

Kim, B., Schein, A., Desmarais, B. A., & Wallach, H. (2017). *A network model for dynamic*

textaul communications with appliation to government email corpora. Paper

presented at the 10th Political Networks Conference, The Ohio State University,

Columbus, OH.

Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social

network. *American Journal of Sociology*, *115*(2), 405-450.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin, 108*(3), 480-498.

Lazega, E., Mounier, L., Snijders, T., & Tubaro, P. (2012). Norms, status and the dynamics

of advice networks: A case study. *Social Networks, 34*(3), 323-332.

Lazer, D., Rubineau, B, Chetkovich, C., Katz, N., & Neblo, M. (2010). The coevolution of

networks and political attitudes. *Political Communication, 27*(3), 248–274.

Leifeld, P., Cranmer, S. J., & Desmarais, B. A. (2017). *btergm*. Temporal Exponential

Random Graph Models by Bootstrapped Pseudolikelihood. R package version 1.9.0.

Levy, M., Lubell, M., Leifeld, P., & Cranmer, S. (2016). *Interpretation of geometrically*

weighted degree estimates in exponential random graph models. Paper presented at

the 9th Political Networks Conference, Washington University, Saint Louis, MO.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in

social networks. *Annual Review of Sociology*, *27*(1), 415-444.

Messing, S., & Westwood, S. J. (2014). Selective exposure in the age of social media:

Endorsements trump partisan source affiliation when selecting news

online. *Communication Research*, *41*(8), 1042-1063.

Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential

of online political discussion groups. *New Media & Society*, *6*(2), 259-283.

Pietryka, M. T. (2016). Accuracy motivations, predispositions, and social information in

political discussion networks. *Political Psychology, 37*, 367-386.

Prior, M. (2009). The immensely inflated news audience: Assessing bias in self-reported

news exposure. *Public Opinion Quarterly, 73*(1), 130-143.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential

random graph (*p**) models for social networks. *Social Networks, 29*, 173–191.

Rudolph, T. J., & Popp, E. (2007). An information processing theory of ambivalence.

Political Psychology, 28(5), 563-585.

Shook, N. J., & Fazio, R. H. (2009). Political ideology, exploration of novel stimuli, and

attitude formation. *Journal of Experimental Social Psychology, 45*(4), 995-998.

Snijders, T. A., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New specifications

for exponential random graph models. *Sociological Methodology, 36*(1), 99-153.

Song, H. (2015). Uncovering the structural underpinnings of political discussion networks:

Evidence from an exponential random graph model. *Journal of Communication,

65*(1), 146-169.

Stromer-Galley, J. (2003). Diversity of political conversation on the Internet: Users'

perspectives. *Journal of Computer-Mediated Communication*, *8*(3). 0-0.

Stroud, N. J. (2011). *Niche news: The politics of news choice.* Oxford University Press.

Sunstein, C. R. (2009). *Republic.com 2.0*. Princeton University Press.

Valentino, N. A., Banks, A. J., Hutchings, V. L., & Davis, A. K. (2009). Selective exposure in the Internet age: The interaction between anxiety and information utility. *Political Psychology*, *30*(4), 591-613.

Wasserman, S., & Faust, K. (1994). *Social network analysis*. Boston, MA: Cambridge University Press.

Wojcieszak, M. E., & Mutz, D. C. (2009). Online groups and political discourse: Do online discussion spaces facilitate exposure to political disagreement? *Journal of Communication*, *59*, 40-56.

Table 1. Key TERGM parameters, associated configurations, and their interpretations

Hypothesis	Configuration	Interpretation
RQ: Motivation		A select B's message (B's message is selected by A) based on nodal attributes
H1 & H2: Homophily		A and B select each other's message based on their shared characteristics
H3: Reciprocity		A select B's message when B also select A's message
H4: Multiple path closure (GWESP-OTP)		A select B's message when A has multiple intermediary actors that also leads to B (implies status differentials)
H4: Multiple cyclic closure (GWESP-ITP)		A select B's message when B has multiple intermediary actors that also leads to A (implies lack of status differential)
H5: Multiple activity closure (GWESP-OSP)		A select B's message when they have similar patterns of message selection patterns (implies similarity in latent attributes)
H5: Multiple popularity closure (GWESP-ISP)		A select B's message when their messages are similarly selected by others (implies similarity in latent attributes)
H6: Preferential attachment (GWD-in)		A select B's message when many others also selected B's message

* Preferential attachment is measured using geometrically weighted in-degree distribution statistics, which measures *unevenness* of in-degree distribution. Therefore, *negative* GWD-in statistic means *positive* preferential attachment pattern (Levy et al., 2015).

Table 2. Bootstrapped TERGM estimates (95% bias-corrected and accelerated confidence intervals within brackets)

	Final Model	Interaction I	Interaction II	Interaction III
Edges (Intercept)	**-1.890** [-2.932; -1.392]*	**-1.819** [-2.732; -.304]*	**-1.823** [-2.807; -1.169]*	**-1.936** [-2.937; -1.098]*
Motivation and homophily				
Consistency motivation (in-ties) (RQ)	**.034** [.009; .113]*	.037 [-.004; .113]	**.037** [.010; .113]*	**.037** [.010; .113]*
Consistency motivation (out-ties) (RQ)	.025 [-.044; .077]	.019 [-.112; .071]	.019 [-.112; .071]	.019 [-.043; .071]
Understanding motivation (in-ties) (RQ)	-.052 [-.080; .022]	-.049 [-.103; .022]	-.049 [-.103; .022]	-.049 [-.078; .022]
Understanding motivation (out-ties) (RQ)	**.028** [.005; .076]*	**.036** [.012; .075]*	**.035** [.011; .087]*	**.035** [.011; .075]*
Same candidate preference (H1)	-.032 [-.070; .047]	**-.135** [-.211; -.111]*	-.033 [-.079; .047]	-.032 [-.079; .047]
Similar policy preference (H1)	-.108 [-.212; .006]	-.091 [-.225; .042]	-.090 [-.230; .042]	.094 [-.764; .272]
Similar evaluative criteria (H2)	**.407** [.399; .415]*	**.385** [.260; .404]*	.295 [-.359; .639]	**.389** [.255; .405]*
Interaction (H7)				
Time trends (linear)		.079 [-.059; .262]	**.083** [.021; .171]*	**.144** [.063; .235]*
x Same candidate preference		**.051** [.038; .071]*		
x Similar evaluative criteria			.046 [-.176; .242]	
x Similar policy preference				-.095 [-.253; .214]
Endogenous structural effects				
Reciprocity (H3)	**.768** [.560; 1.068]*	**.768** [.559; 1.068]*	**.768** [.507; 1.068]*	**.768** [.560; 1.068]*
Path closure (gwesp-OTP: H4)	.057 [-.053; .094]	.057 [-.053; .125]	**.057** [.025; .125]*	.057 [-.053; .094]
Cyclic closure (gwesp-ITP: H4)	**-.066** [-.076; -.061]*	**-.066** [-.076; -.061]*	**-.066** [-.080; -.061]*	**-.066** [-.076; -.061]*
Activity closure (gwesp-OSP: H5)	**.035** [.033; .043]*	**.035** [.033; .041]*	**.035** [.033; .043]*	**.035** [.033; .043]*
Popularity closure (gwesp-ITP: H5)	**.113** [.083; .232]*	**.113** [.083; .232]*	**.113** [.098; .232]*	**.113** [.083; .232]*
Popularity spread (gw-indegree: H6)	**-4.123** [-5.342; -3.541]*	**-4.120** [-5.342; -3.537]*	**-4.121** [-4.810; -3.259]*	**-4.123** [-5.342; -3.541]*

* = zero outside the 95% bias-corrected and accelerated confidence interval using 1000 replications, with significant results being bolded.

Note: All models control for age, gender (including homophily), education, regional origins (including homophily), offline talk frequency, media use frequency, candidate preference, hedonic motivations, Activity spread (gw-outdegree), being isolate, and multiple two-paths (gwdsp). Full results, including interaction models, are reported in the Online Supporting Information.
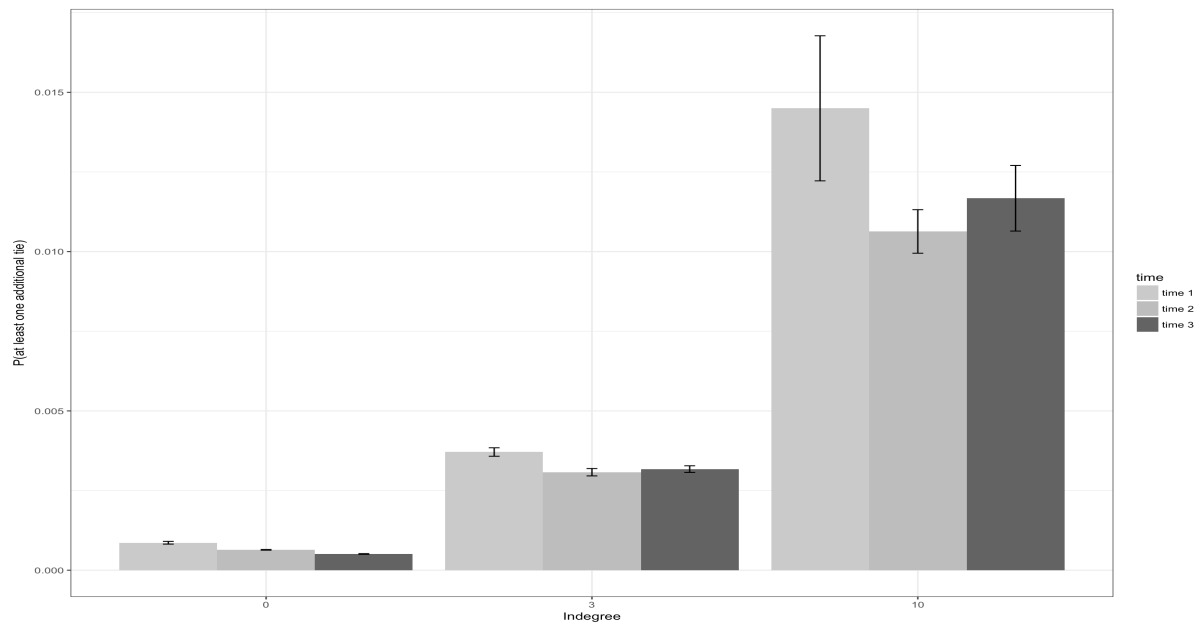
Figure 1. Mean predicted probabilities of *receiving* at least one additional tie (i.e., message being selected by others) as a function of existing incoming ties at 10% (= zero), 50% (= three), and 90% (= ten existing ties) percentile of the in-degree distribution. For each receiver node, we derived the mean edge probabilities of all other nodes (excluding any nodes that are already connected) sending a tie to the target node conditional on the rest of the network and on the model specification.
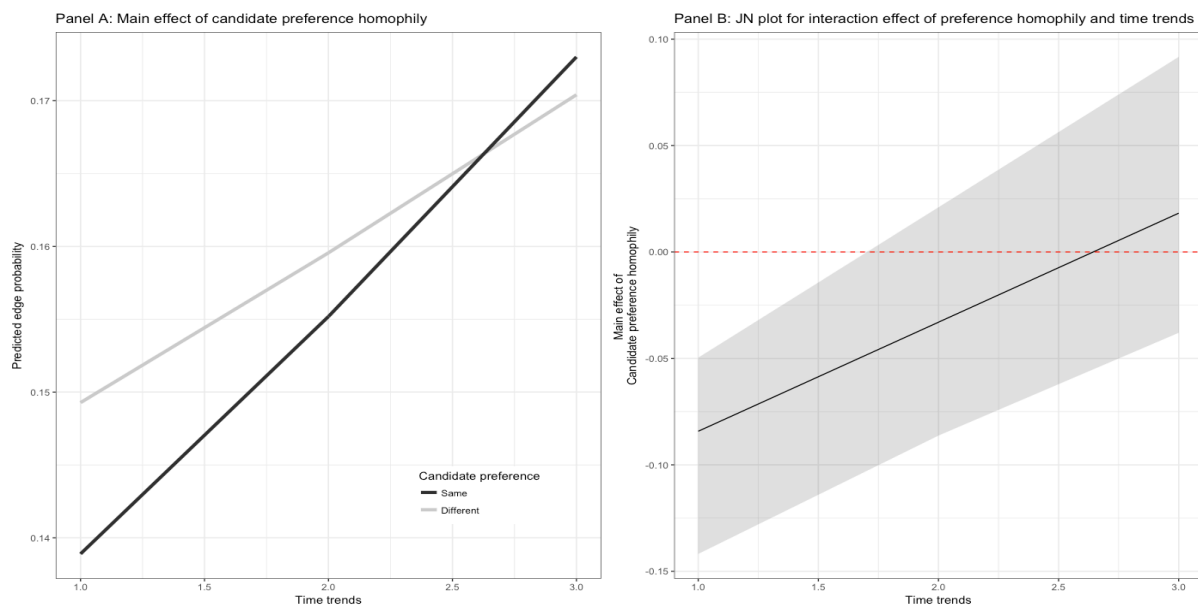

Figure 2. Interaction effects between time trends and candidate preference homophily. Panel A depicts conditional main effects of candidate preference homophily at each time point, and Panel B depicts Johnson-Neyman regions of significance as a function of time.