**The dynamics of message exposure in online political discussion forums:**

**Effects of motivation, homophily, and endogenous network processes**

**Expectations and Hypotheses**

***Motivations at Individual Level***

From the perspective of the cognitive consistency principle, all things being equal, individuals are more likely to choose to view the message that conforms their expectations (McPherson, Smith-Lovin, & Cook, 2001; Iyengar & Hahn, 2009). Yet at the same time, prior findings suggest that partisans with high consistency motivation are no more likely to “avoid” potentially dissonant messages (Garrett, 2009; Garrett & Stroud, 2016). This leads us to expect that higher consistency motivation is associated with higher level of information seeking behavior, irrespective of messages’ congeniality. At the same time, it is also expected that their messages are more likely to *be selected by others*, presumably since those with high consistency motivation communicate clear, strong partisan messages (Ahn, Huckfeldt, & Ryan, 2014). These two expectations culminate to our first set of hypotheses:

**H1a**: Consistency motivation is positively associated with the propensity of one’s messages being selected and read by others in the online discussion forum.

**H1b**: Consistency motivation is positively associated with one’s propensity of selecting others’ messages in the online discussion forum.

From the angle of the understanding principle, those with higher understanding motivation are likely to seek out and carefully processing relevant information, similar to the findings that need for cognition positively predict a host of information seeking behaviors (Cacioppo et al., 1996; Tsfati & Cappella, 2005). Yet compared to consistency motivation, those with understanding motivations are less likely to clearly communicate partisan messages, let alone they are presumably less expressive of their partisan viewpoints. This would lead them to be less likely to be selected by others compared to those with higher consistency motivation. Formally, we expect:

**H2a**: Understanding motivation is negatively associated with the propensity of one’s messages being selected and read by others in the online discussion forum.

**H2b**: Understanding motivation is positively associated with one’s propensity of selecting others’ messages in the online discussion forum.

In addition to consistency and understanding motivations, hedonic motivation – or an idea that people seek to gain pleasure and enjoyment – is another important motivational underpinning of why people use media and interact with each other. That is, those who found using online discussion forum and interacting with others more pleasurable and enjoyable would be generally inclined to be remain active and more participatory than otherwise. Yet while it is expected that those with higher hedonic motivations to be more active (i.e., more likely to view others’ messages and engage with others), it is not entirely clear whether and how such hedonic motivation is related to the propensity of *being selected by others*. Therefore, we simply expect following:

**H3**: Hedonic motivation is positively associated with the propensity of selecting others’ messages in the online discussion forum.

***Homophily at Dyadic Level***

The cognitive consistency principle further leads us to hypothesize a positive impact of partisan preference homophily in their message selection dynamics. Either based on an explicit application of political preferences or based on de facto preference homophily based on other similar characteristics, research has repeatedly suggested that people can selectively construct their social environment around them (Kossinets & Watts, 2009; McPherson et al., 2001). Within the present context, this mean an ego (“focal respondent”) and alters (“potential discussion partner”) are more likely to select each other’s messages if they share same political preferences. Therefore, we posit that:

**H4**: Same candidate preference (H4a) and similar policy preference (H4b) within a dyad increase the propensity of selecting each other’s messages.

In addition, we expect voters of similar candidate evaluation criteria are more likely to select each other’s message, irrespective of their congeniality towards their preference. This is based on the expectation that such information is of high utility to make relevant judgments regarding whom they should (or should not) support for. Prior literature agrees while voters actively glean relevant information from their social networks, they also appear to value political expertise more than shared preferences (Ahn, Huckfeldt, & Ryan, 2014). Hart et al.’s (2009) research, for instance, have found that disconfirmation bias is substantially reduced when encountered with messages of higher informational value. Since messages that are similar in terms of judgmental criteria (on which others make candidate evaluations) may contain highly relevant information and signal utilities, voters are more likely than otherwise to select such messages – especially when they are motivated to make accurate evaluations towards political candidate. Formally, we expect that:

**H5**: Similarity in candidate evaluation criteria is positively associated with the propensity of selecting each other’s messages in the online discussion forum.

***Endogenous Influence of Network Structure***

*Reciprocity.* The notion of reciprocity, or the extent of which the relationships between actors in a social network are symmetric (Wasserman & Faust, 1994), represents one of the fundamental dynamic processes in which how individuates create and maintain their social relationship (Snijders, 2011; Wasserman & Faust, 1994). Previous studies therefore often find positive tendency towards reciprocity in many empirically observed social networks. Such positive tendency towards reciprocity is often found within an online discussion context as well (e.g., Hagemann, 2002; Graham & Wright, 2013).

Often in online discussion forums, users not only intentionally seek for certain information, but they also spontaneously exchange, respond, and react to others’ opinions and messages. This presumably may take a number of possible forms such as providing simple feedback cues (“likes” or “dislikes”), recommend such posts to others, and leaving a comment, etc. Yet the most simple and frequent form of such “interaction” may manifest as continuous, interactive “discussion thread” – message exchange sequences – among a set of members. This also implies that such interaction patterns may have create a situation of which an actor *i* and actor *j* mutually choose to view each other’s messages, direct and return their attentions to each other – provided that leaving a reply or comments to an original message necessitate a responder to actually click and read that message at first place. Based on this expectation, we hypothesize that reciprocity would be one of the significant and positive predictor of presence of ties within an online discussion network:

**H6**: There would be more than expected by chance likelihood of reciprocity in message selection pattern within a dyad.

*Transitivity, cyclic closure, and local hierarchy.* Transitivity and cyclic closure may represent another fundamental social process of which how individuals select which messages to read, determining the overall message exposure patterns online. The concept of transitivity, or “triadic closure,” denotes situation where nodes *i* is more likely than chance to form a relation to another node *j* when they are connected to *k* other nodes (Holland & Leinhardt, 1975). In contrast, cyclic closure denotes similar situation for node *j* to form a tie to node *i* when they are connected to *k* other nodes, as can be seen in Table 1 below.

It is worth noting that transitive closure can signify several different underlying mechanisms of which one can select potential alters in social network; While the most common explanation for transitive closure is that it reflects a local spread of social relations (e.g., “friends of my friends are my friends”), such a pattern also reflects the closure of structural hole, in that node *i* circumvents brokerage role of other node *k* in reaching out another node *j*. Another, equally plausible possibility is that a positive tendency for transitivity reflects a hierarchical nature of a given network, such that node *i* would seek to create a tie towards a higher status node *j* given the exiting relations with intermediate-status node *k*. This expectation is especially true when the network exhibits negative tendency towards cyclic closure (e.g., *j* is less likely to form a tie to low status individual *i* despite the positive tendency of *i* to form a tie *j*), where such pattern is interpreted as the lack of generalized exchange due to local status hierarchy in a given network.

It is important to nothing that, within the context of “message selection” dynamics in an online political discussion forum, the information of whether or not *k* has chosen to view *j*’s messages is not available to *i* when *i* choose to view *j*’s messages (unless such information is explicitly visible via some functionalities in the system). Therefore, it is somewhat less likely that transitivity would reflect local spreads of social relationship, which requires actors to be aware of others’ social relationship in choosing others to interact. Within the context of predicting triadic configuration, it is therefore more plausible to assume that transitivity patterns arise from the hierarchical nature of underlying criteria in which people choose each other’s messages. Indeed, it is well documented that people’s political expertise level is not evenly distributed (Delli Carpini & Keeter, 1996; Downs, 1957; Verba, Schlozman, & Brady, 1995), and people routinely rely on and seek guidance from those who are more politically attentive and knowledgeable (Huckfeldt, 2001; McClurg, 2006). Therefore, one possible source of such hierarchical organization of network structure can be an individual’s need for having political experts around and choose to view messages of those local experts. Assuming the underlying tie-generative process is indeed driven by such substantive interests, it is conceivable that the uneven level of political expertise within a triad would be manifested via a hierarchically organized message selection dynamic. Therefore:

**H7a**: There would be more than expected by chance likelihood of transitive closure in message selection pattern among set of three actors.

**H7b**: There would be less than expected by chance likelihood of cyclic closure in message selection pattern among set of three actors.

*Structural equivalence and profile similarity.* Another important local configuration that help us understand the nature of message selection dynamics in online forums is the concept of structural equivalence and profile similarity. In addition to the hierarchical nature of underlying criteria in which people choose each other’s messages, they choose to interact with each other because they both connected to the same way to other actors in the network. That is, similar to the notion of structural equivalence, they maintain similar pattern of connections to all other actors in the network, such that they choose to view messages from the many same alters (“activity closure”), or they are chosen by same many alters (“popularity closure”: see Table 1 below for the respective diagram), which signals the common properties of a given dyads. This may be viewed as structural bases of homophily, whereby the formation or maintenance of ties are driven by similarity in choices with respect to other actors in the system (DiMaggio, 1986). Therefore, we expect following:

**H8a**: There would be more than expected by chance likelihood of activity closure in message selection pattern among set of three actors.

**H8b**: There would be more than expected by chance likelihood of popularity closure in message selection pattern among set of three actors.

*Preferential attachment.* Several studies indicate that a structure of large, online social network tends to follow power-law distribution. While the existence of skewed degree distribution is rather common (Barabási & Albert, 1999; Snijders, 2011), it appears that such tendencies are more pronounced in online context. For instance, Fisher, Smith, and Welser (2006) found highly imbalanced distribution of message posting and attraction in Usenet newsgroup discussions. Likewise, Himelboim’s (2011) analysis suggests a sharp inequality in ability to draw attention and elicit further engagement with a given message from a large number of users in online discussion groups. This is at least partly explained by the principle of “preferential attachment,” the idea that well-connected nodes or actors draw more connections by virtue of their already exiting connections. Within an online discussion forum, one often employs certain heuristic cues such as the number of “views” and “likes” in selecting which messages to click, which signals utility based on popularity of a message. Therefore, a message that has large number of engagement cues (such as views or likes) usually draw disproportionate reactions by its self-reinforcing dynamics, leading to highly imbalanced distribution of message selections among members. Therefore, we expect:

**H9**: There would be more than expected by chance likelihood of selecting messages when such messages 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; Holbrook & McClurg, 2005), hence are expected to pay close attention to political messages both in online and offline. Not only a heightened attention to politics in general more likely to make them to do so (Song & Boomggaarden, 2017; Zaller, 1992), but they also may need more information to reduce uncertainties or anxieties regarding their decisions as the election day approaches (Atkin, 1973; Downs, 1957). While literature generally suggests that strong partisans and interested voters arrive their decisions early in the election campaign cycle (Lazarsfeld, Berelson, & Gaudet, 1944/1968; Fournier, Nadeau, Blais, Gidengil & Nevitte, 2004), the nature and extent of changes in campaign environment (e.g., campaign competitiveness) may prompt even strong partisans to seek out confirmatory information. Specifically, increases in uncertainty regarding the ultimate consequences of election outcome may further propel confirmatory information seeking behavior (Carnahan, Garrett, & Lynch, 2016; Knobloch-Westerwick & Kleinman, 2012; Valentino, Banks, Hutchings, & Davis, 2009). Literature also suggests if there’s no reason to believe counter-attitudinal information at hand is useful for reducing decision-related uncertainty and anxiety, then individuals are more expected to rely on confirmatory evidence (Valentino et al., 2009). To the extent that changes in campaign environment (e.g., campaign competitiveness) *over time* induce more anxiety and uncertainty regarding the election outcome (which is a reasonable assumption to make), then this further suggests that the effect of various forms of homophily may increase over time until the election day rather than being constant over the course of campaign. Therefore, we posit:

**H10**: The effect of preference homophily in message selection increases over time.

**Data and Methods**

In order to test our predictions, we draw a unique set of whole network panel with survey responses collected during the 2012 South Korean presidential election. The data were collected from an online discussion forum hosted on research firm’s server where participants’ posting and viewing activities during 27 day-period until Election day (from November 23 to December 19, 2012) is unobtrusively logged. A market research firm invited 400 participants from a nationally representative panel, of which a total of 334 participants remained on the discussion forum and completed three waves of panel surveys. Surveys measured participants’ candidate evaluations and its criteria, policy preferences, motivations for using the online discussion forum, and other key covariates of interest. Participants were instructed to freely post and read each other’s opinions regarding upcoming election as they normally would in other online forums in return for a monetary incentive of $100. Activity log data regarding participants’ message viewing and posting activities were later retrieved from the research firm’s computer server and matched with participants’ survey responses.

At the start of the wave 1 survey, 22 participants (6.5%) out of all 334 participants did not initially identify their candidate preference nor had favored one of two major candidates based on relative thermometer ratings. Since we control for actors’ candidate choices and regard candidate preference homophily as a key predictor in our model, we limit the analysis to those with known candidate choices across all three survey waves (N = 312).[[1]](#endnote-1)

**Construction of Networks and Analysis Strategy**

Based on electronical log of participants’ message browsing behaviors, we derive a “message viewing” 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 for otherwise. Based on the dates of three panel survey responses (W1 = Nov 27th to 29th, W2 = Dec 11th to 13th, W3 = after the election day, which was Dec 19th, 2012), we partition log data in a way that it closely matches with survey dates in creating a longitudinal panel series of message exposure networks (e.g., log data from Nov 27th to 29th were regarded as the 1st wave of the network panel), except for the last wave of the network panel. Since the 3rd wave of the survey was conducted *after* the election day whereas electronic log data were collected *only until* the election day, we regard the last three days of log data (Dec 17th to 19th) as the last panel in network.[[2]](#endnote-2) In addition, since the log data were available from November 23rd, the log data *before* the first wave of panel survey (Nov 27th) or *in-between* each survey waves were regarded as lagged observation of the respective network panel. Specifically, we treat data from Nov 23rd to 26th as the lagged observations of the first network while treating data from Nov 27th to 29th as the 1st wave of the network. Likewise, log data from Nov 30th to Dec 10th constitute lagged observation of the second network (Dec 11th to 13th) while log data from Dec 14th to 16th constitute lagged observation of the last network (Dec 17th to 19th).[[3]](#endnote-3)

**Measures**

**Motivations for using online discussion forum.** 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) whether they visit online discussion forums (including discussion forums other than current study) primarily “to justify my opinion of the issue” or “to confirm that my opinion on the issue is correct.” Understanding motivations (α = .81, *M* = 5.26, *SD* = .82) and hedonic motivations (α = .75, *M* = 4.47, *SD* = 1.04) were assessed in a similar manner, respectively using four (e.g., “to make an accurate and objective assessment of the issue”) and three items (e.g., “it is interesting and fun”).

**Preference homophily**. We define three different measures of political preference homophily based on (a) candidate choice, (b) ideological policy preference, and (c) candidate evaluative criteria. First, a candidate preference homophily was defined in a way that a tie was identified as homophilous (coded as “1”) if a given dyad shares the same candidate preference (“1” supporting Moon Jae-in vs. “0” supporting Park Geun-hye; W1: *M* = .60, *SD* = .49; W2: *M* = .66, *SD* = .47; W3: *M* = .61, *SD* = .48). Next, ideological 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 based on respondents’ dyadic Euclidean distance, *d*, out of liberal vs. conservative oriented policy preferences towards economic and north-Korea issues. Policy preferences were measured three times across panel surveys, and respective Euclidean distances were later converted to similarity measures by taking 1 / (1 + *d*), in a way that a greater value of the similarity would represent higher degree of homophily. Lastly, we define candidate evaluative criteria homophily (*M* = .48. *SD* = .15, range = 0 to 1) in a similar manner, 1 / (1 + *d*), using a dyadic Euclidean distance *d* in terms of relative importance of competence/impression (e.g., policy, competence, or perceived personal characters such as integrity) and personal background dimensions (e.g., party affiliation, political career, place of origin, etc.) in making candidate evaluations. Since candidate evaluative criteria was measured only once (at Wave 1 survey), we regard this measure to be invariant across waves.

**Network-endogenous measures.** Reciprocity of message selection relation was measured by whether a pair of actors had mutual “selection” ties with each other. For measures tapping a series of triadic configurations (transitive closure, cyclic closure, activity closure, and popularity closure: see Figure 1 for details), we rely on *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 of the triangle distribution is *weighted to produce a decreasing return* following a decay parameter (for a detailed discussion of this measure, see Hunter, 2007 and 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 activity spread and popularity spread, 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-degree based on message selection activities across the network (for details see Hunter, 2007). We expect these terms to be significantly and highly negative, which signify differential message selection activities across the network.

**Control variables.** In addition to focal predictor variables, we control for a host of variables that are known to be related to the extent of political discussion in order to establish a plausible baseline in our analysis. First, we control for participants’ sociodemographic factors, including *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 control for two demographic homophily variables, one based on their gender and the other based on their regional origin (all coded as 1 if a dyad share same gender or same regional origin) since demographic homophily may be confounded with respondents’ candidate preference homophily (McPherson et al., 2000). We also control 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), media use frequency (measured in hours, W1: *M* = .76, *SD* = .42; W2: *M* = 1.56, *SD* = 1.66; W3: *M* = 1.65, *SD* = 2.32), and internal discussion efficacy (from “Not at all agree” = 1 to “Strongly agree” = 7, *M* = 4.72, *SD* = .98). Media use frequency was defined as the average hour of exposure to internet news, newspaper and television news exposure regarding the upcoming election, and internal discussion efficacy were gauged using a four-item composite measure tapping how competent and efficacious an individual is in typical political discussion settings.[[4]](#endnote-4)

**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 bootstrapping resampling technique as described in Desmarais and Cranmer (2012). The integral part of this approach is to regard the ties in a given network to be a random variable (1 for existence of ties, and zero for 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 account for the nonindependence of observations inherent in network data. The ERGM framework is now regarded as the most versatile yet flexible method for evaluating underlying generative properties of a network, as exemplified in recent applications to various domains such as policy collaboration (Cranmer et al., 2017; Lai, She, & Ye, 2015; Shumate, Fulk, & Monge, 2005), diffusion of online news (Kim, Baek, & Kim, 2015), mediated campaigning (Song et al., 2017), friendship formation (Goodreau et al., 2009; Lewis et al., 2008), and offline political discussion (Song, 2015).

Since our analytical strategy requires all cell entries are defined as binary, we opted for dichotomizing multiple number of selection instances within a same dyad by employing mean number of message selection instances across all dyadic pairs as a threshold. Therefore, our model only speaks to relatively routine, repeated message selection dynamics in a given network panel rather than entire message selection dynamics including accidental, spontaneous selection behaviors. Also, in applying a longitudinal inferential network analysis technique, we also regard observation at a given time point is dependent only upon the previous state of the network (i.e. lagged observation). In capturing temporal dependencies, we include series of lagged endogenous network statistics which might be relevant in messages selection behaviors as additional control variables, along with few additional endogenous network statistics (such as *isolates* and *two-paths*) in order to control temporal or lower-order effects in estimating the effect of key parameters. Details on the applied models are provided in the Appendix. Table 1 below summarize key model terms included in our analysis, with their graphical depiction and substantive interpretation of the effects.

[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 compare the network characteristics from the observed vs. simulated networks (Hunter, Goodreau, & Handcock, 2008). The *gof* results indicate that model specification is satisfactory (results available upon request). All analyses were based on maximum pseudolikelihood estimation with bootstrapped confidence intervals (Desmarais & Cranmer, 2012), as implemented in the *btergm* package in R (Leifeld, Cranmer, & Desmarais, 2017).

**Results**

Table 2 below reports the estimated parameters from the final TERGM specifications along with its 95% confidence intervals (based on bias-corrected and accelerated CIs using 1000 replications, with significant results being bolded), and this is also graphically reported in Figure 1 below (full results are available upon request). Relevant to our main interest, the leftmost model specification (“Final Model” in Table 2) includes the effects of motivation and homophily while properly controlling for hypothesized network structural influence, while a series of interaction models from 2nd to 4th columns test whether the effects of various preference homophily terms increases 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, Figure 1 and 2 About Here]

Our first set of hypotheses posit that messages written by individuals with certain motivations are more likely to be selected by others (H1a & H2a), as well as such individuals are more likely to select others’ messages (H1b, H2b & H3). We found mixed support for these expectations, with some of the hypothesized effects fell short of significance level although all of the coefficients were in line with the hypothesized direction. For final model, we found the effect of consistency motivation being nonsignificant in predicting outgoing messages selection instances (“select others’ messages”: b = .025, 95% bootstrap CI = [−.044, .077]), so as to understanding motivations predicting incoming ties (b = −.052, [−.080, .022]). In contrast, we found a weak (but significant) support for consistency motivation predicting in-ties (b = .034, [.009, .113]) and understanding motivation predicting out-going ties in message selection network (b = .028, [.005, .076]), supporting H1a and H2b. Empirical patterns indicate that those who pursue to better understand the outside world are more likely to select and read others’ messages in online discussion forums (compared to those who are low on understanding motivations), while on average people are more likely to select and read messages written by those with higher consistency motivation. In addition to those findings, hedonic motivation is found to be significantly and positively related to outgoing message selection instances (b = .102, [.087, .133]), supporting H3.

Concerning our dyadic-level homophily variables, neither candidate preference (H4a: b = −.032, [−.070, .047]) nor ideological policy preference homophily (H4b: b = −.108, [−.212, .006]) found to be related to the message selection instances, fail to confirm H4. Such null effects of two preference homophily terms indicate that consistency-driven dynamics (i.e., whether one shares same candidate preference or ideological policy preference) is not likely to be related to whether people choose to select and view each other’s messages. Instead, we have found consistent and quite substantial effect of candidate evaluative criteria similarity, such that the more similar a dyad in terms of their candidate evaluative criteria, the more they likely are to expose themselves to another’s messages (H5: b = .407, [.399, .415]).

Our next set of hypotheses concerns endogenous structural effects of network itself. As shown in Table 2, we have found consistent and robust support for these predictions, such that reciprocity (H6: b = .768, [.560, 1.068]), multiple cyclic closure (H7b: b = −.066, [−.076, −.061]), multiple activity closure (H8a: b = .035, [.033, .043]), multiple popularity closure (H8b: b = .113, [.083, .232]), and preferential attachment (*Popularity spread*, H9: b = −4.123, [−5.343, −3.541]) were all strongly supported, controlling for the tendency for not having 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, notably the effect of preferential attachment (or uneven degree distribution) was most strongest and substantial, as the negative incoming degree distribution parameter indicates (H9: *b* = −4.123). Figure 2 gives substantive interpretation of the effect, such that predicted probabilities of receiving at least one additional message selection tie (excluding ties that are already connected) from other participants in the forum sharply increases as a function of existing in-degree of a node, irrespective of time periods. This suggests that messages selection dynamics are largely based on self-organizing dynamics, consistent with the notion that people are disproportionately drawn upon and more likely to expose themselves to already popular messages in a forum (Himelboim, 2008).

In addition to the effect of preferential attachment, participants in the online forum are approximately 2 times (*reciprocity*, conditional odds ratio = 2.15) more likely to read potential alter’s message if that alter were already have read his or her message. Likewise, an individual is approximately 4 to 12 percent more likely to read potential alter’s message for every one person increase in a same way that they are both tied to based on outgoing (*multiple activity closure*: conditional OR = 1.035) or incoming connection patterns (*multiple popularity closure*: conditional OR = 1.121), which signals latent shared characteristics between a dyad. They were also slightly less likely to form a closed three-cycle, suggesting the network has a slight tendency against generalized exchange that returns to a lower status individuals. The only exception for this pattern was the multiple path closure term (H7a: b = .057, [−.053, .094]), although the direction of the effect was again in the expected direction.

Our last hypotheses predicted that as the election approaches, the impact of preference homophily in predicting message selection dynamics would be increased. Among tested interaction terms, only candidate preference homophily is found to significantly interact with time trends (Interaction model I: *b*interaction = .051, [.038, .071]). Specifically, the effect of candidate preference homophily is found to be linearly increasing over time, in a way that message selection among a dyad that share the same candidate preference is more likely as the later in the election period, as plotted in Figure 3. Panel B of the Figure 3 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 conditional main effect: b = −.135, [−.211, −.111]) but this effect gradually disappears as the preference towards same candidate choice increases. No other interaction terms emerged as significant.

[ Figure 3 About Here]

Discussion and Conclusion

Our results suggests XXXX.

Reference

TBA

Footnotes

Table 1. TBA

**Panel A: transitive closure and cyclic closure.**



*Note*: Positive transitivity closure and negative cyclic closure means node i is more likely than chance to send a tie to j, but j are less like than chance to send a tie back to i, suggesting network has a hierarchical structure.

**Panel B: Activity closure and popularity closure.**

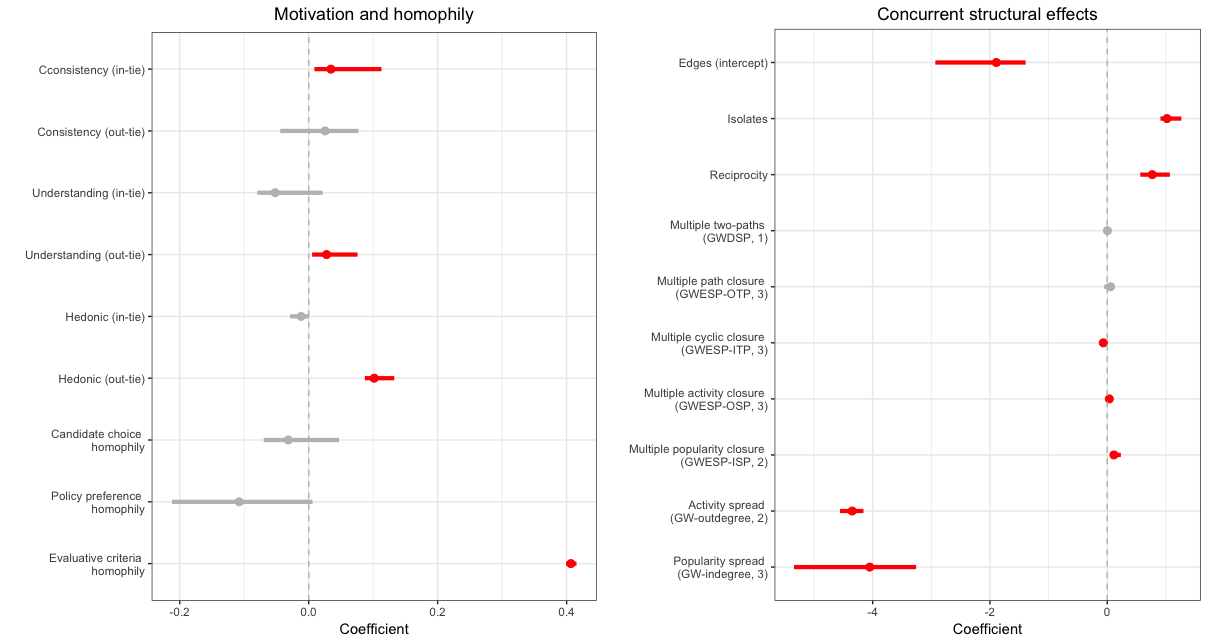


*Note*: Positive shared activity closure and shared popularity closure means that node i and node j are selecting and being selected by many common actors, suggesting that they are in common in terms of certain nodal characteristics. This signifies structural bases of homophily in their characteristics.

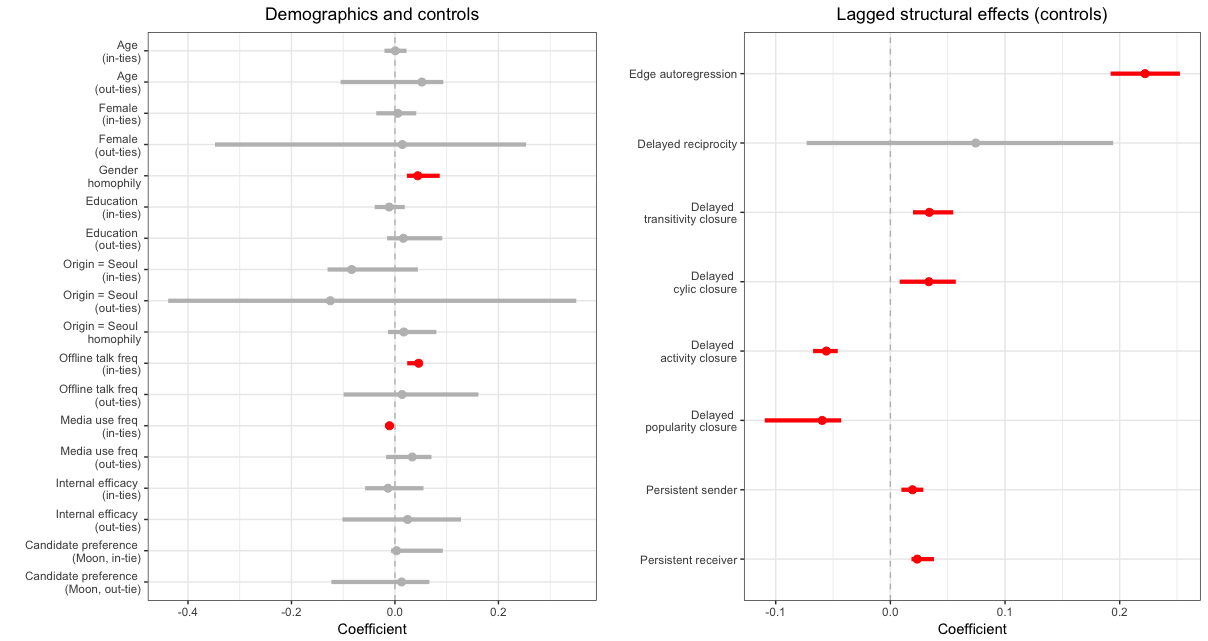
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Final Model** | **Interaction I** | **Interaction II** | **Interaction III** | |
| Edges (Intercept) | **-1.890** [-2.932; -1.392]\* | **-1.819** [-2.732; -.304]\* | **-1.823** [-2.807; -1.169]\* | **-1.936** [-2.937; -1.098]\* | |
| ***Motivation and homophily*** |  |  |  |  | |
| Consistency motivation (in-ties) (H1a) | **.034** [.009; .113]\* | .037 [-.004; .113] | **.037** [.010; .113]\* | **.037** [.010; .113]\* | |
| Consistency motivation (out-ties) (H1b) | .025 [-.044; .077] | .019 [-.112; .071] | .019 [-.112; .071] | .019 [-.043; .071] | |
| Understanding motivation (in-ties) (H2a) | -.052 [-.080; .022] | -.049 [-.103; .022] | -.049 [-.103; .022] | -.049 [-.078; .022] | |
| Understanding motivation (out-ties) (H2b) | **.028** [.005; .076]\* | **.036** [.012; .075]\* | **.035** [.011; .087]\* | **.035** [.011; .075]\* | |
| Hedonic motivation (in-ties) | -.012 [-.029; .001] | -.012 [-.038; .001] | -.013 [-.032; .001] | -.013 [-.038; .001] | |
| Hedonic motivation (out-ties) (H3) | **.102** [.087; .133]\* | **.102** [.094; .130]\* | **.102** [.096; .130]\* | **.102** [.094; .105]\* | |
| Same candidate preference (H4a) | -.032 [-.070; .047] | **-.135** [-.211; -.111]\* | -.033 [-.079; .047] | -.032 [-.079; .047] | |
| Similar policy preference (H4b) | -.108 [-.212; .006] | -.091 [-.225; .042] | -.090 [-.230; .042] | .094 [-.764; .272] | |
| Similar evaluative criteria (H5) | **.407** [.399; .415]\* | **.385** [.260; .404]\* | .295 [-.359; .639] | **.389** [.255; .405]\* | |
| ***Interaction (H10)*** |  |  |  |  | |
| 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*** |  |  |  |  | |
| Isolates | **1.003** [.793; 1.264]\* | **1.005** [.793; 1.152]\* | **1.005** [.895; 1.264]\* | **1.003** [.793; 1.264]\* | |
| Reciprocity (H6) | **.768** [.560; 1.068]\* | **.768** [.559; 1.068]\* | **.768** [.507; 1.068]\* | **.768** [.560; 1.068]\* | |
| Multiple path closure (H7a) | .057 [-.053; .094] | .057 [-.053; .125] | **.057** [.025; .125]\* | .057 [-.053; .094] | |
| Multiple cyclic closure (H7b) | **-.066** [-.076; -.061]\* | **-.066** [-.076; -.061]\* | **-.066** [-.080; -.061]\* | **-.066** [-.076; -.061]\* | |
| Multiple activity closure (H8a) | **.035** [.033; .043]\* | **.035** [.033; .041]\* | **.035** [.033; .043]\* | **.035** [.033; .043]\* | |
| Multiple popularity closure (H8a) | **.113** [.083; .232]\* | **.113** [.083; .232]\* | **.113** [.098; .232]\* | **.113** [.083; .232]\* | |
| Multiple two-paths | .003 [-.007; .007] | .003 [-.007; .007] | .003 [-.007; .009] | .003 [-.007; .007] | |
| Activity spread | **-4.395** [-4.557; -4.153]\* | **-4.392** [-4.557; -4.152]\* | **-4.392** [-4.557; -3.994]\* | **-4.395** [-4.557; -4.153]\* | |
| Popularity spread (H9) | **-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. All models control for age, gender (including homophily), education, regional origins (including homophily), offline talk frequency, media use frequency, and candidate preference. | | | | |

Table 2. Bootstrapped TERGM estimates (95% BCa confidence intervals within brackets).

Panel A: key predictors



Panel B: Control variables



*Figure 1*. Parameter estimates and 95% confidence intervals from the final model.

Note. Significant model terms are denoted in red.

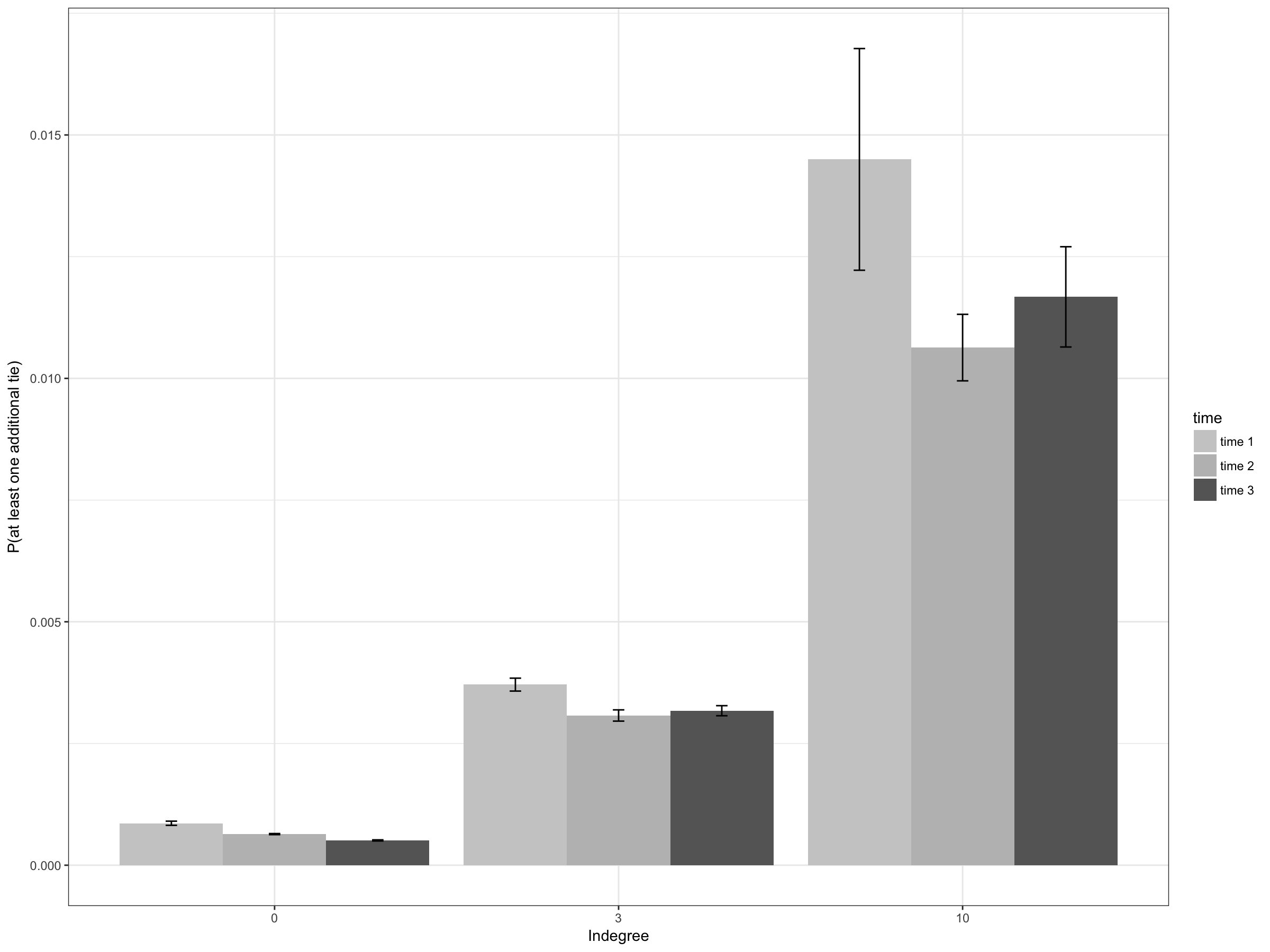


Figure 2. Mean predicted probabilities of *receiving* at least one additional tie (i.e., message being selected by others) as a function of existing incoming ties at 10% (zero), 50% (three), and 90% (ten) percentile of the in-degree distribution. Predicted probabilities are based on all eligible nodes of respective in-degrees. For each receiver node, we derived the mean edge probabilities of all other nodes (excluding any nodes that are already connected to a target node) sending message selection to the target node conditional on the rest of the network and on the model specification.

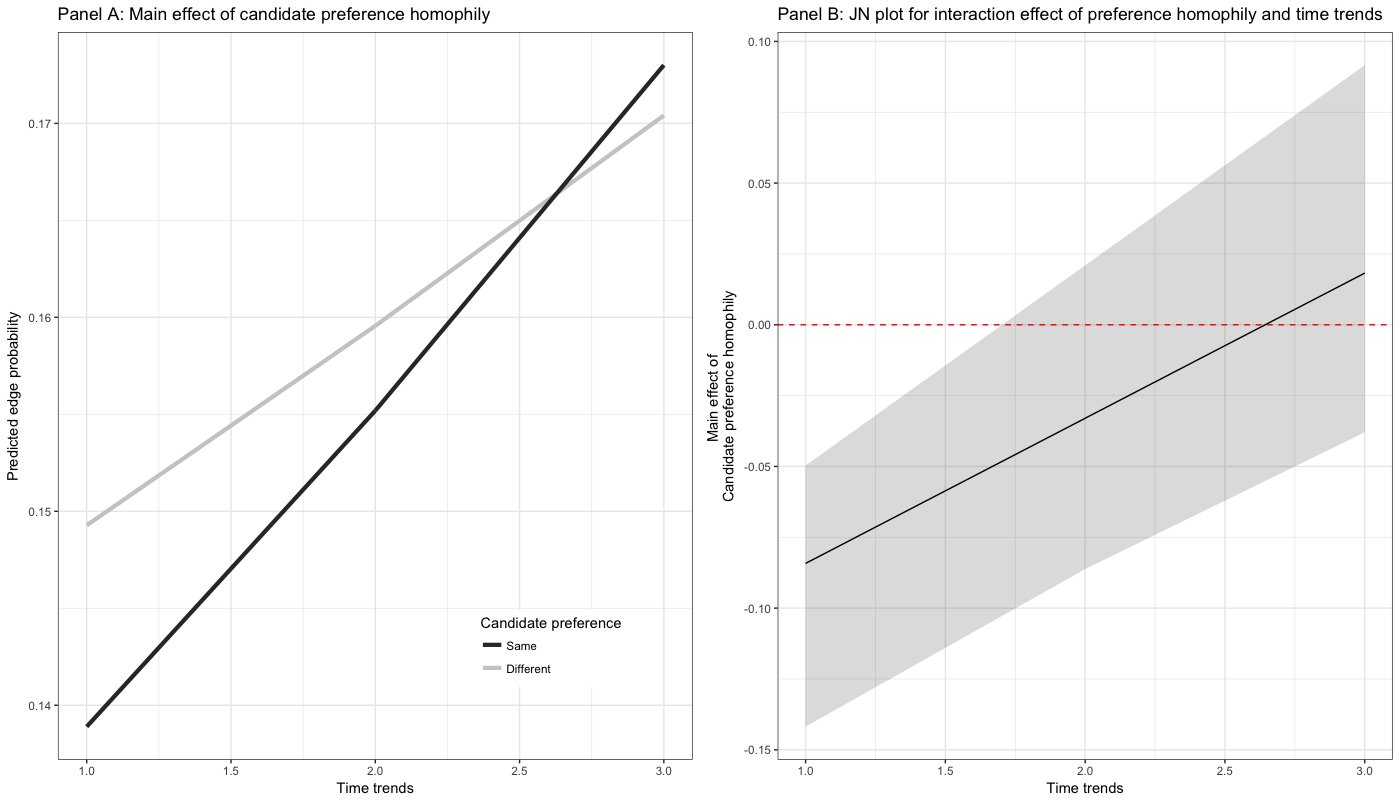


Figure 3. Interaction effects between time trends and candidate preference homophily.

1. ERGM framework requires all key nodal covariates to be complete in estimation, therefore any missing values on nodal covariates are not allowed. To address the missingness in our data, we have also estimated an identical model with multiple imputation technique on candidate preference (imputation N = 5), yet the substantial conclusion has not been changed by the inclusion or exclusion of those 22 missing cases. [↑](#endnote-ref-1)
2. Since participants’ key characteristics such as candidate evaluations and preferences were rather highly stable across survey waves (mean correlations across waves = .61 to .89), we regard participants’ characteristics may drive the creation of network ties, but not the other way around. This also assumes that participants’ characteristics are relatively hard to be changed within such a short period of time as a function of mere message exposure online, which is rather a standard assumption to make (Lazer, 2001). [↑](#endnote-ref-2)
3. Instead of partitioning the behavioral log data into three-wave panel survey dates (t = 3), we also estimated models with daily slices of log data (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 results suggest that our results and conclusions are reasonably robust against potential model misspecification and methodological issues in construction of the networks. [↑](#endnote-ref-3)
4. Items include: “I am competent at presenting my own opinions in a discussion,” “I can express my ideas in a coherent manner,” “I make full use of my subject knowledge in a discussion,” and “I feel competent persuading others in a discussion.” [↑](#endnote-ref-4)